MOLECULAR SIMULATION OF PROTEINS TO UNDERSTAND DYNAMICS AND ITS RELATION WITH FUNCTION

Thesis by

Neshatul Haque

In Partial Fulfillment of the Requirements
For the Degree of
Doctor of Philosophy

SCHOOL OF LIFE SCIENCE
DEPARTMENT OF BIOTECHNOLOGY
AND BIOINFORMATICS
UNIVERSITY OF HYDERABAD
TELANGANA, INDIA

FEBRUARY, 2017

Dedicated to my wife

Sana Parveen

DEPARTMENT OF BIOTECHNOLOGY AND BIOINFORMATICS
SCHOOL OF LIFE SCIENCE, UNIVERSITY OF HYDERABAD,
500046, HYDERABAD, INDIA

CERTIFICATE

This is to certify that thesis entitled "**Molecular Simulation of Proteins to Understand Dynamics and its Relation with Function"** submitted to the University of Hyderabad by **Mr. Neshatul Haque** for the degree of Doctor of Philosophy, is based on the studies carried out by him under my supervision. This work has not been submitted before for the award of degree or diploma from any University or Institution.

Dr. N. Prakash Prabhu
Supervisor

Head
Department of Biotechnology and Bioinformatics

Dean
School of Life Science

CERTIFICATE

This is to certify that the thesis entitled **"Molecular Simulation of Proteins to Understand Dynamics and its Relation with Function"** Submitted by **Neshatul Haque** bearing registration number **11LTPH03** in partial fulfilment of the requirements for award of Doctor of Philosophy in the school of **Life Science** is a bonafide work carried out by him under my supervision and guidance.

This thesis is free from plagiarism and has not been submitted previously in part or in full to this or any other university or institute for award of any degree or diploma.

Part of this thesis has been:

A. Published in the following publications:

1. Lid closure dynamics of porcine pancreatic lipase in aqueous solution N. Haque, N.P. Prabhu, Biochimica et Biophysica Acta (BBA)-General Subjects 1860 (10), 2313-2325	**ISSN 0304-4165**	**Chapter 2 of Thesis**
2. Lid dynamics of porcine pancreatic lipase in non-aqueous solvents N. Haque, N.P. Prabhu, Biochimica et Biophysica Acta (BBA)-General Subjects 1860 (10), 2326-2334	**ISSN 0304-4165**	**Chapter 3 of Thesis**
3. Insights into protein–TNS (2-p-toluidinylnaphthalene-6-sulfonate) interaction using molecular dynamics simulation N. Haque, N.P. Prabhu, Journal of Molecular Structure 1068, 261-269	**ISSN 0022-2860**	**Chapter 5 of Thesis**

B. Published in the following conference:

1._____, (National/International)

2._____, (National/International)

Further, the student has passed the following courses towards fulfilment of coursework requirement for PhD degree was award.

	Course code	Name	Credits	Pass/Fail
1	BT801	Analytical technique	4	Pass
2	BT802	Res. ethics, Stats. & Data analysis	3	Pass
3	BT803	Lab work and seminar	5	Pass

Dr. N. Prakash Prabhu

Supervisor

Head of Department of Biotechnology and Bioinformatics

Dean of School of Life Science

Acknowledgement

It is a pleasure to acknowledge the freedom, support, encouragement, cooperation and acute scientific criticism by Dr. N. Prakash Prabhu during the course of my work in his laboratory. I consider it as a privilege and honor to learn from him, the academic and social responsibility. I am also very thankful to my doctoral committee members, Dr. I. A. Qureshi and Dr. V. Vindal for their precious time and suggestions.

I would like to express my gratitude to Prof. A. S. Raghavendra (former Dean), Prof. P. Reddanna (Dean, School of Life Science), Prof. P. Prakash Babu, Prof. Niyaz Ahmed (former Head) and Prof. A. Kumar Kondapi (Head of department of Biotechnology and Bioinformatics) to provide necessary facilities required during these years. I am also very thankful and indebted to my teachers and faculties in the school for their honest and gentle way of teaching the subject, from the day of my masters to PhD. I would also like to thank school's staffs, Mr. Rajshekhar, Mr. Rahul, Mr. Shekhar, Mr. Krishana, Mr. Murthy, Mr. Rajesh (bioinformatics facility) and others for their invaluable support. It will be indispensable to acknowledge the financial support provided by CSIR (JRF/SRF), without which it was impossible to complete my PhD with tranquility.

I feel very lucky to have a pleasant working environment with our lab group members, Tejaswi, Kiran and Bramhini. I am eminently thankful to Tejaswi for the healthy scientific discussions which helped a lot in shaping my scientific ideas and thoughts. I am also very thankful to Krishnakanth for helping me in developing my scientific programming skill. I want to thank former and present lab members, Chaitanay, Suraj, Hyma, Samshul, Deepak, Pinki, Audre, Krishnakanth, Sagar, Manikyaprabhu, Swathi, Dhanyasree, Sharath, Abhilash, Sreevidya, Subhasree, Archi, Haider, for making my stay

in the lab happening and eventful. I am also indebted to my seniors and friends, Farhan, Khubaib, Arif, Mujahid, Majid, Suhail, Savita, Ashutosh, Nishant, Shivendra, Narendra, Sai, Arshad, Tanvi, Sravni, in the school of life science for their moral support when I extremely needed them.

I want to pay special homage and gratitude to late Chandramouli Malleda who helped me in the very early stage of the computation work. He was extremely gentle and modest person I have ever seen. May his soul rest in peace.

It would be unvirtuous, if I do not acknowledge my seniors and batch mates from other departments of the university. I would like to express my heartfelt gratitude to all my friends, Moinuddin, Qasim, Zulquarnain, Zikrullah, Taufeeq, Sajjad, Raees, Meer, Lateef, Danish, Imran, Akhtar, Shareef, Hamid and many more, to teach me social and personal responsibility by words and sometime by action.

I am very much thankful to my parents and other family members for having faith in me and supporting me. It is their blessings and their faith in almighty ALLAH, which always charged me with confidence. I am also very much thankful to my wife for all the love and care which she has showered on me to protect me from any kind of tension and depression during these years. I also want to acknowledge her for valuable personal, social and scientific suggestions and discussions.

The most important of all I would like to thank almighty ALLAH for his kindness, benevolence and beneficence without which this long journey was absolutely impossible.

Neshatul Haque

Table of Content

Chapter 1

Molecular Dynamic Simulation:

A Short Account of Theoretical Science

"...if we were to name the most powerful assumption of all, which leads one on and on in an attempt to understand life, it is that all things are made of atoms, and that everything that living things do can be understood in terms of the jigglings and wigglings of atoms."

Richard Feynman

Understanding of science always made scientists optimistic about calculating every property of a molecule by qualitative observation and more precisely by empirical or semi-empirical equations. The barrier of approximation is always minimized by advances in the knowledge of science. The science of understanding physical world could be said to be complete, only if, the phenomenon like interstellar motion (of stars, planets and moons) and the motion of sub-atomic particles (electrons, protons and neutrons) could be explained by one single equation, no matter how long the equation may be.

1.1 History of Theoretical Treatment of the Molecules

The use of Newton's laws of motion in the study of atoms and molecules can be traced back to 1930s. Vibrational spectroscopists used Hook's Law to treat the force holding two atoms in a molecule. Prof. Donald H. Andrew [1] assembled a set of steel balls connected by springs (Fig. 1), and calculated their natural frequencies by small perturbation in the assembly. He concludes his investigation by saying,

> *"It seems reasonable to conclude that the molecular forces are directed along the lines associated with the chemical bonds and that mechanically, they are analogous to spiral springs. It appears also that the great majority of Raman lines have values of the frequency closely corresponding to fundamental mechanical frequencies in the molecule. ……. The mechanical model should prove very valuable in making these identifications and in ascertaining the structure of the molecules."* [1]

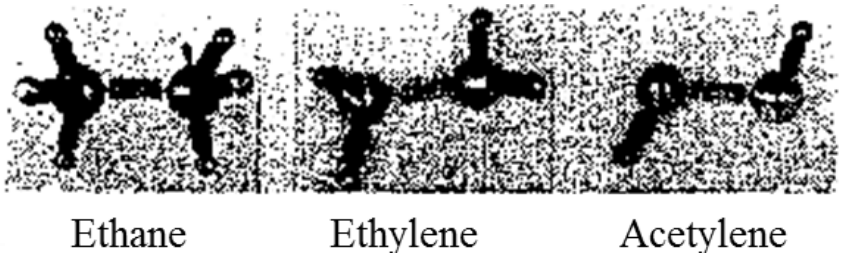


Figure 1. Steel ball and spring model of Prof. D. H. Andrew. Adapted from [1]

Prof. Terrell L. Hill [2] proposed that if two atoms in a molecule come very close to each other (less than their van der Waals radii), a repulsive force comes in origin (steric effect), and the molecule may undergo stretching or bending of the bonds. He formulated the quantitative calculation of the steric energy, and presented a very naïve exposition to geometry and energy of the molecules. A similar, but independent work was carried out by Prof. F. H. Westheimer [3] on the rate of racemization of sterically hindered optically active diphenyl derivatives. Similar approach of mechanics was also applied by Prof. I. Dostrovsky [4] to show the role of steric hindrance in the bimolecular nucleophilic substitution reaction (S_N^2 reaction) in neopentyl halides.

Although the foundation of theoretical calculations in the chemical world was laid down in mid of 20th century, yet there was a long way to go. The massive barrier in the development of the field of theoretical chemistry was manual and approximate calculation of the complicated equations.

1.2 Earliest use of Computing Machine

The journey of more accurate and precise calculation was only possible in the beginning of 1950s when the electronic computers were made available to the scientists. Development and testing of Monte Carlo methods [5], was among the earliest scientific work performed by Los Alamos Maniac I (**M**athematical **A**nalyzer, **N**umerical **I**ntegrator, **a**nd **C**omputer). Although the Monte Carlo method was developed for nuclear reaction simulation, it has been proved to be one of the robust algorithms in all areas of science. Prof. John G. Kirkwood in 1950 [6] calculated the excess entropy and equation of states by solving complicated integral equations using International Business Machine (IBM). In an attempt to deal with statistical mechanical problem, Monte Carlo method was implemented on a two dimensional system of hard spheres, interacting with each other by simple interaction potential [7] as well as Lenard-Jones type interaction potential [8]. Within statistical error limit, the spheres were observed to exert acceptable pressure, in a given volume and temperature. No evidence of phase transition was observed [7]. E. Fermi and co-workers [9], simulated a system of one-dimensional coupled masses (64 particles) attached together by springs, and numerically solved the non-linear equations of motion on Los Alamos MANIAC I. The numerical integration of Newton's equations of motion of the coupled system of masses in Fermi's experiment gave hope to solve the "many-body problem".

1.3 History of Molecular Dynamics Simulation

1.3.1 Two Body Interaction Potential

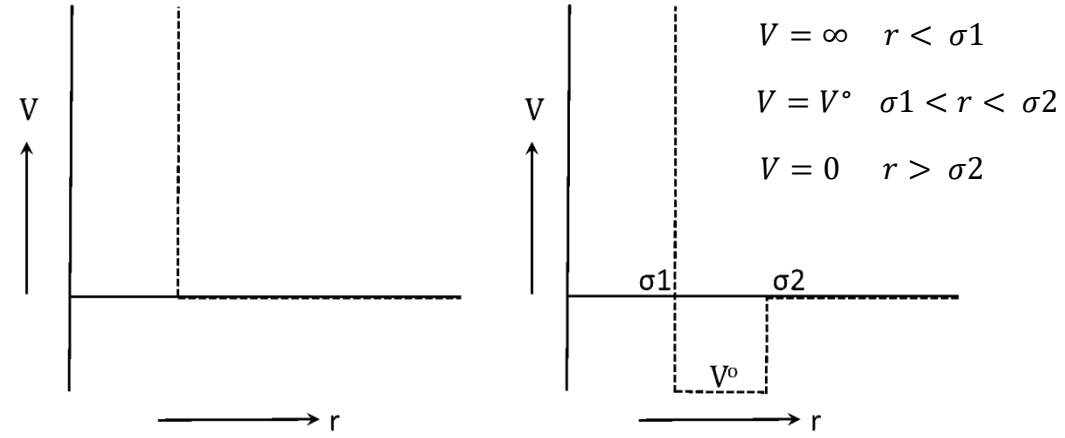


Figure 2. Hard sphere potential (left), square well potential (right), V is the potential energy and r is the separation between the spheres [10, 11].

With the advent of fast computing machines, it became possible to handle intensive calculations of multiple simultaneous equations. Following molecular dynamics was challenging though interesting topic for the scientists. Simplified model with reasonable approximation was required to formulate the scheme of calculation. Molecular dynamics simulation for the first time on a condensed phase model was performed using "hard sphere potential" (Fig. 2) [10, 11] with the nature of only collision interaction (if the two particles are at a distance of their van der Waals radii) to understand the phase transition. It was further improved by defining a "square well potential" (Fig. 2) for more realistic calculation [12]. A detailed scheme of the protocol was published with convincing result of the calculations showing two different densities in equilibrium (phase transition).

Among the earliest contributors in developing non-bonded interaction potential function, the crystallographer Prof. A. I. Kitaigorodosky is well renowned. He considered the additivity of the energy terms and transferability of the parameters within similar class of molecules. He refined the constants of the bond stretch and angle bending by studying the crystals of small organic molecules [13, 14]. In 1960, G. H. Vineyard group at Brookhaven National Laboratory studied the damage of metallic copper by radiation using the machine calculation. Born-Mayer type interaction potential was defined as central repulsive force for two-body interaction. In equilibrium configuration the cohesive and the repulsive forces just balances each other. The approximate interaction potentials could not give the exact understanding of real molecular interactions, but proved very helpful in dealing with the real system [11, 13, 15, 16]. Prof. Aneesur Rahman at Argonne National Laboratory simulated a system of 864 particles of liquid argon interacting with Lennard-Jones potential (continuous interaction potential) [17] for better understanding of the interacting classical particle in large number. He could achieve a convincing agreement of the calculated values with that of experimental values in terms of self-diffusion coefficient and pair correlation function. Liquid water, for the first time, was simulated by him in 1971 [18]. The methodological contribution in molecular dynamic simulation using molecular mechanics methods by these scientists was remarkable and paved the way for more realistic simulations along with the parallel evolution of computers.

1.3.2 Interaction Potential for More Than Two Atoms

Prof. James B. Hendrickson performed the computational study of the organic molecules [16]. He computed the conformational energy of the hydrocarbon molecules as the function of their geometrical parameters. Keeping bond length constant he could show the dependence of angle bending, torsion angle and van der Waals interaction on the conformational energy. Prof. N. L. Allinger also computed the conformational energy of the hydrocarbons but introduced the 'bond stretching' parameter in the empirical energy function for the molecules like ketones, saturated hydrocarbons and olefins [19-21]. A relation between the force constant and bond order was obtained by quantum mechanical

calculations in a delocalized electron system of hydrocarbons. By defining the standard way to handle the delocalized system, Allinger extended the molecular mechanics method to determine the structure of the molecules in a more generalized way, with the name "MMI force field". After the development of MMI [22], a series of updated and sophisticated versions were launched as MM2, MM3, and MM4 [23-25]. Allinger deposited his molecular mechanics program, MMI/MMPI, to the Quantum chemistry exchange program (QCPE 318) in November 1976 [26]. Allinger's work was well acknowledged by Edward M. Engler in 1973 [27]. Engler compared his force field results with that of Allinger's and found that most of the results were reproducible as already shown by experimental methods. Finally he expresses his endorsement to the molecular mechanics method refined by Allinger's as:

> "*We now support Allinger's conclusion that the molecular mechanics method, in principle, must be considered to be competitive with experimental determination of the structures and enthalpies of molecules."*

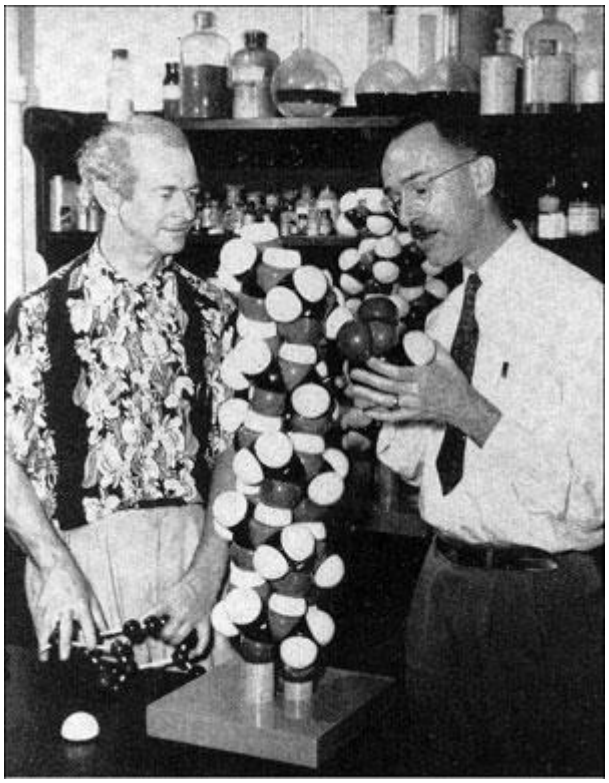


Figure 3. Pauling and Corey with their model of the alpha helix, 1951. File photo from pnas.org [28]

1.4 Studies with Proteins

Three dimensional structure of protein was first deciphered by X-ray diffraction of myoglobin crystals in Davy Faraday Laboratory in 1958 [15] by Prof. John Kendrew. In 1962 Prof. J. Kendrew and Prof. M. Perutz shared the Nobel Prize in chemistry for the atomic structures of protein. Before the great discovery of protein structure by X-ray crystallography, Linus Pauling and Robert Corey (Fig. 3), together modelled the 3D structure of the α-helix and β-sheets of polypeptides [29, 30]. It was a great breakthrough in the science to have the knowledge of structure of a protein at atomistic detail which opened an entirely new field of research. The understanding of protein structure helped us in many ways to fight against diseases and drug discovery research [31-35]. It became inevitable for the scientists to investigate the dynamic behavior of atoms in such a large macromolecular assembly of covalently bonded amino acid residues of proteins. The processes of protein folding was investigated using simplified model where fine

structural details were averaged to represent the protein [36]. This approximation was crucial to minimize the calculation burden. The protein folding model could fold back to near-native structure endorsing the idea that extended structure collapses first and then local non-covalent interactions shapes the 3D structure. Also, it was proposed that the loss of entropy is compensated by the gain of enthalpy. Prof Harold A. Scheraga in 1975 wrote a computer program, ECEPP (Empirical Conformational Energy Program for Peptides) [37] which calculated the conformation of the most favored potential energy of polypeptides and proteins. Prof. J. A. McCammon in 1977 performed probably the first molecular dynamics study of a protein, bovine pancreatic trypsin inhibitor at atomistic detail (except hydrogens which were merged with respective heavy atoms) [38]. This was the first report of its type, where Newton's laws of motion were numerically integrated to calculate the force and energy of every atom. This calculation was run for 9.2 ps without considering any solvent effect. The local atom fluctuations were shown to have diffusional character referring to fluid-like interior of the protein. It was further shown by the X-ray crystallographic experiments, that proteins attain same conformation with marginal changes in flexible regions of the protein represented by thermal or b-factor [39, 40] suggesting their dynamic nature. Later, Prof. S. Lifson's group was working on the generalization of the force field which could be used across the molecular family of distinct structure and properties, and came up with idea of Consistent Force Field (CFF) method [41].

> *"CFF method uses all relevant and available experimental data of whole families of molecules in order to select the best potential energy functions and to determine their constant parameters by a least-squares procedure, such as to obtain a best fit to the experimental data."*

Lifson and co-workers, in a stepwise manner, optimized the intermolecular and intramolecular interaction in alkanes [42, 43], then in alkenes [44] and then in amide crystals and hydrogen bonds of amides [45, 46]. Their parameter optimization with intermolecular, intramolecular and hydrogen bond interactions in amide were further useful in writing efficient energy functions for proteins.

Peter A. Kollman and co-workers in 1981 developed a force field and computer program, called "AMBER" (Assisted Model Building with Energy Refinement) [47] which was optimized for a wide variety of molecules including proteins and nucleic acid. The complete description of the scheme of the parameter refinement were published later in 1984 [48]. Martin Karplus and co-workers in 1983 came up with another computer program and their own force field, "CHARMM" (Chemistry at HARvard Macromolecular Mechanics). van Gunsteren and co-workers developed "GROMOS" (Groningen molecular simulation), a program for molecular simulation [49, 50] and Jorgensen group developed OPLS force field in 1988 [51]. Out of many molecular simulation engines "GROMACS" (Groningen MAchine for Chemical Simulation)

became most popular because of its open source nature [52]. It was developed by H. J. C. Berendsen, and David van der Spoel from GROMOS team. It did not have any force field of its own, but was compatible with all the popularly available force fields like CHARMM, AMBER, GROMOS, OPLS and their variants.

History of molecular dynamics simulation has witnessed a great evolution since 1950s till date. The hybrid classical and quantum mechanical treatment to the biomolecular systems for simulating the bond breaking and bond formation was first performed by A. Warshel and M. Levitt [53]. Nowadays QM/MM simulation could be relatively easily performed using many of the above simulation packages. This branch of science reached its zenith in 2013 when Prof. Michel Levitt, Prof. Martin Karplus and Prof. Arieh Warshel were jointly awarded Nobel Prize for "The development of multiscale models for complex chemical systems" [54]. The computational cost of the all atom simulation is always very enormous. So it becomes a limitation to investigate the large macromolecules or macromolecular assembly with sufficient sampling. Robert E. Rudd and Jeremy Q. Broughton in 1998, formulated a scheme, CGMD (Coarse Grained Molecular Dynamics) [55] simulation. CGMD became very popular and has been used for a number of simulation study of large system [56-58].

Till date, numerous simulation studies of small molecules, proteins, nucleic acid, and their complexes have been performed [59-64]. These studies have presented many interesting results such that DNA has huge flexibility and can be stretched to double of its size [65-67], quadruplex DNA (parallel d[G]4) could be very stable when a chain of Na^+ ions were also simulated in the central channel, and in the absence of the chain of ions the complex get distorted but regains the structure when ions were added again [68, 69]. Moreover, molecular dynamic (MD) simulation has been continuously used to understand the processes of protein folding and associated metastable and transition states [70-72]. Direct calculation of transition state using molecular mechanics is nearly impossible. By theory, transition state of the protein is a state which can be achieved with the probability of 0.5 from either of the meta-stable states. To make that happen in the MD calculation, large numbers of simulation should be started from the conformation of the folding trajectory till the fraction of the trajectory which goes to the native state before going to the unfolded states is equal to 0.5 [73, 74]. This is a time consuming process and computationally expensive. According to statistical mechanics, only a small fraction of all trajectories is statistically relevant. Various schemes have been proposed to sample the transition state, like milestoning, bias annealing, dominant reaction pathways [75-79] etc. However, further studies in this area is necessary to improve the accuracy of calculations.

1.5 Molecular Dynamics Simulation in Modern Science

MD simulation is proved to be a useful tool in the modern day science. There has been an exponential rise in publication of research articles on the use and development of

MD simulation (Fig. 4). Availability of cheaper and efficient computers along with the open source MD simulation softwares, has greatly contributed to this exponential rise. It has potential to answer those problems for which present day experimental techniques has limited scope. In the material science, MD simulation has been used to predict the properties of alloys [80] , calculate the thermal conductivity of nano-assemblies [81], study the behavior of materials in different pH and ionic solutions [82], understand the heat transfer properties of nanotubes [83], interaction of polymer-nanoparticle composite materials [84], study the deformation behavior of the nano-crystalline metals [85, 86] etc. In the chemical and biochemical science, MD simulation has been used for investigating huge range of scientific questions. Researchers all around the world are using MD simulation to understand the problems ranging from,bond-breaking and bond formation in simple organic and enzymatic reactions [87-91] to functional role of gigantic molecular assemblies like nucleic acid-protein complexes [64, 92-94]. However, there is much more diversity in the use of MD simulations in the biochemical science like studies of protein dynamics [60, 95], protein folding barriers and kinetics [70-79], protein-protein interactions [96-98], protein-drug interactions [32, 96-98], protein-membrane interactions [99, 100], drug-membrane-interaction [101-103], membrane dynamics [104-106], voltage-gated ion channels [107, 108], amyloid fibril formation [109, 110] etc.

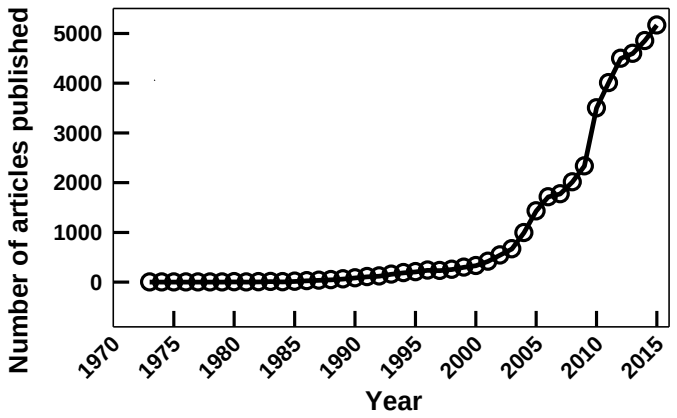


Figure 4. Data is obtained from medline trend with the search term "molecular dynamic simulation"

This thesis also largely adapts MD simulation methods to understand the dynamic nature of protein segments and their role in activity of the protein along with the effect of solvents on the dynamics of these segments. Also, molecular assembly of surfactants and their effect on the protein dynamics and function has also been examined. Further, self-assembly of small molecular dye and formation of protein-small molecular complexes has been investigated with atomistic detail using MD simulation.

Chapter 2

Lid Closure Dynamics of Porcine Pancreatic Lipase in Aqueous Solution

2.1 Abstract

Pancreatic lipases hydrolyse fatty acids in dietary pathway. The activity of porcine pancreatic lipase (PPL) is controlled by a lid domain along with a coenzyme, colipase. The active open-state conformation of the protein could be induced by detergents or bile salts which would be further facilitated by binding of colipase. In the absence of these interactions, the lid preferably attains a closed conformation in water. Molecular dynamic simulation was used to monitor the lid movement of PPL in open and closed conformations in water. Free energy surface was constructed from the simulation. Energy barriers and major structural changes during lid opening were evaluated. The lid closure of PPL in water from its open conformation might be initiated by columbic interactions which initially moves the lid away from domain 1. This is followed by major dihedral changes on the lid residues which alter the trajectory of motion. The lid then swirls back towards domain 1 to attain closed conformation. This is accompanied with the conformational changes around β5- and β9-loops as well. However, PPL in closed conformation shows only the domain movements and the lid remains in its closed conformation. PPL in closed conformation is stable in water and the open conformation is driven towards closed state.

2.2 Introduction

Lipases [EC 3.1.1.3] are a subclass of esterases isolated from various species and share a well-conserved α/β-hydrolase fold [111-113]. Lipases hydrolyse naturally occurring oils of plant seeds and fats of animals which on further oxidation release huge amount of their stored energy [114]. Lipases are activated at oil-water interface to access water insoluble substrates for hydrolysis [115-118]. Like other esterases, lipases also follow the well-characterized acyl enzyme pathway for hydrolysis reaction [117]. The activity of lipases is reported in a range of solvents from polar to nonpolar, such as ethanol, heptane, and octanol [118-123]. Lipases break (hydrolysis) [115-118], and make (esterification) [119-123] ester bonds in polar and non-polar solvents, respectively. Due to their ability to catalyze the organic reactions such as hydrolysis, acidolysis, aminolysis, esterification, and transesterification, lipases are widely used in food, pharmaceuticals, leather, and cosmetics industries [124, 125].

Pancreatic lipases are crucial enzymes for the digestion of dietary long chain triglycerides. Upon digestion in the duodenum these bulky non-polar molecules are broken down into small polar β-monoglycerides and fatty acids. These polar fractions are then easily transited to the enterocytes [126]. Optimal activity of the pancreatic lipases could be obtained in the presence of a coenzyme known as colipases [127]. Surfactants such as sodium deoxycholate and sodium taurodeoxycholate, at above their critical micelle concentration (CMC) inhibit the activity of pancreatic lipases which can be restored in the presence of colipases [128-130]. Colipases are highly homologous across

mammals and a colipase from one mammal can activate the lipase of another mammal [131].

Porcine pancreatic lipase (PPL) is a 50 kDa protein, with 448 amino acid residues. The crystal structure of PPL-colipase complex in the open state was resolved in 1996 by Hermoso *et al.* (PDB id: 1ETH) [132]. PPL consists of two domains: α/β-fold *N*-terminal domain (residues: 1-336) and β-sandwich *C*-terminal domain (residues: 337-448) named domain 1 and domain 2, respectively. The domains are linked through a hinge residue Ala333. The active site constituted by Ser153, Asp177, and His264 is covered by a flap or lid region stretched between Cys238 and Cys262. The conformational changes around the lid control the enzyme activity of the lipase. The open conformation of lid allows substrates to reach the active site residues which are otherwise not accessible in the closed conformation [133-137]. Open conformation of PPL could be stabilized by either ionic/non-ionic detergents or bile salts in aqueous solution. Addition of colipase could further enhance the stabilization of open conformation [132, 138]. Colipase is held with the lipase through non-covalent interactions across the lid and *C*-terminal domain [132].

The structural stability and plasticity of the various lipases have been investigated by both experimental and computational approaches [121-123, 139-142]. Studies on the atomistic-level changes in lipases with respect to their structure, solvent environment, and substrate specificity are getting momentum in the recent years [139-144]. However, to the best of our knowledge, the lid dynamics of mammalian lipase has not been investigated at atomistic-detail. Our work attempts to identify the kinetic pathway and the possible barriers during the lid closure of PPL from its open conformation (PPL-O) in aqueous environment and to compare its dynamics with colipase bound PPL (PPL^{+colip}) and PPL in closed conformation (PPL-Cl). The results suggest that the open conformation is less stable in water, in the absence of colipase, and moves towards domain 1 to attain the closed conformation. The lid first moves away from domain 1 and later alters its trajectory and moves towards domain 1. However, the conformational fluctuations in PPL-Cl are mainly due to domain-domain movement, and the lid dynamics of PPL-Cl is highly constrained in water.

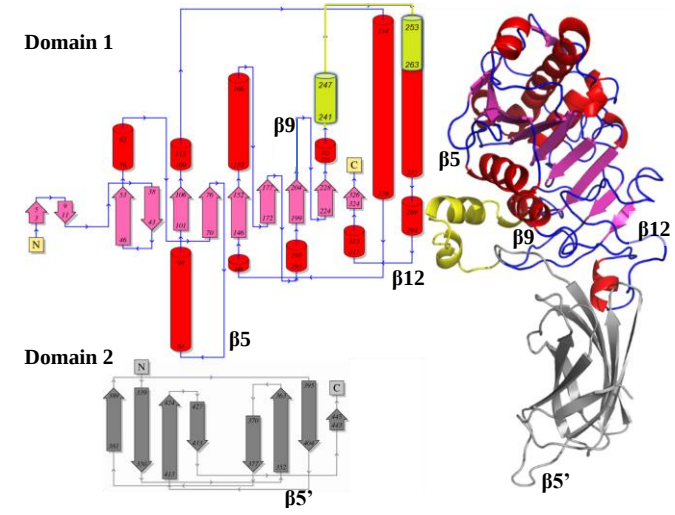


Figure 1. Topology diagram of PPL in open conformation without colipase (left panel). Rods represent α-helices and arrows represent β-strands. Lid region is marked in green and domain 2 in gray. Right-side panel presents the cartoon diagram of PPL-O with the same color codes as in topology diagram. Major β-loops are labeled.

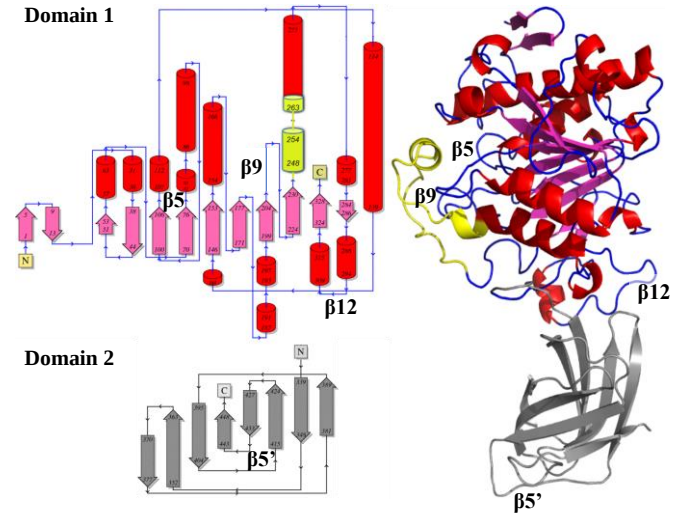


Figure 2. Topology diagram of PPL in closed conformation modeled using Modeler 9.1 (left panel). Rods represent α-helices and arrows represent β-strands. Lid region is marked in green and domain 2 in gray. Right-side panel presents cartoon diagram of PPL-Cl with the same color codes as in topology diagram. Major β-loops are labeled.

2.3 Experimental Methods

2.3.1 Structure of PPL in Open and Closed Conformations

The crystal structure of colipase bound PPL in open conformation was obtained from protein databank (PDB id: 1ETH). The structural coordinates of lipase-colipase complex (PPL^{+colip}) was extracted by removing other co-crystallized molecules. Further, colipase free PPL with open conformation was obtained from PPL^{+colip} by removing the structural coordinates of colipase (PPL-O) (Fig. 1). Since the crystal structure of closed conformation of PPL without colipase (PPL-Cl) is not available, it was modeled using Modeller 9.11[145] against rat pancreatic lipase-related protein 2 (RPLRP2) in closed conformation (PDB id: 1BU8) [146] as a template. RPLRP2 shows 84% of similarity and 99% of sequence coverage with PPL and the structure was resolved with 1.8Å resolution in the absence of colipase. The quality of the modeled structure was further validated by verify 3D [147] and procheck [148].

2.3.2 Molecular Dynamics (MD) Simulation

MD simulations of PPL-Cl, PPl-O, and PPL^{+colip} were carried out using GROMACS 4.5.4 [52, 149-151] with OPLS-AA force field. The protein was initially solvated in a cubic box with SPC/E water. The structures were energy minimized and equilibrated under NVT followed by NPT conditions. After the equilibration, the production simulations were carried out at 300 K for 50 ns. 1 nm cutoff was used for short range columbic and Lennard-Jones interactions. Long range electrostatic interactions were treated by Particle mesh Ewald (PME) scheme [152]. Stochastic velocity rescaling scheme with time constant of 0.1 ps was used for temperature control. Pressure was maintained at 1 bar by Parrinello-Rahman [153] type pressure coupling with a time constant of 2 ps. For all the bond constraints, LINCS algorithm [154] was used. Coordinates, velocities, trajectories, and energies were collected at every 2 ps for further analysis. GROMACS utility tools were used for analyzing the trajectories. For structure visualization and analyses PyMol [155] and UCSF-Chimera [156] were used. PDBsum web server [157] was used for creating 2D representations of the protein structures. PDB2PQR [158, 159] and APBS [160] were used to calculate solvation energies and electrostatic forces, respectively. For comparative analysis, the changes occurred only on the lipase chain (PPL^{-colip}) of colipase bound lipase were extracted from the simulation of $PPL^{+colipase}$.

2.4 Results

2.4.1 Modeling of PPL-Cl

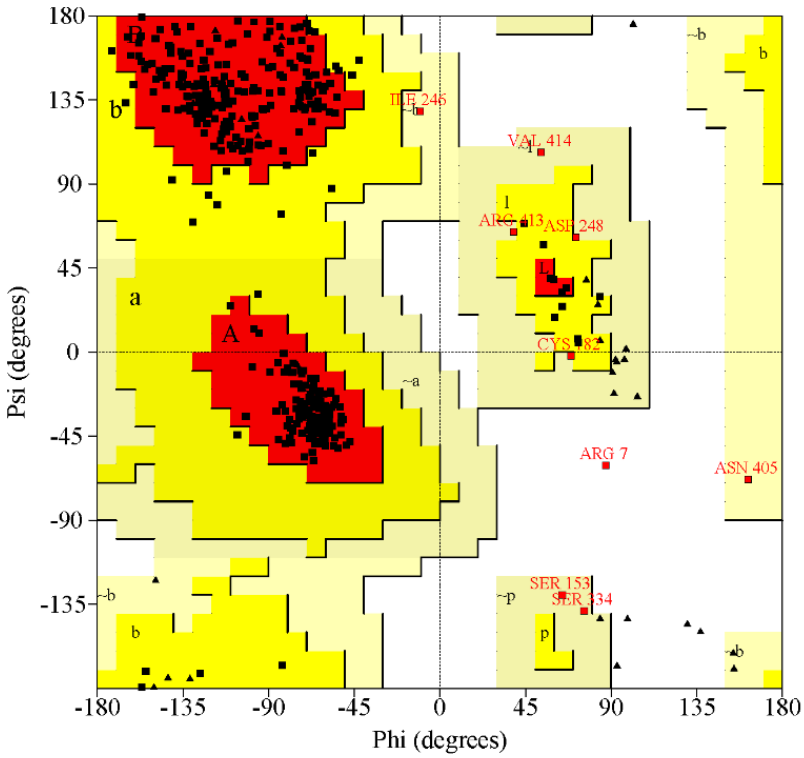
The structure of PPL in closed conformation (PPL-Cl) was modeled (Fig. 2.) using RPLRP 2 as a template. Ramachandran plot of the modeled structure showed only 0.3% of the residues in the disallowed region. Verify 3D results also suggested that the obtained model is acceptable (Fig. 3). The modeled structure was further energy minimized *in vacuo* using OPLS-AA force field. Major structural differences between PPL-O and PPL-Cl were observed around the lid region whereas the other regions of the proteins showed only minimal differences. In PPL-Cl, the lid remained closer to domain 1 and was characterized by a small helix flanked with long unstructured loops. The lid of PPL-O was well separated from both the domains and possessed two short helices connected through a loop. These two helices were further connected to the protein chain through short loops (Fig. 1). PPL in closed conformation lost some of the H-bonding and salt bridge interactions present in the open conformation of PPL as observed in human pancreatic lipase (HPL) [161]. Notably, the interactions of lid residues R257 with core residues D80 and Y268, and E254 with D258 of lid and K269 of core were lost in the closed form. However, D250 formed new H-bonding interactions with E254 and Y115. It was observed in horse pancreatic lipase that the interaction between E254 and Y115 is one of the stabilization factors of the lid in closed conformation [162]. Further, MD

simulation studies show that disruption of these lid residue interactions, E254 with D250 and Y115, by point mutations could initiate lid opening movements in PPL (Chapter 3).

2.4.2 Preliminary Trajectory Analysis

Root mean square deviations (RMSD) of C_α-atoms were calculated using gromacs utility tool. PPL-Cl and PPL-O showed considerable changes in their RMSD values whereas the lipase chain of colipase bound PPL (PPL^{-colip}) showed relatively less changes (Fig. 4A). A sharp increase in the RMSD of PPL-Cl was noted at above 25 ns whereas PPL-O showed two increasing transitions of RMSD at above 10 ns and 30 ns, respectively. The structural compactness of the proteins was monitored by following the changes in radius of gyration (R_g). The R_g of PPL-O and PPL^{-colip} did not show major changes whereas PPL-Cl showed a slight increase in R_g (Fig. 4B).

Root mean square fluctuation (RMSF), standard deviation of atomic positions during simulation, provides details about plasticity and rigidity of individual residues in the protein. From Fig. 4C, it was observed that the RMSF of residues 240-255 (corresponding to lid region) was more in the case of PPL-O compared to PPL^{-colip} which was higher than PPL-Cl. Further, the residues around 210-216 on domain1 (β9 loop) showed significantly higher RMSF for PPL^{-colip} and PPL-O compared to PPL-Cl. PPL-Cl showed higher fluctuations around the regions 340-360 and 365-380 in domain 2, 298-300 at the domain interface, and around 80-85 in domain 1 (β5 loop) compared to PPL-O and PPL^{-colip}. All the three proteins showed significant fluctuation of surface loop in domain 2 (β5′ loop) covering the residues 405-413. At the end of 50 ns simulation, notable changes were observed in all the three structures which are illustrated in Fig. 5. The overlay of initial and final structures of PPL-Cl obtained from the simulation showed only marginal changes in domain 1 and the lid whereas significant fluctuations were observed around domain 2. In the case of PPL-O, substantial changes were observed around the lid region and the trajectory of lid movement is represented by arrows (Fig. 5B and Fig. 6). A circular motion of the lid was observed with loss of secondary structure in the lid region (Fig. 7B). In PPL^{-colip}, notable changes were observed around the lid region and in domain 2 as well.

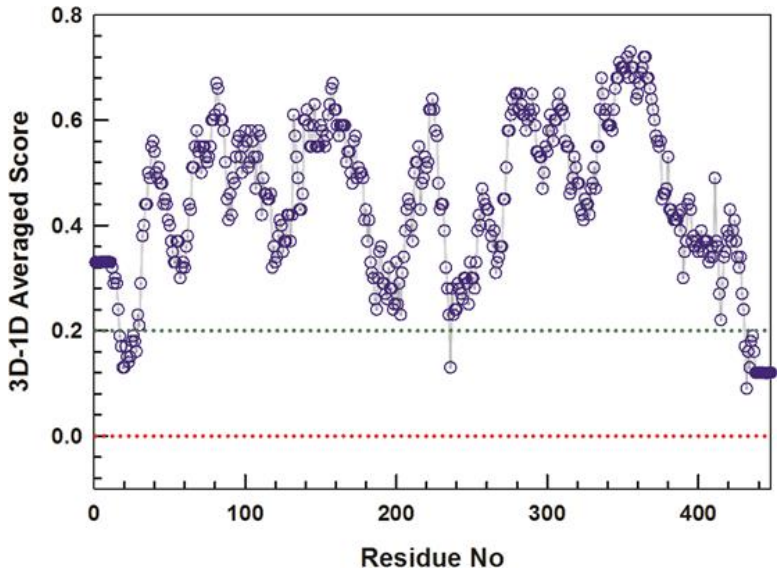


Plot statistics

Residues in most favoured regions [A,B,L]	**344**	**90.1%**
Residues in additional allowed regions[a,b,l,p]	**29**	**7.6%**
Residues in generously allowed regions[~a,~b,~l,~p]	**8**	**2.1%**
Residues in disallowed regions	**1**	**0.3%**
	------	------
Number of non-glycine and non-proline residues	**382**	**100%**
Number of end –residues (excl. Gly and Pro)	**2**	
Number of glycine residues (shown as triangles)	**39**	
Number of proline residues	**25**	

Total number of residues	**448**	

A



B

Figure 3. (A) Ramachandran plot of PPL-Cl modelled using Modeller 9.11. (B) 3D-1D average score of each residues of the modelled protein obtained from Verify 3D.

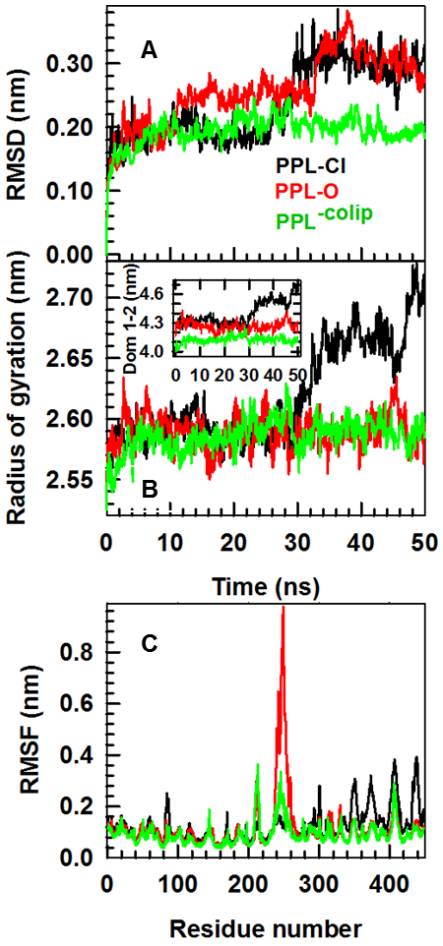


Figure 4. Preliminary trajectory analysis of PPL-Cl (black), PPL-O (red), and PPL^{-colip} (green): (A) RMSD from the corresponding initial structures, (B) change in radius of gyration during simulation and the inset presents change in distance between two domains, (C) RMS fluctuations of each residue in the lipase chain.

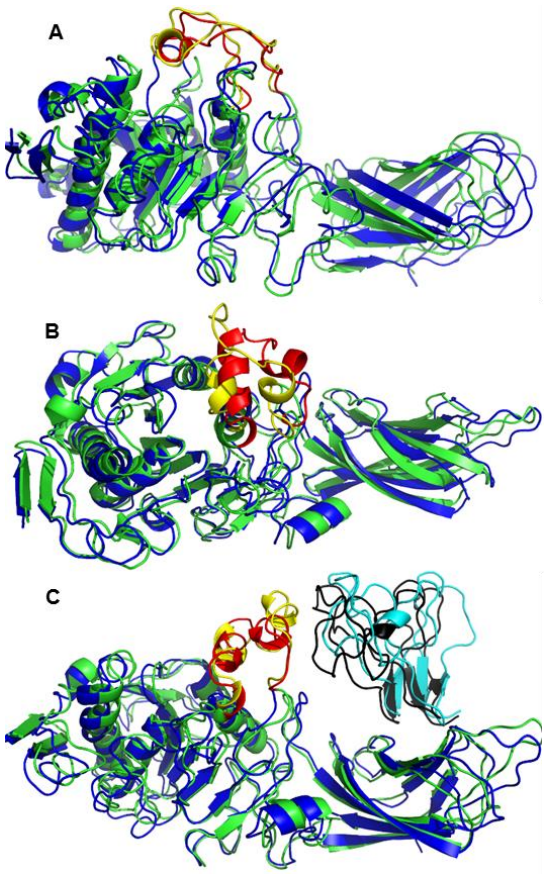


Figure 5. Comparison of initial and final structures of PPL in MD simulation: Green cartoons with lid in red color represent the initial conformation and blue cartoons with lid in yellow color represent the final conformation of (A) PPL-Cl, (B) PPL-O, and (C) PPL^{+colip}. The initial and final conformations of colipase are represented in black and cyan, respectively, in PPL^{+colip}. For more clarity on PPL-O lid movement, please refer Fig. 6 and video V2.

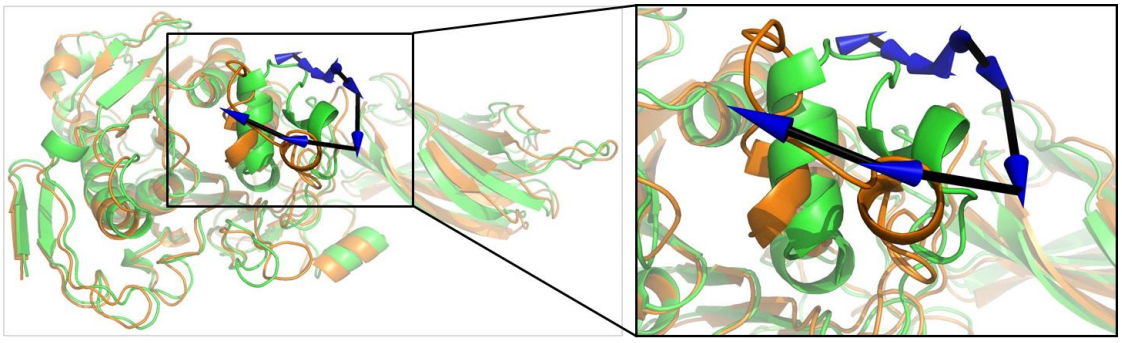


Figure 6. Comparison of initial and final structures of PPL-O in MD simulation: green cartoon represents the initial conformation and brown cartoon represents the final conformation with the lid trajectory represented by blue arrows. Right-side panel presents a closer view of the lid region.

The hydrophobic and hydrophilic solvent accessible surface area (SASA) of PPL-Cl and PPL-O were monitored during the simulations (Fig. 8). The hydrophobic SASA value of PPL-Cl increased over simulation time whereas PPL-O was decreasing. The hydrophilic SASA value of PPL-Cl remained almost same throughout the simulation except initial fluctuations. However, PPL-O showed slight increase in hydrophilic SASA after ~38 ns.

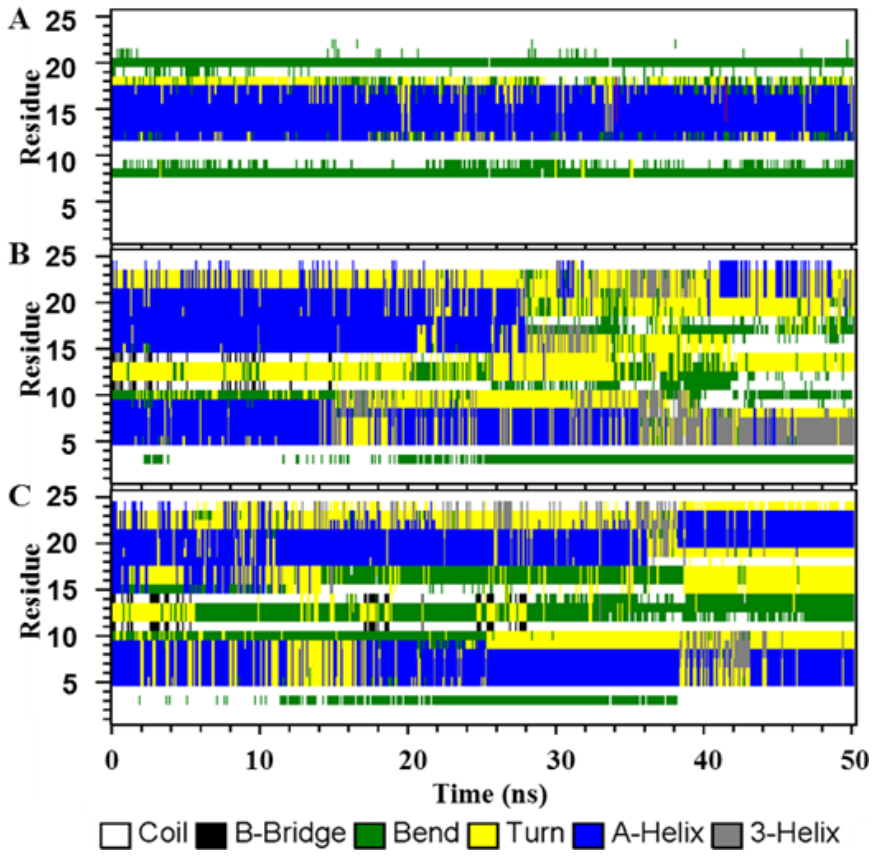


Figure 7. Change in secondary structure of the lid residues (238-262) in (A) PPL-Cl, (B) PPL-O and (C) PPL^{-colip} during the simulation. The secondary structural changes were calculated using DSSP program integrated with Gromacs.

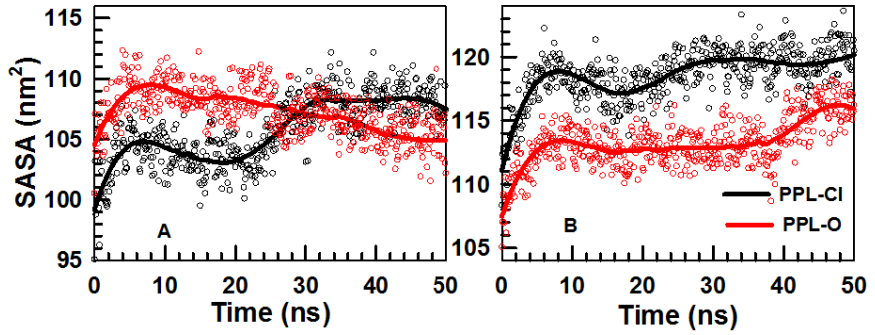


Figure 8. (A) Hydrophobic and (B) hydrophilic SASA changes of PPL-Cl (black) and PPL-O (red) during simulation. The solid lines are for visual clarity.

2.4.3 Electrostatic and Solvation Energy

The mean values of long range and short range electrostatic energies in PPL-Cl and PPL-O were calculated (Table 1). PPL-Cl showed more negative values than PPL-O suggesting that PPL-Cl is more favorable conformation in water. Two different components of solvation free energy changes (ΔG) were also calculated: polar (ΔG^{p}) and non-polar (ΔG^{np}). ΔG^{p} was calculated using APBS tool on the last 10 ns of simulation [163, 164]. ΔG^{np} was calculated by SASA approach, i.e., $\Delta G^{np} = \gamma \times (\text{SASA}) + \beta$, where γ and β are 2.2 kJ mol^{-1} nm^{-2} and 3.84 kJ mol^{-1}, respectively [165]. ΔG^{p} was significantly larger than ΔG^{np} in both the cases. Moreover, ΔG^{p} was more negative for PPL-Cl than PPL-O which reemphasizes that PPL-Cl is the favorable conformation in water.

Table 1. Mean electrostatic energy of PPL in different conformations

Energy	PPL-Cl (kJ/mol)		PPL-O (kJ/mol)	
	average	error	average	error
Coulomb (short-range)	-2203410.00	110.00	-2171590.00	50.00
Coulomb (long-range)	-261988.00	26.00	-259790.00	18.00
Polar Solvation (ΔG^{p})	-10146.34	18.86	-9862.06	15.18
Non-Polar Solvation (ΔG^{np})	505.67	0.42	489.94	0.48

2.4.4 Electrostatic Force on the Lid

To further investigate on the lid movement of PPL-O, the electrostatic forces on four selected charged residues in the lid, Asp248, Asp250, Glu254, and Asp258, were calculated using APBS tool. The trajectories of negatively charged side chain oxygen atoms of these residues were obtained from 500 frames selected at a constant interval of every 0.1 ns. The components of force and position vectors were plotted as a function of time (Fig. 9). Substantial amount of correlation was noted between the force and motion of these residues along *Y*- and *Z*- coordinates till 32 ns. When the net force was positive (repulsion) in *Y*-direction, the position vector increased. When the net force was negative (attraction) in *Z*-direction, the position vector decreased. The *X*-component of the force and position vectors also correlated with each other in the cases of Asp248 and Asp250. In all the cases, there was a sharp transition at around 32 ns following which no correlation between the force and position vectors was observed.

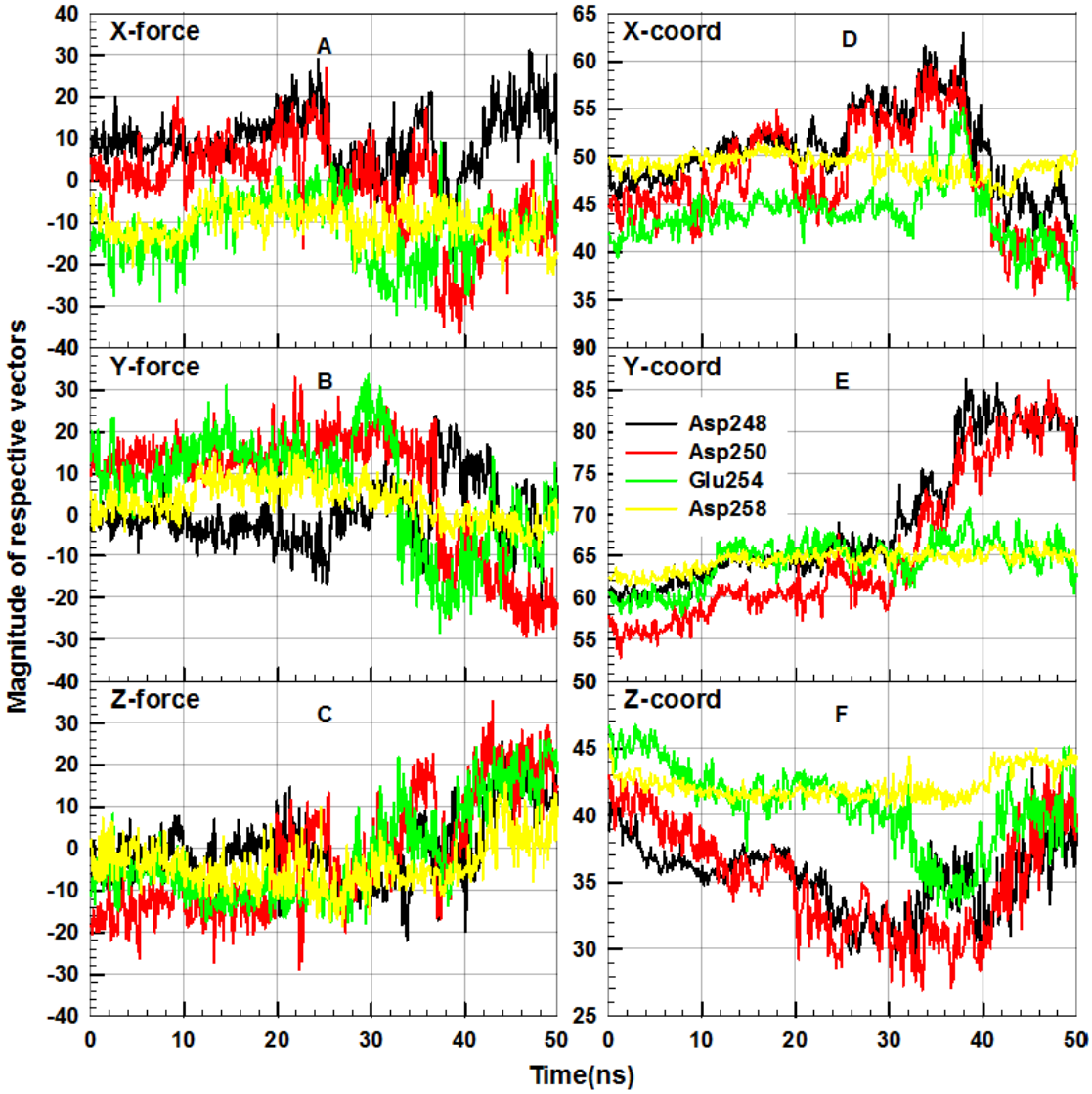


Figure 9. (A-C) Electrostatic force on the lid residues Asp248 (black), Asp250 (red), Glu254 (green), and Asp258 (yellow) along all three coordinates X, Y, and Z. (D-F) The magnitude of change in distance vectors of the four lid residues along the three coordinates.

2.4.5 Essential Dynamics

In order to understand the major dynamic events, principal component analysis (PCA) was done on both Cartesian (cPCA) and dihedral (dPCA) coordinates for all the three proteins. The cPCA and dPCA together could cover the essential dynamics of PPL. The covariance matrix of PPL-O (Fig. 10A and C) inferred that substantial amount of highly correlated motions exist around the lid residues. On the other hand, the covariance matrix of PPL-Cl (Fig. 10B and D) and PPL^{-colip} (Fig. 11) did not show much correlated motions. Further, 80% of the dynamics was covered by first 9, 6, and 21 eigenvectors (EVs) in PPL-Cl, PPL-O, and PPL^{-colip}, respectively.

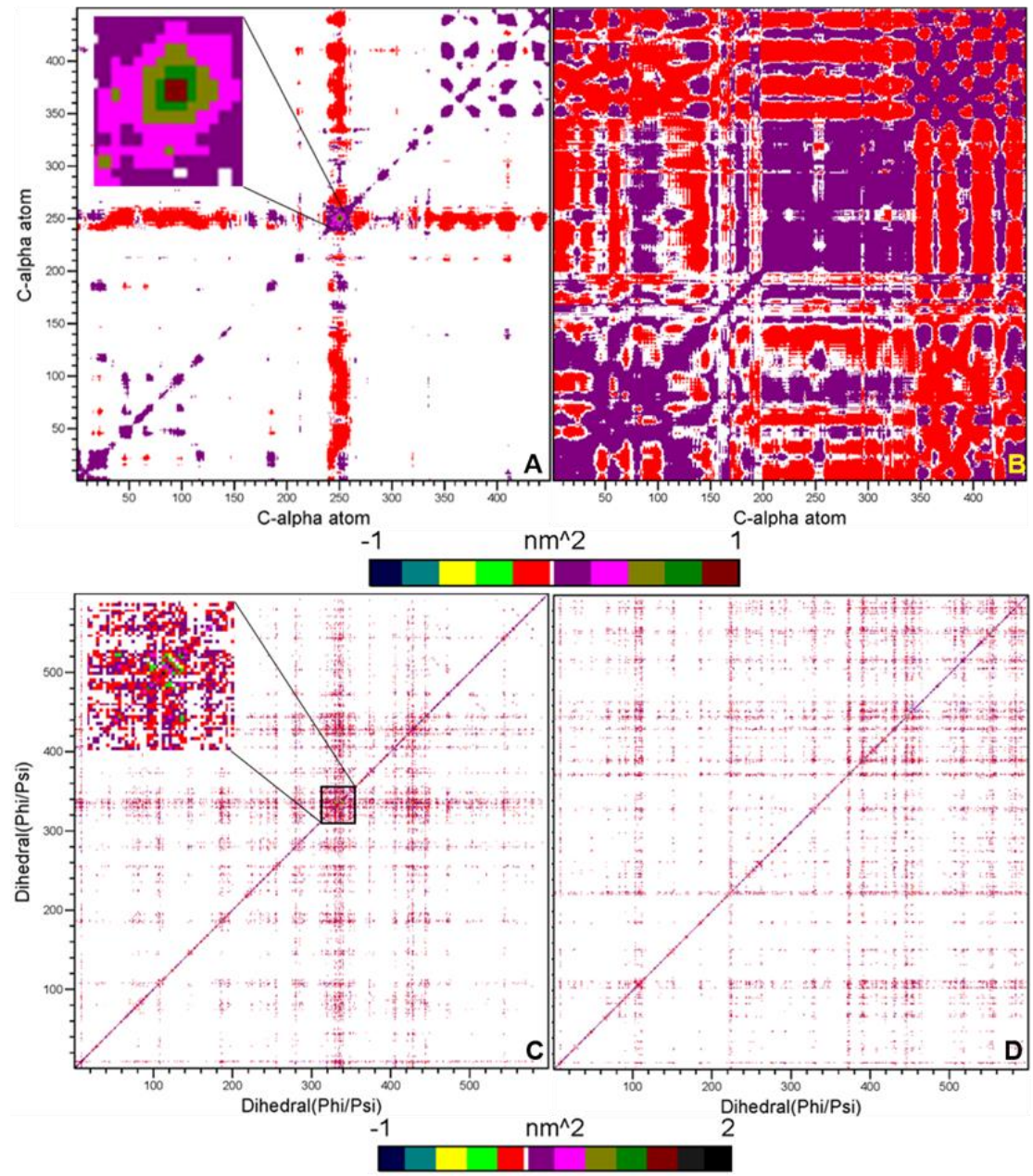


Figure 10. Covariance matrix of C_α motions derived from Cartesian coordinates PCA for PPL-O (A) and PPL-Cl (B). Covariance matrix of C_α motions derived from dihedral PCA for PPL-O (C) and PPL-Cl (D). The color codes represent the amplitude of variations from anti-correlated motions (negative) to correlated motions (positive) with white as null variation. The insets in panels A and C are focused on the correlated motions around the lid region of PPL-O.

The first 5 EVs could cover 74, 78 and 60% of the dynamics, respectively (Fig. 12). Two extreme conformations from the first two EVs are shown in the Fig. 13-15 for PPL-O, PPL-Cl, and PPL^{-colip}, respectively, to visually appreciate the conformational differences. The first EV components of each residues obtained from cPCA and dPCA is presented in Fig. 16 along with RMSF of each residues calculated from first EV subspace. The EV components of cPCA and dPCA showed slightly different fluctuations of residues in the first eigenspace. The detailed description of next four EV components,

RMSF changes, and their projections are presented in the Fig. 17 – 19. In all the cases, the higher EVs showed lower amplitude fluctuations than the first EV.

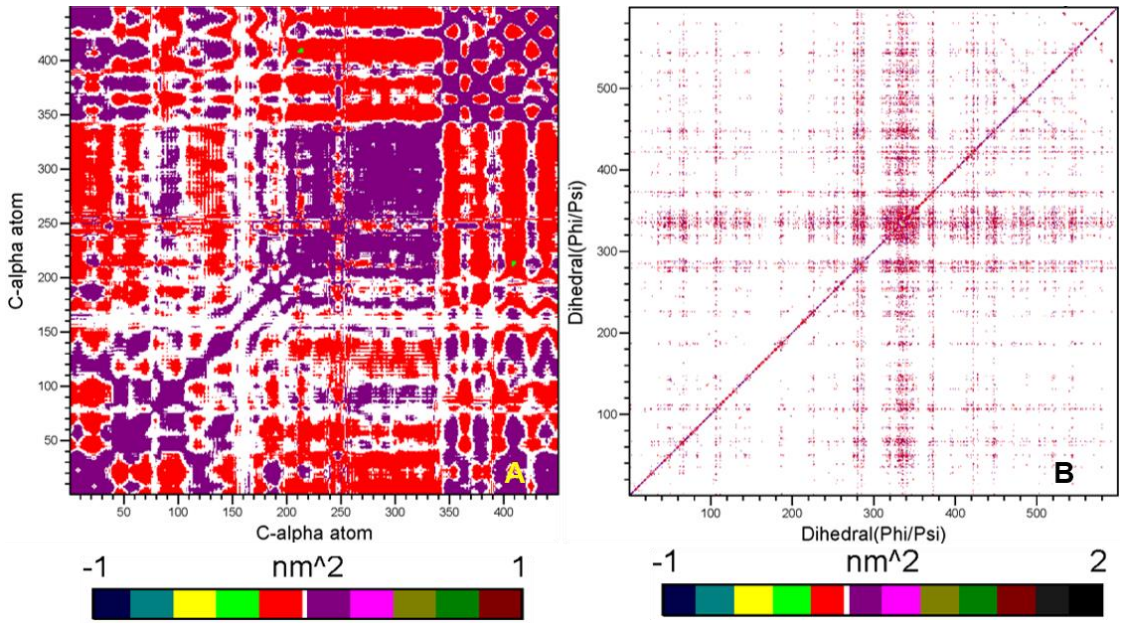


Figure 11. Covariance matrix of C_α motions in PPL-colip obtained from principle component analysis of Cartesian coordinates (A) and dihedral changes (B). The colors represent the amplitude of variations from anti-correlated motion (negative) to correlated motions (positive) with white as null variation.

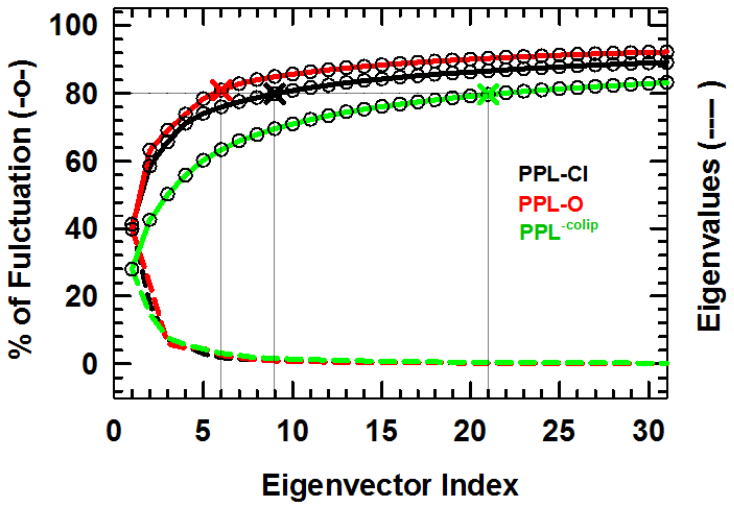


Figure 12. The cumulative percentage of dynamic changes covered (solid lines) and proportional eigenvalues (dotted lines) of each eigenvector during the simulation of PPL-Cl (black), PPL-O (red), and PPL-colip (green).

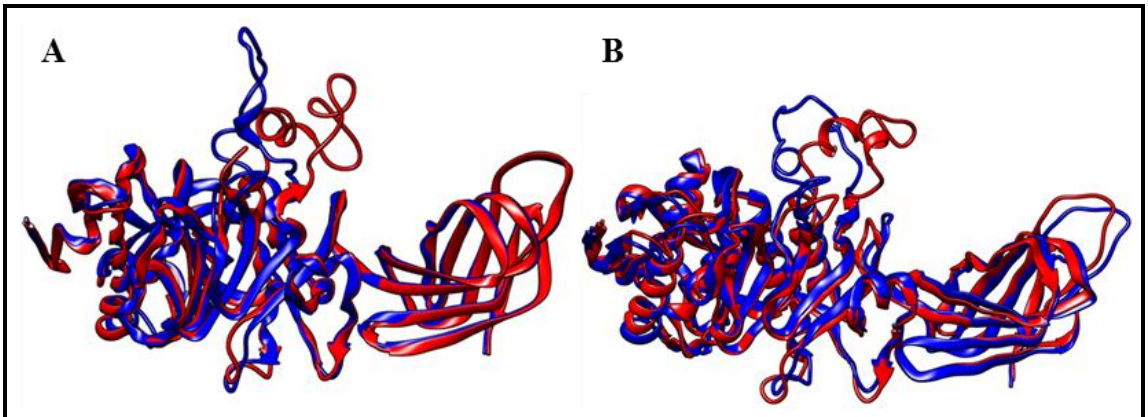


Figure 13. Comparison of two extreme conformations obtained from eigenvector 1 (A) and eigenvector 2 (B), derived from cPCA for PPL-O.

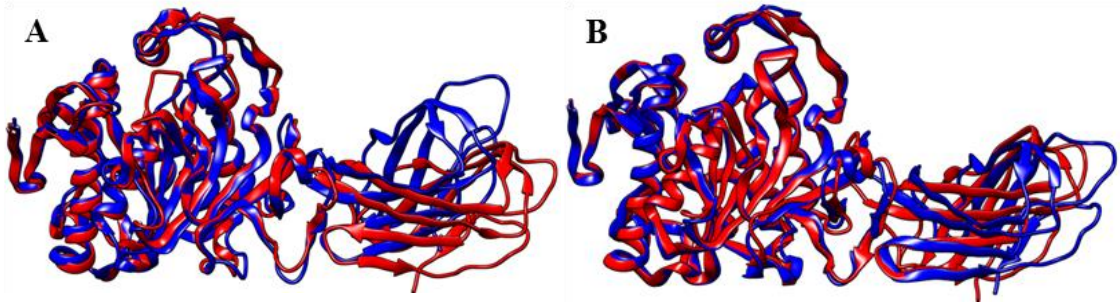


Figure 14. Comparison of two extreme conformations obtained from eigenvector 1 (A) and eigenvector 2 (B), derived from cPCA for PPL-Cl.

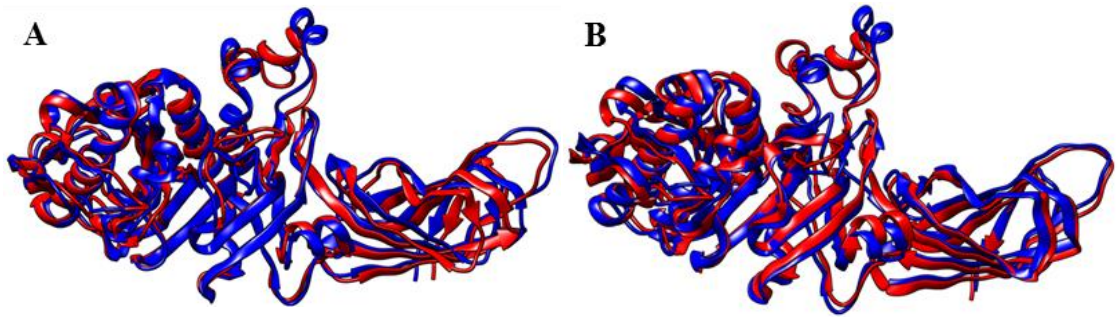
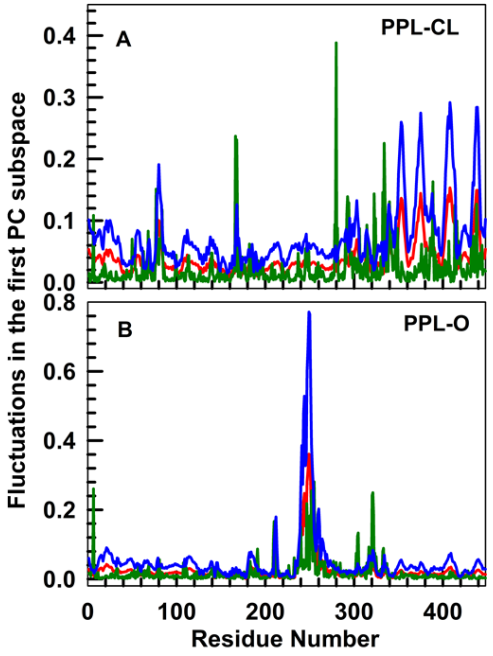


Figure 15. Comparison of two extreme conformations obtained from eigenvector 1 (A) and eigenvector 2 (B), derived from cPCA for PPL^{-colip}.



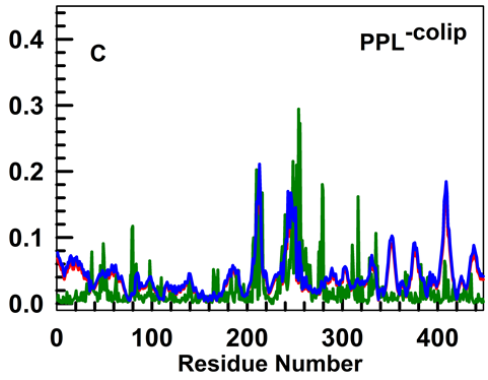
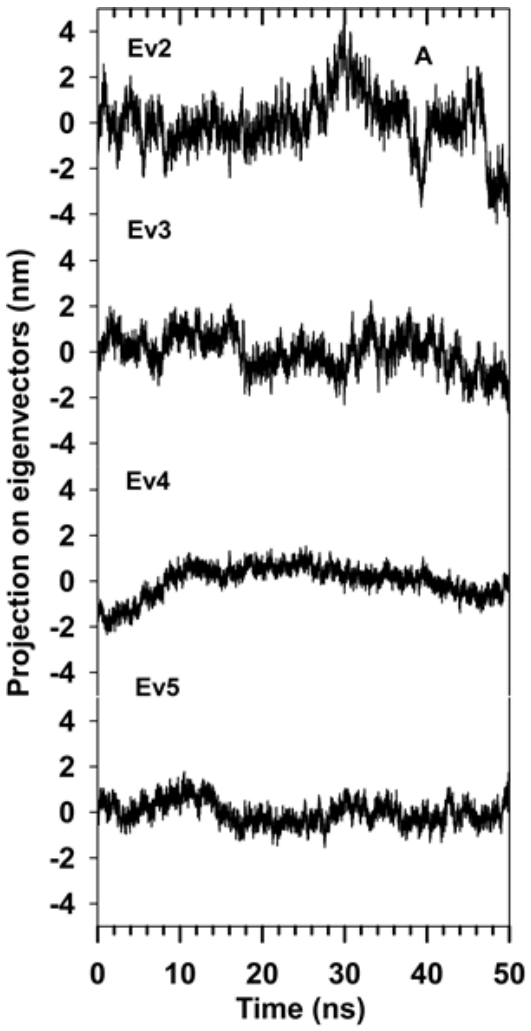


Figure 16. Residue-wise eigenvector components of (A) PPL-Cl, (B) PPL-O and (C) PPL^{-colip} for first eigenvalues of cPCA (red) and dPCA (green) drawn along with the corresponding RMSF values of EV1 (blue).



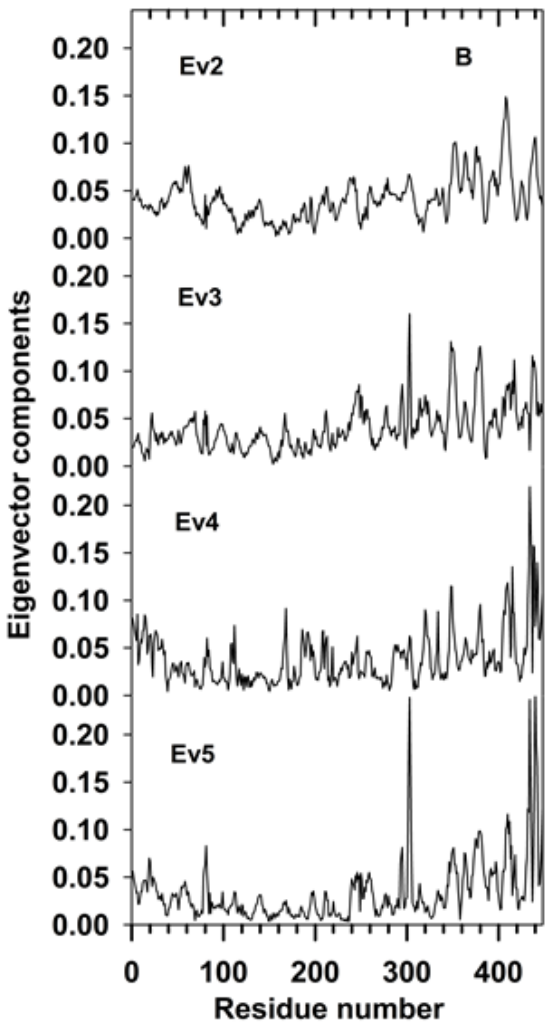


Figure 17. Projection of trajectory onto the Eigenvector (A) and Eigenvector component: weight of each C_α atom on the eigenvectors (B) Ev2-Ev5 of PPL-Cl.

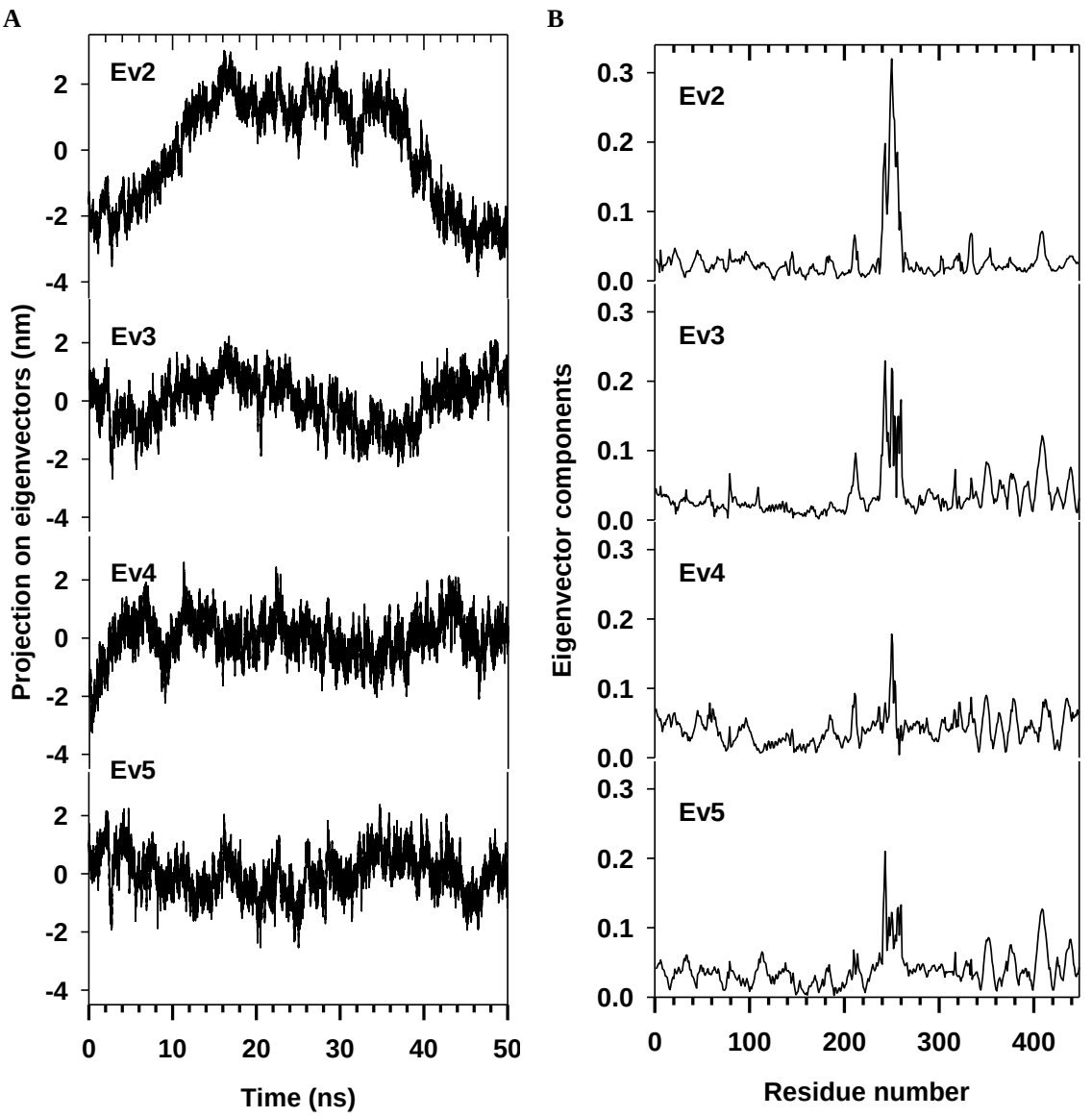


Figure 18. Projection of trajectory onto the Eigenvector (A) and Eigenvector component: weight of each C_α atom on the eigenvectors (B) Ev2-Ev5 of PPL-O.

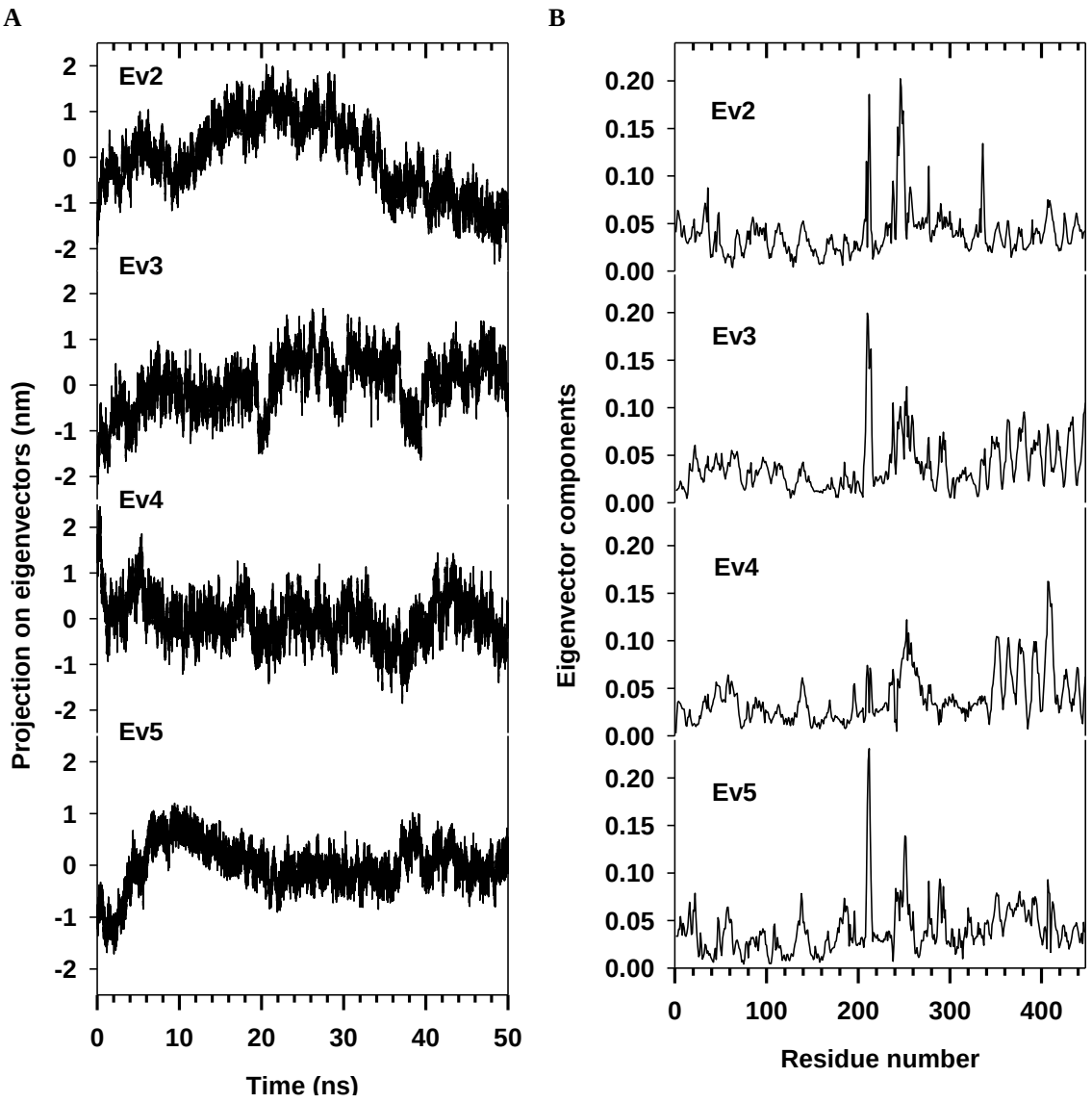


Figure 19. Projection of trajectory onto the Eigenvector (A) and Eigenvector component: weight of each C_α atom on the eigenvectors (B) Ev2-Ev5 of PPL^{-colip}.

2.4.6 Free Energy Surface of PPL-O

Free energy landscape was constructed by projection of the first EV of dPCA against the first EV of cPCA as reaction coordinates. The trajectory on the energy surface is shown in Fig. 20A. Though multiple transitions could be marked on the landscape, the most probable pathway with major transitions on the energy surface was considered for further analysis (Fig. 20B). The Gibbs energy of major transitions along the reaction coordinate is presented Fig. 20B (inset). During the transitions, large conformational changes were noted around the lid of the protein. The essential conformational transitions corresponding to the states with local minima which led to lid closing trajectory were denoted as S0, S1, S2, S3, and S4. The transition states between these local minima were marked as T1, T2, T3, and T4, respectively. The corresponding conformational states of these transition states on the simulation time scale were around 10, 31, 42, and 44 ns. The

largest barrier encountered in the trajectory was at T3 which was estimated to be ~14 kJ/mol. The local minima and the transition state structures captured along the reaction pathway are presented as a movie (video V2)

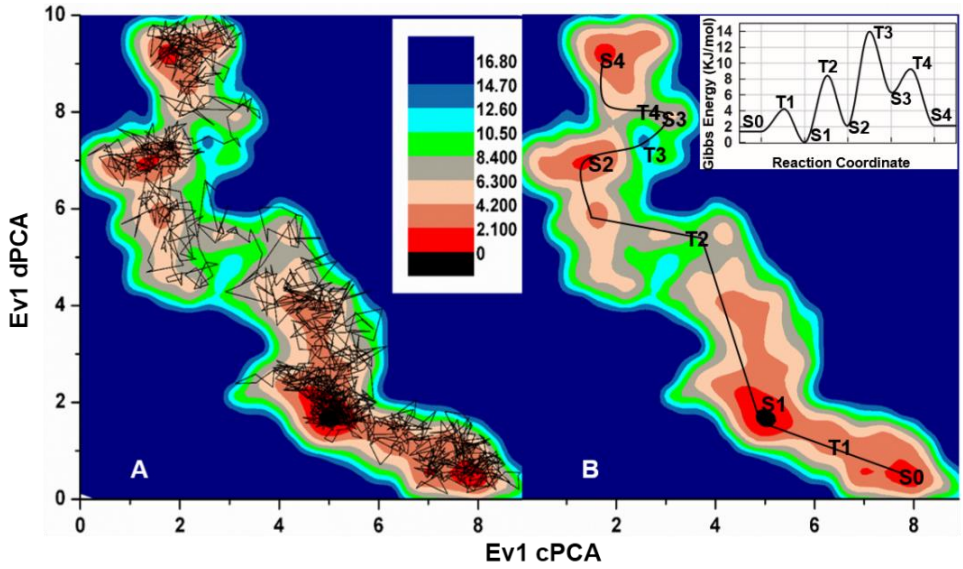


Figure 20. Free energy landscape of PPL-O constructed by projection of first eigenvectors of dPCA against cPCA with the trajectory of movements in black lines. (B) The most probable pathway with major transitions are chosen and the corresponding Gibbs energies of the transitions against reaction coordinate is shown in inset. The color codes represent the energy levels.

2.4.7 Radial Distribution Function

Radial distribution function (RDF) of the solvent molecules around the lid of PPL-O was calculated at an interval of each 1 ns by considering heavy atoms on the lid residues as reference against the oxygen of water. Relative water density around the lid in first hydration shell, i.e., at 0.274 nm, was plotted against simulation time (Fig. 21). The density did not show significant changes during the initial half of the simulation. However, a sharp increase was observed between 30 and 40 ns.

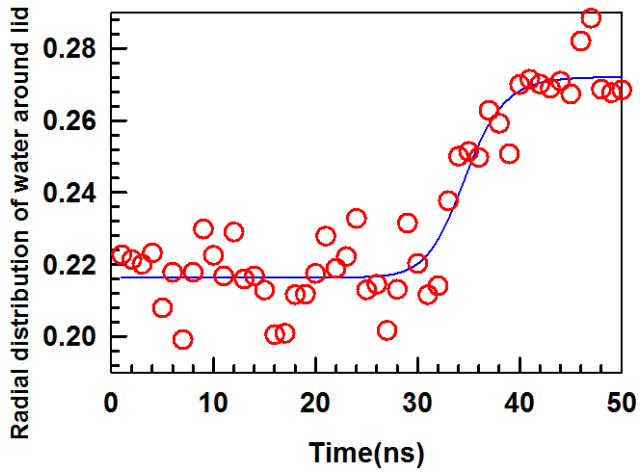


Figure 21. Change in radial distribution of water molecules around the lid region of PPL-O calculated at the distance of 0.274 nm during simulation.

2.4.8 Dihedral Angles

Based on the observations of EV components and RMSF in the first EV space, change in dihedral angles (phi and psi angles) of eighteen amino acid residues in and around the lid during the simulation were analyzed (Fig. 22 and 23). Dihedral angles of PPL-Cl showed fluctuations at the initial period of the simulation, but reached a stable phase almost in all the residues. High fluctuations were observed in PPL^{-colip} and in PPL-O. Particularly, Ile252, Trp253, Glu254, Gly255, Val260 and Cys262 (Fig. 22) showed clear transitions at the time scales corresponding to T1, T2, or T3 states observed in the free energy surface (Fig. 20B) for PPL-O whereas the other residues around the lid showed either early or late transitions (Fig. 23).

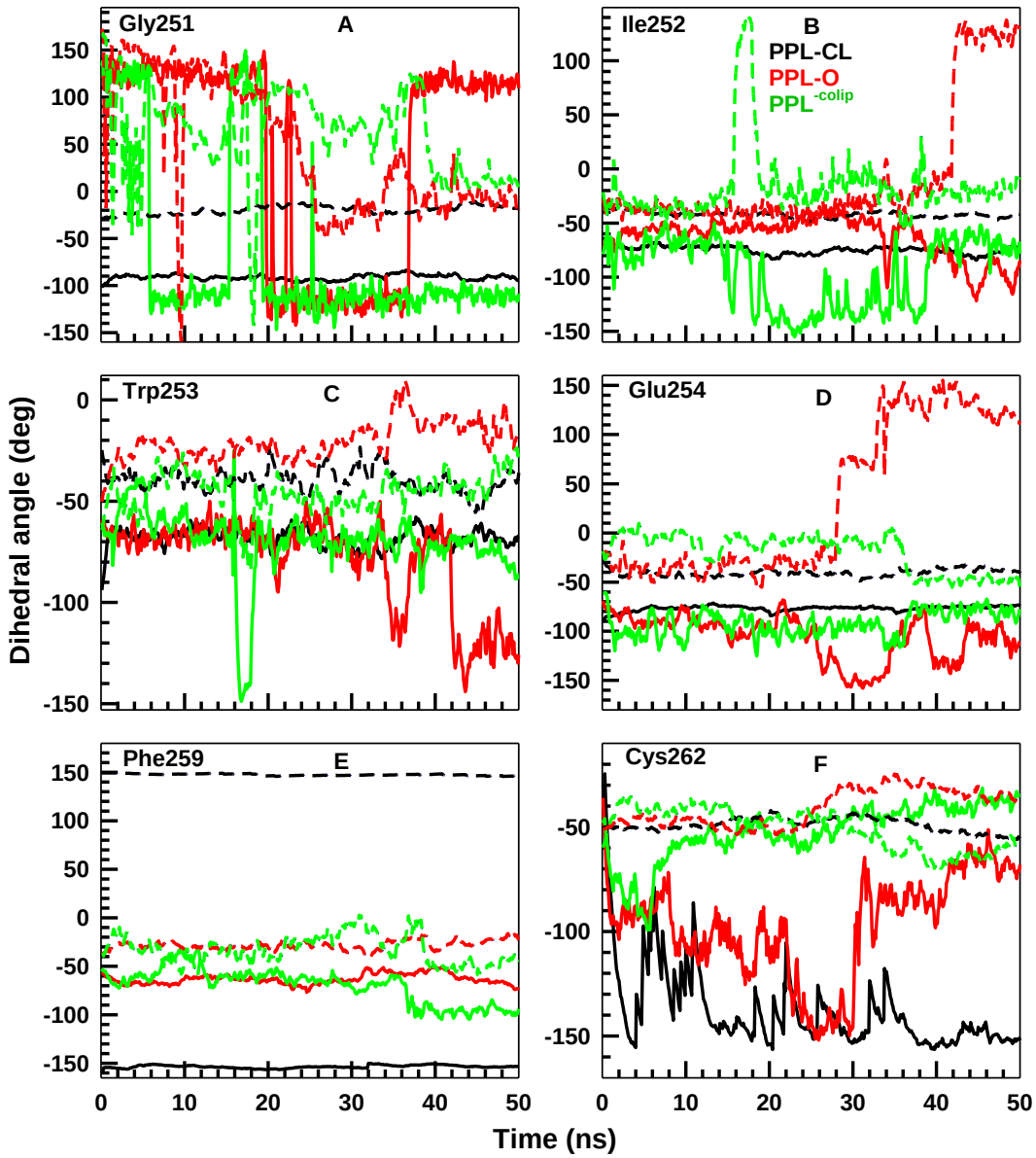
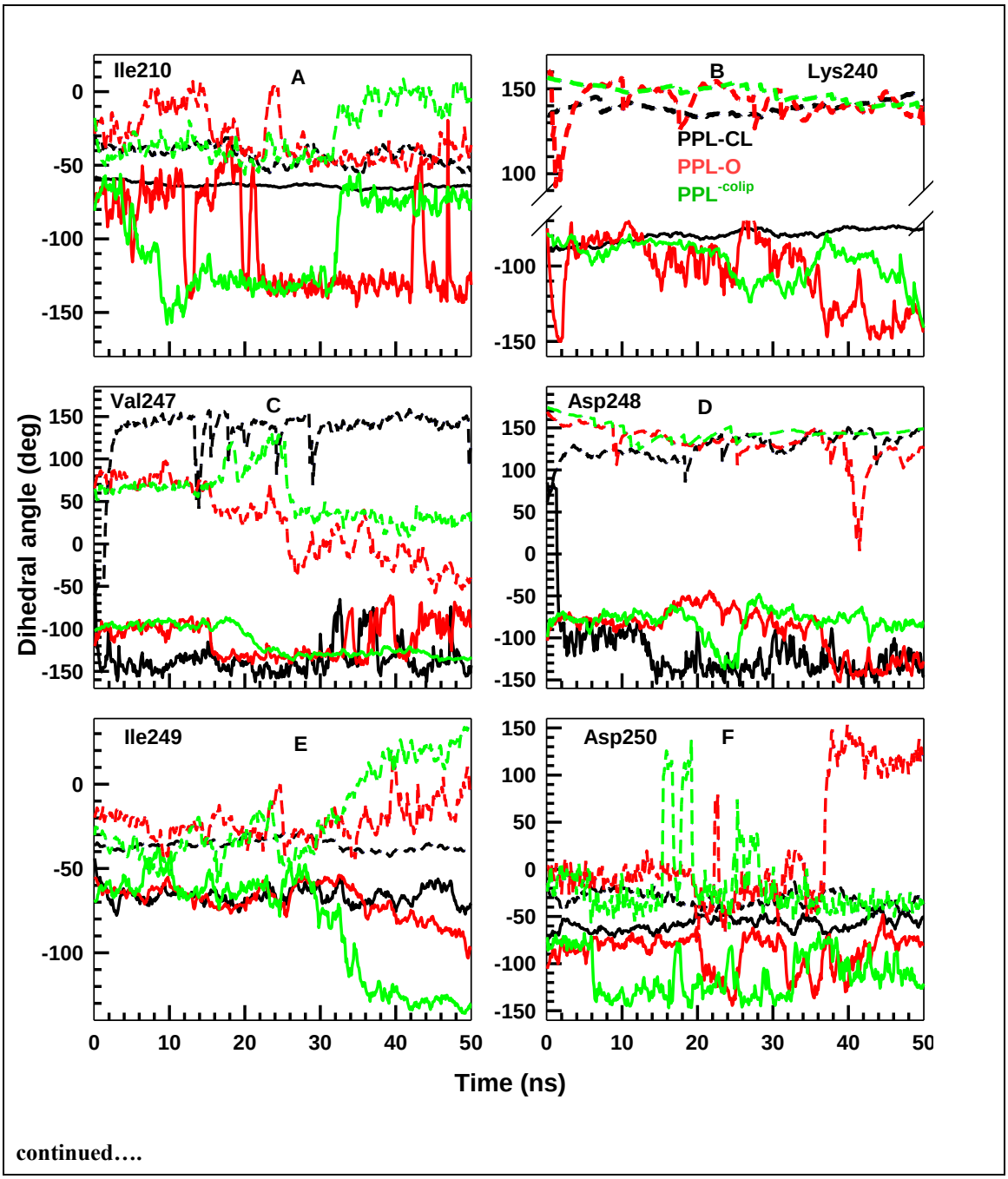


Figure 22. The changes in dihedral angles phi (solid lines) and psi (dotted lines) of selected residues in and around the lid domain for PPL-Cl (black), PPL-O (red), and PPL^{-colip} (green).


Ile210
A
B
Lys240
PPL-CL
PPL-O
PPL-colip
Val247
C
Asp248
D
Ile249
E
Asp250
F
Dihedral angle (deg)
Time (ns)

continued….

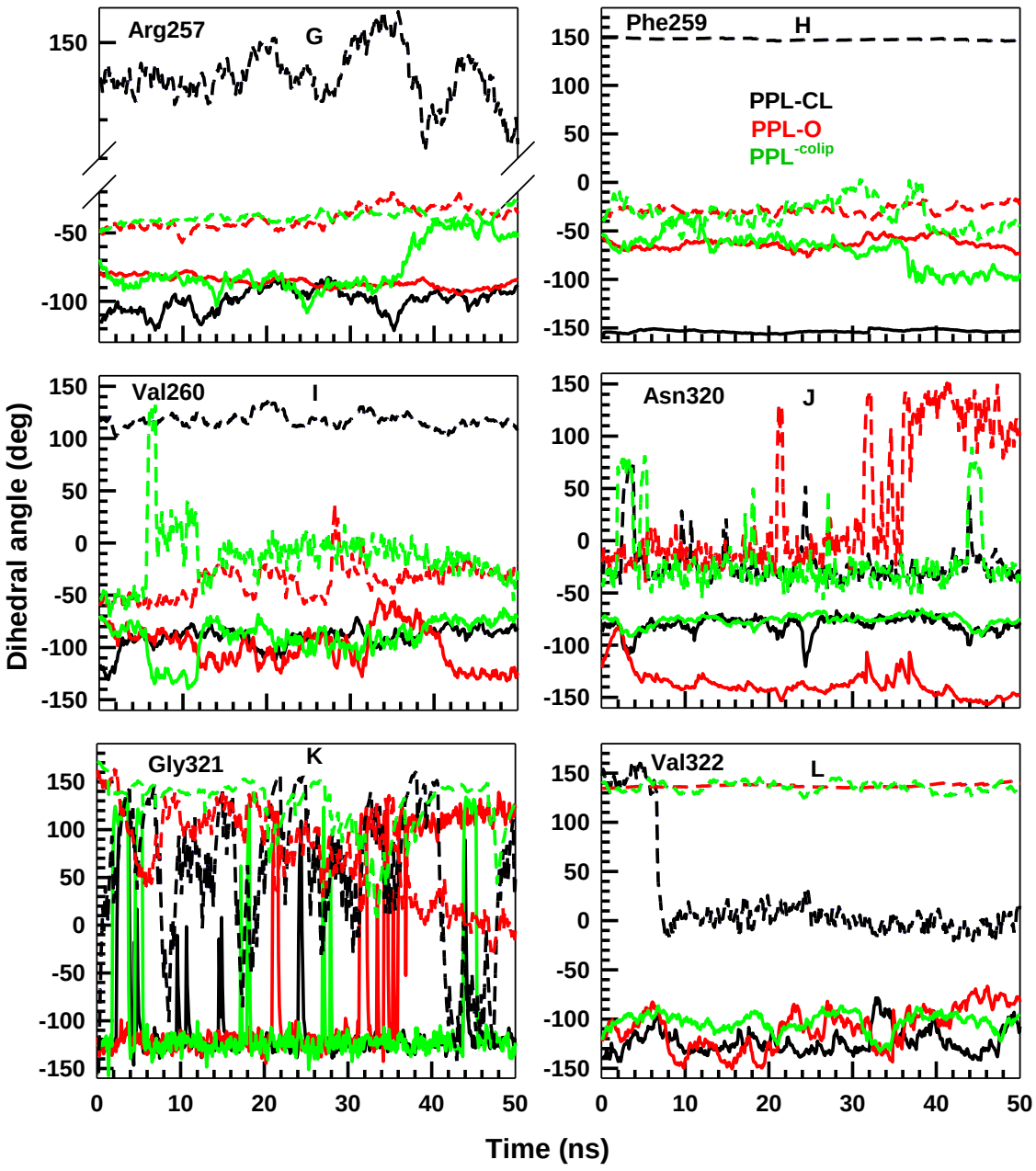


Figure 23. The changes in dihedral angles phi (solid lines) and psi (dotted lines) of selected residues in and around the lid domain for PPL-Cl (black), PPL-O (red), and PPL^{-colip} (green).

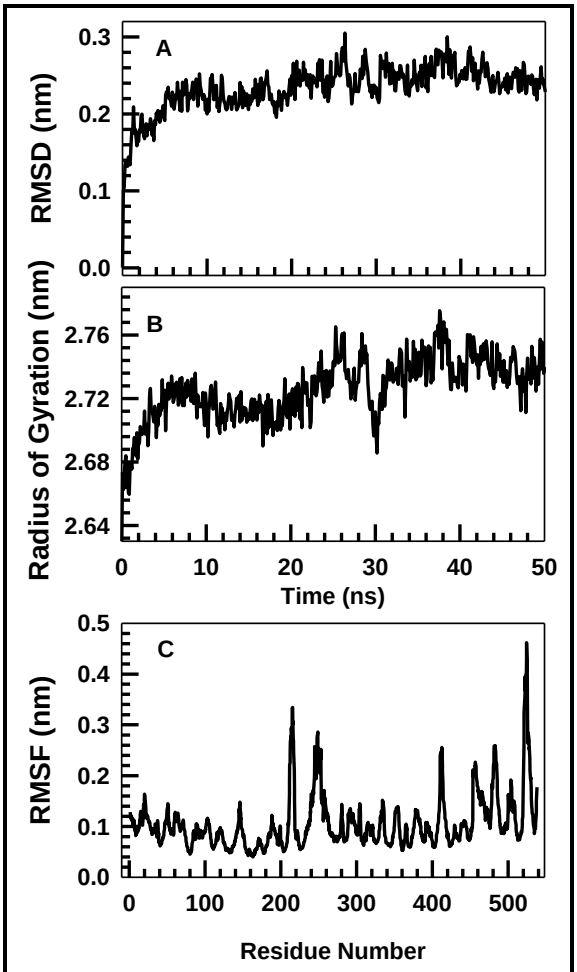


Figure 24. (A) RMSD and (B) radius of gyration of PPL^{+colip} complex. (C) RMS fluctuations of residues in PPL^{+colip} during MD simulation.

2.4.9 PPL-colipase Complex

A 50 ns simulation of lipase-colipase complex (PPL^{+colip}) was also carried out. Analysis of the simulation results showed that RMSD and R_g of PPL^{+colip} (Fig. 24A & B) were higher than the corresponding PPL^{-colip}. Moreover, RMSF of colipase in PPL^{+colip} showed more fluctuations than the backbone of lipase alone in the complex (Fig. 24C). The major conformational changes during the simulation were observed at the interface of lipase and colipase, and around the lid region of the lipase. Out of eight hydrogen bonds between lipase and colipase [132] five of them were stable throughout the simulation. However, the hydrogen bonds formed by Asn241 with Glu15, Ser244 with Leu16, and Gln369 with Glu64 (in the order of lipase-residue with colipase-residue) were found to be weak during the simulation (Fig. 25B and Table 2). Apart from the earlier characterized hydrogen bonds, additional interactions were also found during the simulation, but their existences were transient and were not considered for further analysis.

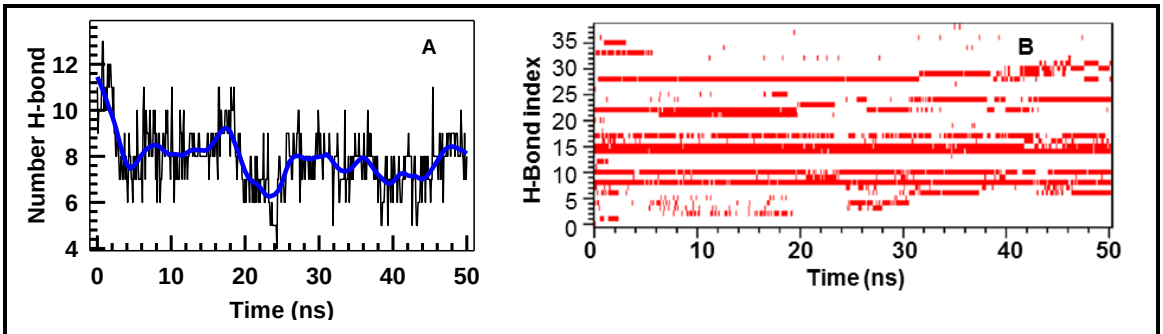


Figure 25. H-bond profile of PPL^{+colip}: (A) Total number of H-bonds found between lipase chain and colipase during the simulation. (B) The existence of H-bonds as indexed in Table 2 during the simulation.

Table 2. The list of H-bonding interactions between lipase and colipase‡

No	Donor	Hydrogen	Acceptor	No	Donor	Hydrogen	Acceptor
1	ASN241N	ASN241H	GLU15OE1 (C)	21	LEU16N (C)	LEU16H	SER244OG
2	ASN241N	ASN241H	GLU15OE2 (C)	22	ARG38NH1 (C)	ARG38HH11	SER244OG
3	ASN241ND2	ASN241HD21	GLY14O (C)	23	ARG38NH1 (C)	ARG38HH11	VAL247O
4	ASN241ND2	ASN241HD21	GLU15OE1 (C)	24	ARG38NH2 (C)	ARG38HH21	SER244OG
5	ASN241ND2	ASN241HD21	GLU15OE2 (C)	25	ARG38NH2 (C)	ARG38HH21	VAL247O
6	SER244OG	SER244HG	GLY14O (C)	26	ARG38NH2 (C)	ARG38HH21	ASP248O
7	SER244OG	SER244HG	GLU15OE1 (C)	27	ARG44NE (C)	ARG44HE	ASN366O
8	SER244OG	SER244HG	GLU15OE2 (C)	28	ARG44NH1 (C)	ARG44HH11	ARG368NH1
9	ASN366ND2	ASN366HD21	GLU45O (C)	29	ARG44NH1 (C)	ARG44HH11	ASP390OD1
10	ARG368NE	ARG368HE	LYS42O (C)	30	ARG44NH1 (C)	ARG44HH11	ASP390OD2
11	ARG368NH2	ARG368HH21	LYS42O (C)	31	ARG44NH2 (C)	ARG44HH21	ASP390OD1
12	GLN369N	GLN369H	GLU64OE1 (C)	32	ARG44NH2 (C)	ARG44HH21	ASP390OD2
13	GLN369N	GLN369H	GLU64OE2 (C)	33	GLU45N (C)	GLU45H	ASN366OD1
14	GLN369NE2	GLN369HE21	GLU64OE2 (C)	34	ASN46ND2 (C)	ASN46HD21	ASN366OD1
15	GLN369NE2	GLN369HE21	ARG65O (C)	35	ASN46ND2 (C)	ASN46HD21	ASN366ND2
16	LYS400NZ	LYS400HZ1	GLU45OE1 (C)	36	ARG65NH1 (C)	ARG65HH11	GLU440OE1
17	LYS400NZ	LYS400HZ1	GLU45OE2 (C)	37	GLY66N (C)	GLY66H	GLU440OE1
18	LYS400NZ	LYS400HZ1	ASN89OD1 (C)	38	GLY66N (C)	GLY66H	GLU440OE2
19	ARG5NH1 (C)	ARG5HH11	ASN241OD1	39	ASN89ND2 (C)	ASN89HD21	LEU442O
20	ARG5NH2(C)	ARG5HH21	ASN241OD1				

‡**The index numbers correspond to the label used in Fig. 25B. The residues marked with (C) belong to colipase.**

2.5 Discussion

2.5.1 Global Changes in Different Forms of PPL

The preliminary analysis of changes in RMSD, R_g and RMSF of PPL-Cl, PPL-O and PPL^{-colip} suggest that the atomic fluctuations of PPL^{-colip} are less compared to others (Fig. 4). However, the region covering the residues 210-220 (β9-loop) and 410-415 (β5'-loop) are found to have significant fluctuations along with the lid region. The fluctuations around β5'-loop may be attributed to high thermal B-factor of this loop in the X-ray crystallographic structure. Moreover, PPL^{-colip} and PPL-O do not show much change in their compactness, but PPL-Cl shows slight increase in its R_g value suggesting trivial loss in its compactness. This might be due to the higher fluctuations observed in domain 2 which could increase the mean distance between the domains during the simulation, particularly at above 25 ns (Fig. 4B). This could be attributed to the loss of some of the hydrogen bonding interactions between the domains (for instance, W339-D279, and K338-G280), though some of the H-bonding interactions are stable in both PPL-O and PPL-Cl (C300-N424). This is further supported by increase in the distance measured between the center of masses of domain 1 and domain 2 (Fig. 4B-inset). Further, the RMSD of PPL-O also increases during simulation which almost reached the same value as PPL-Cl at 50 ns (Fig. 4A). This change could be directly attributed to the fluctuations around the lid region which do not necessitate any significant change in the compactness of the protein, thus not altering the R_g of PPL-O (Fig. 4B).

The SASA analysis suggests that some of the buried hydrophobic residues in PPL-Cl are exposed during simulation which increases the hydrophobic SASA values whereas the hydrophilic SASA almost remains same throughout the simulation (Fig. 8). This might be due to the increase in distance between the domains which also leads to the increase in R_g value of PPL-Cl as discussed earlier. However, in the case of PPL-O, the hydrophobic SASA value comes down indicating that some of the exposed hydrophobic residues are buried which could be attributed to the movement of lid closer to domain 1 along with β5- and β9-loops movements. The hydrophilic SASA of PPL-Cl is higher than the PPL-O throughout the simulation as it was observed earlier from the crystal structure of PPL as well [166]. It is also interesting to note that the distance between the domains is not altered in PPL-O (Fig. 4B-inset). Further, at the end of the simulation SASA values of PPL-O approaches nearer to that of PPL-Cl suggesting that the conformational changes occurring in PPL-O may lead to closed-lid conformation. Also, the mean values of electrostatic interaction energies indicate that the closed-state is energetically favorable compared to the open conformation in aqueous medium. This is further supported by the higher negative value obtained for the polar solvation free energy of PPL-Cl (Table 1) [165].

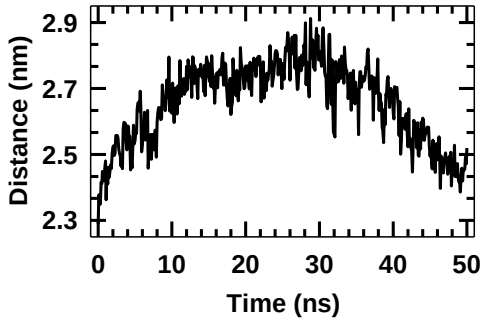


Figure 26. The distance between domain 1 and the lid region of PPL-O during simulation. The residues (150-160) with least RMSF on domain 1 and W253 on the lid were considered as references for distance measurement.

2.5.2 Essential Dynamics and Lid Movement in PPL-O

The major conformational changes in PPL-O occur around the lid region. Analyzing the global movement of the lid using the distance between domain 1 and the lid (Fig. 26 and video file V2) predicts that initially the lid moves away from domain 1 which later change its trajectory and approaches domain 1. This is accompanied with dihedral and secondary structural changes of the residues around it.

Principal component analysis of C_α-atoms of PPL-O suggests that there are highly correlated motions among the residues covering the lid region of the protein. The contribution of other regions could not be identified from cPCA. However, the covariance matrix and the first EV component of dPCA (Fig. 10C and 16B) propose that the region covering the residues 318-322 (surface loop) show correlated motions along with the residues 303-307 (β12-loop) and 210-215 (β9-loop). This emphasizes the associated motions of the domain-domain interface along with the lid. It may also be noted that part of β9-loop constitutes the catalytic groove where substrate binds, particularly the residues L214 and F216 of β9-loop are suggested to be interacting with substrates or inhibitors through van der Walls interactions. Major side chain conformational changes were observed between open and closed

conformations on F216 [132, 166]. It was also proposed that these residues on β9-loop may directly participate in the lipolysis reaction by stabilizing the acyl-enzyme intermediate [167]. Though the covariance matrix of cPCA could not identify the correlated motions between the lid and the other regions, the projection of extreme structures obtained from the first EV of cPCA indicates concomitant changes in these regions (Fig. 13). This could be due to the fact that cPCA corresponds to the changes occurring in 3D-coordinates representing internal as well as global dynamics of the system whereas dPCA is specific to internal dihedral changes irrespective of their positional fluctuations.

From the eigenvector index of PPL-O, it is observed that first 5 EVs could cover 78% of the conformational changes. The analysis of the EVs from 2 to 5 also predicts that major fluctuations are around the lid region alone, the other residue fluctuations are lower in amplitude, except a few noticeable dynamics around the domain-domain interface and in domain 2 (β5' hydrophobic loop). Based on these observations, further residue specific analyzes were carried out to understand the lid movement in PPL-O.

The electrostatic force exerted by charged residues on the lid in all the three coordinates reasonably correlate with the movement of the atomic positions of the residues for initial 30 ns. This infers that the initial lid movement is mainly influenced by electrostatic interactions. The major differences in lid region between open and closed conformations are the salt bridges formed with *N*-terminal domain and within the lid region. For instances, R257 of lid in open conformation makes salt bridge interaction with D80 on β5 loop and D258 of lid interacts with K269. However, in closed conformation the lid makes mostly non-polar interactions with the surrounded regions [161, 168]. Nevertheless, above 30 ns the correlation becomes poor which might be due to predomination of other factors over the electrostatic force.

In order to probe further, radial distribution function of the solvent molecules around the lid region and the dihedral changes of the residues in and around the lid were calculated. The relative water density in the first solvation shell of PPL-O increases around the lid region at above 30 ns (Fig. 21). Correlating this with the increase in hydrophilic SASA values and changes in distance between lid and domain 1suggest that increase in solvation energy could be the major driving factor to move the lid towards closed conformation. In fact, it has been observed that total-water accessible surface area of pancreatic lipase is more in closed conformation, particularly due to increase in water accessibility of charged residues on the lid and β5-loop. Moreover, these lid movements necessitate significant changes in the dihedral angles of residues in and around the lid. The components of first EV from dPCA provide vital information on these dihedral changes (Fig. 16B). From this plot, 18 amino acid residues which could show larger dihedral changes were chosen for further analysis (Fig. 22 and 17).

2.5.3 Free Energy Surface and Lid Movement of PPL-O

The free energy surface obtained by the projection of first EVs of cPCA and dPCA proposes the possibility of four major transition states in the dynamics of PPL-O (Fig. 20) which are discussed in detail:

2.5.3.1 Step 1 - SO –T1- S1: (4 ns -10 ns - 11.6 ns)

High amplitude movements were noticed on the lid residues and on β9-loop. The lid moves away from domain1 due to columbic forces of the charged residues on the lid, particularly D248, D250, E254, and D258. This movement is initiated by breaking the salt bridge between the lid residue R257 and D80 in β5-loop, and the H-bonding interaction of R257 with Y268. During this process, the salt bridge between D258 and K269 in domain 1 is also lost (Fig. 27A). These interactions are found to be crucial in stabilizing open conformation of the lid and conserved in all the classical lipases [161, 168, 169]. These changes influence the hydrophobic part of the β9-loop also to move in the same direction, consequently the active site residues H264 and F78 are moved away from the protein core. Also, significant changes are observed in β5-loop such as F78 moves closer to the active site residue S253 (Fig. 27A-inset) and considerable changes in the backbone dihedrals of residues I79 and E84 (Fig. 27B) which are consistent with the variations observed in the crystal structure of closely associated HPL [161]. Comparing the initial changes in the dihedral angles of lid residues suggest that the transition from S0 to S1 does not necessitate larger dihedral changes, hence, secondary structural changes. The major energy barrier for this transition could be the columbic forces. During this transition, the interface of the domains do not show much conformational changes, however, large dihedral changes occur around C300-C305 disulphide bond (Fig. 28D) which is part of the long loop connecting helix12 and helix13.

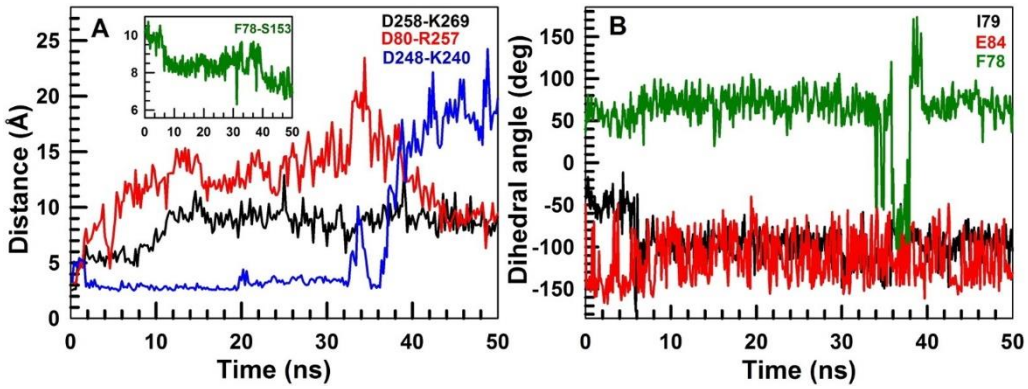


Figure 27. (A) Salt bridge interaction between D258-K269 (black), D80-R257 (red), and D248-K240 (blue) measured as a distance between their side chain atoms during PPL-O simulation. The inset shows the distance between C_α-atoms of F78 and S153. (B) Change in dihedral angles of the residues in β5-loop during PPL-O simulation. Ψ-angle of I79 (black), Φ-angle of E84 (red), and χ_2-angle of F78 (green) are presented.

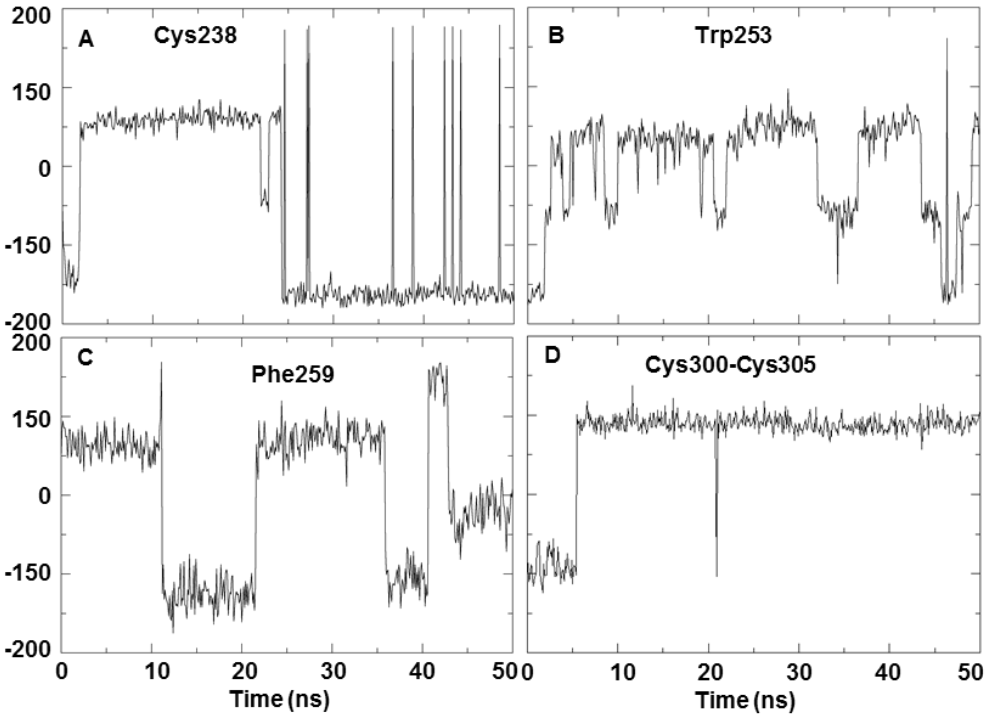


Figure 28. Changes in (A) side chain (χ_1) dihedral angle of Cys238 (B) side chain (χ_1) dihedral angle of Trp253 (C) side chain (χ_2) dihedral angle of Phe259, and (D) disulphide bond dihedral of Cys300-Cys305 during the simulation of PPL-O in water.

2.5.3.2 Step 2 - S1 to S2: 11.6 ns - 31.4 ns - 41.3 ns

The lid which initially moves away from domain 1 (up to ~ 30 ns), starts moving towards domain 1 by adapting a circular pathway (video V2). It has a major dihedral barrier around the hinge regions connecting the lid with domain 1, specifically on the lid disulphide bond (C238-C262), Gly251 and Gly 255. When this dihedral barrier to twist the moving lid towards domain 1 is crossed, the lid produces a considerable positional changes and moves quickly towards domain 1, particularly the loop region 247-260

shows a major movement. The FES also predicts a considerable energy barrier (T2 ~ 8 kJ/mol) for these movements at ~31 ns. This is further verified by measuring the distance between the centers of masses of domain 1 and the lid which shows that the lid starts approaching domain 1 after ~30 ns (Fig. 26). Analysis of dihedral angles predict that major conformational changes occur around the *Φ*-angle of Cys262 (Fig. 22F) along with the side chain dihedral (χ_1) of Cys238 (Fig. S13A) suggesting that the rotation of lid is centered on *–H-N-*C_α- of Cys262 and -C_α-C_β- of Cys238, respectively. Also, the loop region consisting amino acids 249-255 show major movement with relatively less dihedral changes except around Gly 251 and Gly255. This could be due to the fact that the energy requires to change the dihedrals of glycine residues are relatively lesser than other residues, hence, the system might adapt Gly for major rotational movement. The active site residues H264 & F78 and the lid residues R257 & D258 interacting with residues Y268 and K269 of domain 1, which earlier moved away from the protein core, are also found to be moving towards the core.

Though the surface β9-loop does not show much change, β5-loop approaches towards the protein core. The structural comparison of open and closed conformations of HPL suggests that β5-loop needs major drift in the main chain residues of 78 to 84 and in the side chain of F78 [161]. The further movement of F78 towards the active site S153 (Fig. 27A-inset) decreases its solvent exposure. Also, the side chain of F78 showed considerable rotation during this transition along its χ_2 angle (Fig. 27B). Comparison of PPL with that of guinea pig pancreatic related protein 2 (GPLRP2) which has shortened lid and increase in semi-polar residues in the β5-loop with increase in hydrophilic/lipophilic ratio in the open forms suggested that the interactions between lid and β5-loop have major role in substrate specificity of the proteins [170]. Moreover, this transition also leads to loss of helices in the lid region and breaking of salt bridge between K240 and D248 within the lid (Fig. 27A). These changes are accompanied with trivial changes in domain 2 and the interface of domains. The parallel beta-sheet formed between the regions 225-229 and 324-328 is disrupted during this lid closure movement.

2.5.3.3 Step 3 - S2 to S3: 41.3 ns – 42.5 ns – 43.1ns

As shown in Fig. 26, the lid moves further towards domain 1 during this simulation period. This movement happens around the lid region from the residues 241 to 255 with the loss of helicity. The major dihedral changes around Gln245, Ile 252, and Trp253 (Fig. 22) could be the energy barrier for this transition which twists the loop region of the lid and facilitates it to move further towards domain 1. It may be noted that the H-bonding interaction between Gln245 and Gly237 plays an essential role in flexibility and proteolytic-resistance of the lid in pancreatic lipases and the PLRP2s lacking this interaction are susceptible to proteolysis [171]. These changes might be associated with the side chain movements of Trp253 through rotation along its χ_1 angle and the rotation of phenyl ring of Phe259 along its χ_2 angle (Fig. 28). During these

transitions, the center of β9-loop also moves towards the protein core. However, the active site residues show less conformational changes.

2.5.3.4 Step 4 - S3 to S4: 43.1 ns – 44.1 ns – 47.8 ns

During this transition, the loop region of lid covering the residues from 249 to 253 show major movements. The other regions of the protein have insignificant changes. The transition energy barrier, T4, may be due to the side chain movements of Ile249 and 252 in the loop region of lid and marginal changes around Asp250 at the center of the loop (Fig. 22 and 23). In addition, β9-loop moves along with lid region, but no changes are observed around the active site region.

2.5.4 H-bonding Interactions and Dynamics of Water

Comparison of open and closed states of PPL suggests the presence of at least one water molecule, in both the cases, at the interface of the domains involved in hydrogen bonding with the side chains of Arg340 and Asn424 (Fig. 29A). Apart from this, the open state also possess two more water molecules at the interface: one of them involved in H-bond with the amide backbone atoms of Tyr311 and Met308, and another one with the amide backbone of Gly283 and the side chain of Pro306 (Fig. 29C). These additional water molecules found at the interface might restrict the dynamics of domain 2 in PPL-O, thus, not altering its R_g value. In the lid region, a water molecule is found to be interacting with the side chain of Arg266 and the backbone of Gly331 in both the conformations. However, this along with two other water molecules (Fig. 29B) interacting with the side chains of charged residues in the lid (Glu254, Asp248 and Asp258) move away from the interior of lid at the early phase of simulation (within 10-20 ns) suggesting that the rearrangement of solvent-mediated interactions is also essential during lid closure. In case of PPL-Cl, it is noted that one of the active site residue, Asp177, always interacts with a water molecule through its side chain suggesting that this could be a stabilizing factor for the charged side chain of Asp177 in the closed state.

Further, the analysis of H-bonding interactions between the residues in domain 1 and domain 2 show that there are at least three stable H-bonding interactions exist in both the states: Cys300 and Asn424, Trp339 and Asp279, and Arg338 and Gly280. However, the open conformation possesses two additional H-bonding interactions between Asp329 and Asp388, and Asn295 and Glu386 which could also contribute to the restrained dynamics at the interface of the domains, thus, limiting the changes in the hydrodynamic radius of PPL-O.

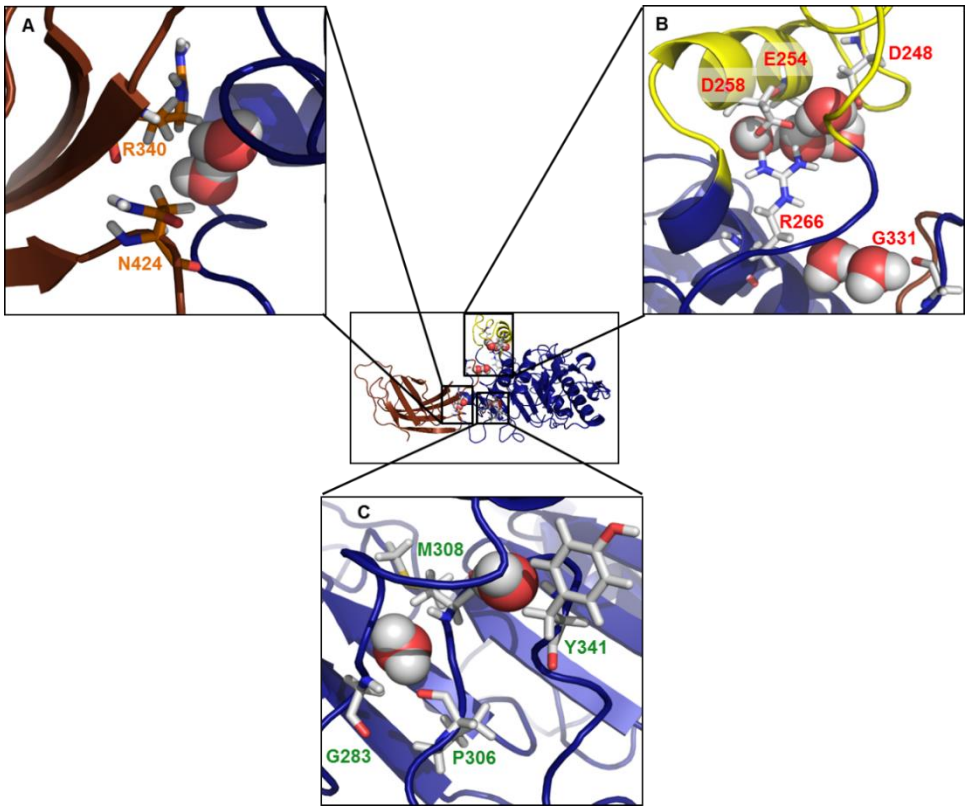


Figure 29. Selected water molecules in PPL-O: The central panel presents the water molecules (shown as spheres) on lid (yellow ribbon), domain 1 (blue ribbon) and domain 2 (brown ribbon). Focused view of the specific regions: (A) domain-domain interface, (B) lid and (C) domain 1 region which is closer to lid and domain 2.

2.5.5 Dynamics of PPL-Cl and PPL^{-colip}

PPL-Cl has more R_g and SASA values [166] compared to PPL-O and PPL^{-colip}. This is probably due to the higher fluctuations on domain 2 along with β5-loop and on the interfacial region between the domains in PPL-Cl. The overall fluctuations include the changes around the lid region, particularly the residues which interact with colipase. The first eigenvector component obtained from the simulation of PPL-Cl also suggests that the major fluctuations are around β12-loop and a β sheet (370-374) in *C*-terminal domain. PPL-O and PPL^{-colip} have comparable RMSD values suggesting that the open conformation constrained the dynamics of *C*-terminal residues. It could be an inherent property of PPL to have more fluctuations in the closed state to facilitate the binding of colipase whereas constrains the dynamics of domain 2 in the opened state where the activity is more.

2.5.6 Dynamics of Colipase

Major fluctuation in the colipase is found around lipase-colipase interfaces and the solvent exposed helix which mainly consist of hydrophobic residues. However, the presence of colipase has tremendously reduced fluctuations on lipase chain including the lid region. The residues around β9-loop and at the interface of lid and colipase show more

fluctuations (Fig. 24C) which essentially participate in substrate accessibility of the protein. The dynamics of colipase and its essential role in protein activity is further discueed in Chapter 4.

2.5.7 Lid Opening Dynamics of Different Lipases

The sequence of mammalian pancreatic lipases (PLs) show close similarity and highly conserved active site region. Also, the lid, β5-loop, and β9-loop are found be well conserved in these proteins. PLs share common sequence and structural features with PLRP proteins. However, their lid regions show changes including deletions of amino acid residues [170]. These changes on the lid alters substrate specificity of the lipases, in fact, PLRP2s can accommodate phospholipid and galactolipids as their substrates. Besides, they show higher flexibility in lid region and no interfacial activation [167, 170, 171]. Further, microbial lipases, which are structurally similar to the catalytic *N*-terminal domain of pancreatic lipases, show similarity in the lid conformation-dependent accessibility of the active site region. The studies on these proteins suggest that lid closer movement in aqueous environment is driven by constricting the hydrophobic face of the amphipilic helix of the lid towards inner core of the protein [172-174]. Though in some lipases the process can be considered as a rigid body movement since the secondary structural changes are minimal (for instance, *Burkholderia cepacia* lipase [174, 175]), other microbial lipases show major secondary structural changes during the lid closure (for instance, *Candida rugosa* lipase [176]) as observed in PPL. However, the changes observed in other regions are trivial in both microbial and mammalian pancreatic lipases. Further, the lid closure movements followed using MD simulation in microbial lipases show a single trajectory such as continuous rolling movement of helix 5 in *Burkholderia cepacia* and a linear motion of the lid in *Candida rugosa* lipases [174, 176] whereas PPL initially moves away from the catalytic domain and later turns back to approach closed conformation (video V2).

2.6 Conclusions

The dynamics of PPL-Cl, PPL-O and PPL^{-colip} were followed in water using an all-atom force field. PPL-Cl shows large fluctuations due to domain-domain movements which does not alter the lid conformation significantly. In PPL-O, the domain-domain movement is less and the lid moved towards the closed conformational state. The lid movement is initially influenced by columbic interactions which moved the lid further away from domain 1. After, about 30 ns, the lid swirls back towards domain 1 and approached closed conformation. The major energy barriers during these movements are implied by dihedral changes on the lid residues. Rearrangements of solvent-mediated interactions also contribute to the lid movement. However, PPL in the presence of colipase shows less fluctuation. The results clearly propose that the lid closure is not a uni-directional movement and the polar environment could drive the lid towards closed conformation at ambient conditions in the absence of colipase.

Chapter 3

Lid Dynamics of Porcine Pancreatic Lipase in Non-Aqueous Solvents

3.1 Abstract

Understanding the dynamics of enzymes in organic solvents has wider implications on their industrial applications. Pancreatic lipases, which show activity in their lid open-state, demonstrate enhanced activity in organic solvents at higher temperatures. However, the lid dynamics of pancreatic lipases in non-aqueous environment is yet to be clearly understood. Dynamics of porcine pancreatic lipase (PPL) in open and closed conformations were followed in ethanol, toluene, and octanol using molecular simulation methods. *In silico* double mutant D250V and E254L of PPL (PPL_{mut}-Cl) was created and its lid opening dynamics in water and octanol were analyzed. PPL showed increase in solvent accessible surface area and decrease in packing density as the polarity of the surrounded solvent decreased. Breaking the interactions between D250-Y115, and D250-E254 in PPL_{mut}-Cl directed the lid to attain open-state conformation. Major energy barriers during the lid movement in water and octanol were identified. Also, the trajectories of lid movement were found to be different in these solvents. Only the double mutant at higher temperature showed lid opening movement suggesting the essential role of the three residues in holding the lid in closed conformation. The lid opening dynamics was faster in octanol than water suggesting that non-polar solvents favor open conformation of the lid. This study identifies important interactions between the lid and the residues in domain 1 which possibly keeps the lid in closed conformation. Also, explains the rearrangements of residue-residue interactions during the lid opening movement in water and octanol.

3.2 Introduction

Water, as a universal solvent and reaction media, is used in almost all enzymatic reactions. In the real biological world, enzymes don't act in isolation inside the cells. In fact, they are compartmentalized in various forms of microenvironments like completely embedded into membranes, on the surface of membranes, and in the cytosolic microenvironments [177, 178]. In pursuit of addressing these facts, experiments were carried out in organic solvents by replacing the bulk water around enzymes and constraining them in a comparatively smaller volume of 'water pool' with surfactants at the interfaces [179-187]. Such reverse micelle systems paved way for entirely a new arena of molecular enzymology. The experiments conducted in reverse micelle systems showed similar activity as observed in bulk water in many cases [179, 180], and even higher activity in certain cases [184, 185]. It was also observed that some of the enzymes retain their activity even at subzero temperatures in reverse micelles [182, 187]. Further, the activity and substrate specificity could be fine-tuned by adjusting pH, water activity (w_o = [water]/[surfactant]), surfactants, and hydrophobic media [188-190].

Klibanov and co-workers showed that certain enzymes can remain entirely active in organic solvents [121, 122, 191]. Organic solvents as a medium for enzymatic reactions was not initially appreciated, because of already prevailing notion that they

denature proteins by weakening their hydrophobic interactions [192-194]. However, many enzymes were later shown to be functional in organic solvents [195-199]. Besides being active in organic solvent medium, enzymes also exhibit noteworthy properties like molecular memory, stability, tolerance toward hostile conditions, substrate specificity, stereo-selectivity, and regio-selectivity [121, 200, 201]. Moreover, the enzymes like lipases, esterases and proteases could show transesterification reactions in organic medium, which is of great biotechnological and industrial importance [121-123]. Though the use of organic solvents in enzymology is unequivocally inevitable, the structural dynamics of enzymes in such solvents at atomic resolution are not well investigated. Only limited attempts have been made to investigate the mode of action of enzymes in non-aqueous solvents using molecular dynamics simulation [175, 176, 202-211].

Lipase is one of the most widely used enzymes in industries [124, 125]. However, the solvent- and temperature-dependent activity and substrate specificity of lipases, particularly of mammalian pancreatic lipases are less understood [121, 123, 200]. The active state of pancreatic lipases is generally determined by the conformation of lid (open-active or closed-inactive) which could be influenced by different solvent and temperature conditions [176, 202]. The present study uses porcine pancreatic lipase (PPL) as a model to examine the structural changes acquired by PPL, particularly its lid region, in different organic solvents. The results highlight the major interactions stabilizing the closed conformation of lid. Also, the lid opening trajectory of PPL from its closed conformation is analyzed in water and octanol.

3.3 Experimental Methods

Molecular dynamic (MD) simulations of PPL in open (PPL-O) and closed (PPL-Cl) conformations were carried out in four different solvent conditions, ethanol, toluene, octanol, and octanol with water molecules found in the crystal structure of PPL having PDB id 1ETH (octanol + crystal-water), following the protocol mentioned in Chapter 2. Topologies of the organic solvents were obtained from the virtual chemistry database [212, 213]. An *in silico* double mutant of PPL-Cl was created by replacing the residues Asp250 and Glu254 with Val and Leu, respectively in PyMOL 1.3 [155]. The mutant protein (PPL_{mut}-Cl) was energy minimized using OPLS-AA force field. Further, MD simulations of PPL-Cl and PPL_{mut}-Cl were performed at 400 K in water and octanol. A list of all the simulations is given in Table 1. The packing density of the structures were calculated by implementing Voronoi cell, an algorithm which creates atomic volumes by constructing hyperboloid interfaces [214]. The diffusion coefficients of PPL under different solvent conditions were calculated by applying corresponding mean square displacements in the Einstein relation [215]. The conformational entropies of the proteins were calculated by implementing the quasi-harmonic approximation of the eigenvectors.

Table 1. Index of all the Simulations

S.No	Simulation description	Abbreviation	Time (ns)
1	PPL-colipase (open conformation) in water	PPL^{+colip}	50
2	PPL-open conformation in water	PPL-O	50
3	PPL-open conformation in ethanol		50
4	PPL-open conformation in toluene		50
5	PPL-open conformation in octanol		50
6	PPL-open conformation in octanol+crystal-water		50
7	PPL-closed conformation in water	PPL-Cl	50
8	PPL-closed conformation in ethanol		50
9	PPL-closed conformation in toluene		50
10	PPL-closed conformation in octanol		50
11	PPL-closed conformation in octanol+crystal-water		50
12	PPL-closed conformation in water†		50
13	PPL-closed conformation in octanol†		50
14	PPL-mutant-closed conformation in water	PPL_{mut}-Cl	50
15	PPL-mutant-closed conformation in ethanol		50
16	PPL-mutant-closed conformation in toluene		50
17	PPL-mutant-closed conformation in octanol		50
18	PPL-mutant-closed conformation in water†		50
19	PPL-mutant-closed conformation in octanol†		50

†simulations carried out at 300 and 400 K.

3.3.1 Convergence and Comparison of PPL Simulation

Essential dynamics (ED) method [216] filters out anharmonic motions (essential subspace) from near constraint motions (near constraint subspace) in a molecular dynamics trajectory. Essential subspace is defined by eigenvectors with higher eigenvalues which could capture the highest fluctuations. Here, we performed principle component analysis (PCA) in the Cartesian (cPCA) as well as dihedral (dPCA) coordinate spaces to characterize essential dynamics. Further, root mean square inner product (RMSIP) [217] of eigenvectors was used to compare the simulation of the PPL in different conditions. For individual PPL simulations, the RMSIP of first ten eigenvectors of each trajectory was used to compare the essential subspace. Two subparts of the simulation were subjected to RMSIP for calculating the convergence of sampled conformations, both in Cartesian and dihedral essential subspace. The value approaching

one would indicate that the simulation is more converged. To evaluate the overlap between two different trajectories, essential subspaces obtained from first ten EVs of two different simulations were subjected to RMSIP. If the simulations are similar, the value approaches one.

3.4 Results

3.4.1 Dynamics of PPL in Non-Aqueous Solvents

The effect of solvent polarity on the lid dynamics of PPL was monitored using MD simulation of the protein in open and closed conformations in three different non-aqueous solvents: ethanol, octanol and toluene, and in octanol + crystal-water. Root mean square deviation (RMSD) of C_α-atoms of PPL in closed conformation (PPL-Cl) in different solvents is presented in Fig. 1A. The RMSD values indicated that the overall fluctuations were higher in non-aqueous solvents compared to water at the initial phase of simulations. Even the inclusion of crystal water in non-aqueous environment did not affect the fluctuation. However, RMSD values increased in the presence of water and ethanol after 25 ns. Comparison of this with RMSD values of PPL in open conformation (PPL-O) suggested that there were only small amplitude fluctuations except in water and toluene (Fig. 1D). In toluene, there was a significant increase in the fluctuations at the early time scale whereas RMSD of PPL-O in water increased at different stages as discussed in previous chapter.

Change in radius of gyration (R_g) of PPL-Cl and PPL-O in different solvents were monitored (Fig. 1B and 1E). The results showed that in non-aqueous solvents R_g values were lesser than water except in the case of ethanol. Further, R_g of PPL-Cl was higher than PPL-O in all the solvent conditions. In order to understand the contribution of individual residue fluctuation in the overall dynamics of the protein, root mean fluctuations of C_α atoms were calculated for both PPL-Cl and PPL-O (Fig. 1C and 1F). In case of PPL-Cl, more fluctuations were observed in the presence of water and ethanol particularly on the residues in domain 2. However, in case of PPL-O the maximum fluctuations were observed in the presence of toluene among all the non-aqueous solvents.

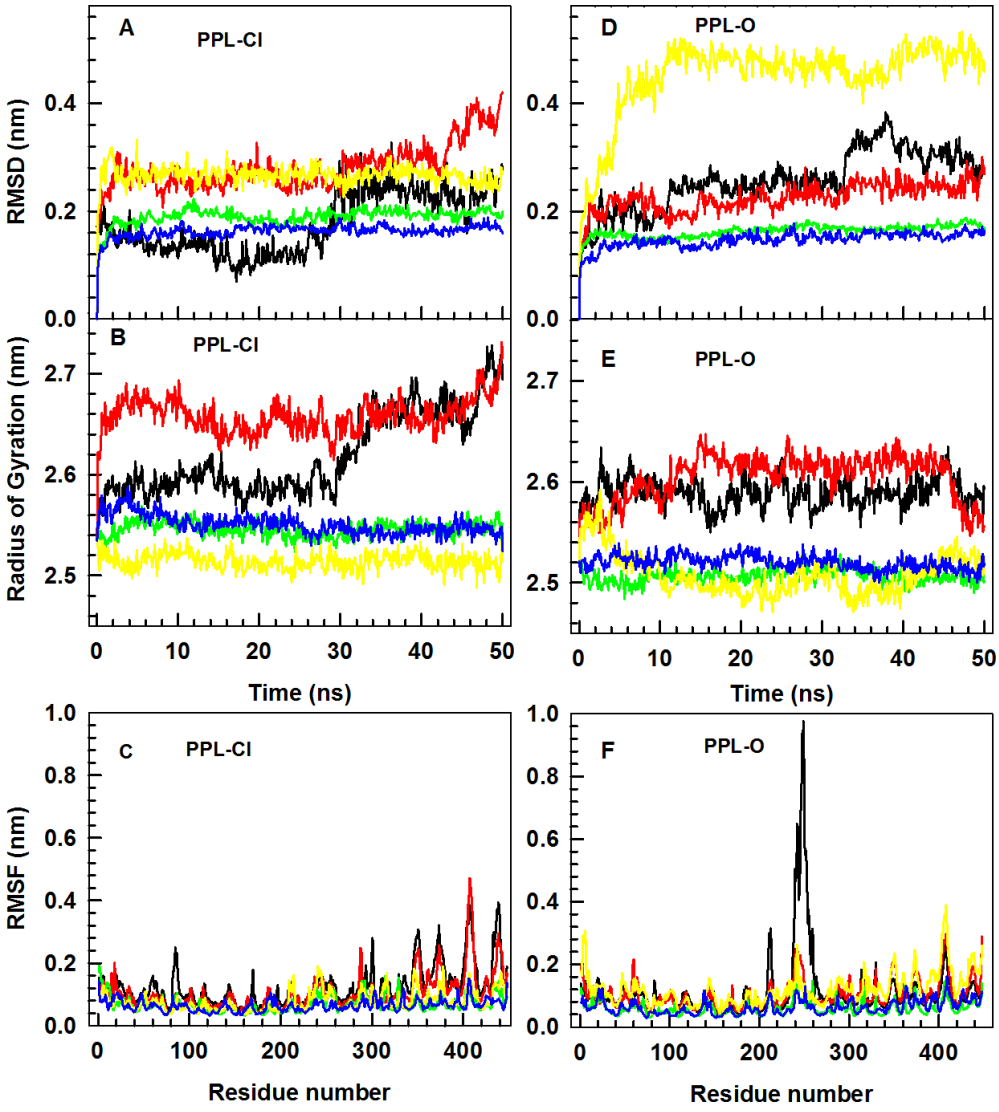


Figure 1. Changes in RMSD, radius of gyration, and RMSF of PPL-Cl (A-C) and PPL-O (D-F) in water (black), ethanol (red), octanol (green), toluene (yellow), and octanol + crystal-water (purple) calculated from 50 ns of MD simulation.

3.4.2 Solvent Dynamics Around PPL

To examine the dynamics of different solvents around PPL, the radial distribution function (g(r)) of the solvent from the protein heavy atoms were calculated (Fig. 2). Water, ethanol, and octanol showed the first maxima around 0.275 nm, though there were differences in their density. Toluene showed the first maximum at longer distance i.e. around 0.4 nm. Diffusion coefficient of PPL in different conditions were evaluated using gromacs tool [215] (Table 2). The values, in general, decreased with increase in the size and decrease in the polarity of solvents. Moreover, in all the solvent conditions, PPL-Cl showed higher diffusion than PPL-O.

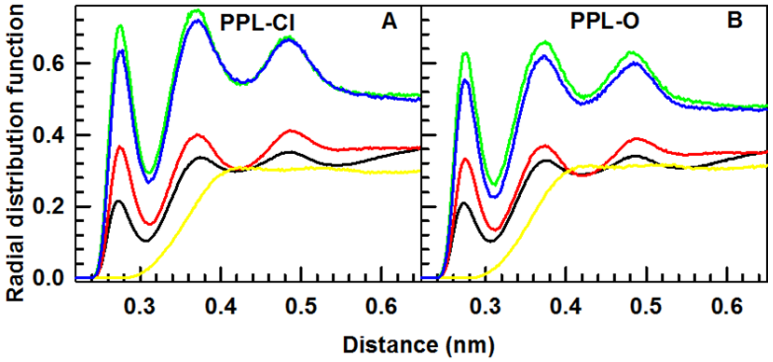


Figure 2. Radial distribution function of solvent molecules calculated by distance between heavy atoms of the protein and solvent atoms for PPL-Cl (A) and PPL-O (B) in water (black), ethanol (red), octanol (green), toluene (yellow), and octanol + crystal-water (purple).

Table 2. Diffusion coefficient of PPL in different solvent environments

Solvent (dielectric constant & radius in Å)	**Diffusion Constant ($1x10^{-7}$ cm^2/s)**	
	PPL-Cl	**PPL-O**
Water (80 & 1.4)	7.04±0.19	5.15±0.01
Ethanol (24.3 & 2.3)	3.37±0.09	2.92±0.06
Toluene (2.3 & 2.7)	3.84±0.04	3.82±0.07
Octanol (10 & 3.4)	0.01848±0.0026	0.01645±0.0036
Octanol + crystal-water	0.01494±0.0032	0.05159±0.0014

Further, solvent accessible surface area (SASA) of PPL was calculated in different solvents (Fig. 3A-B). In polar solvents, water and ethanol, PPL showed higher solvent accessibility than in non-polar solvents. It was also noted that PPL-Cl had higher SASA values compared to PPL-O in polar solvents. Packing density of PPL in different solvent environments were also calculated (Fig. 3C-D). In the presence of water and ethanol, PPL had lower packing density whereas in the presence of octanol and toluene the protein showed higher packing density. It was also observed that the open conformation of PPL was densely packed than the closed conformation in the respective solvent conditions.

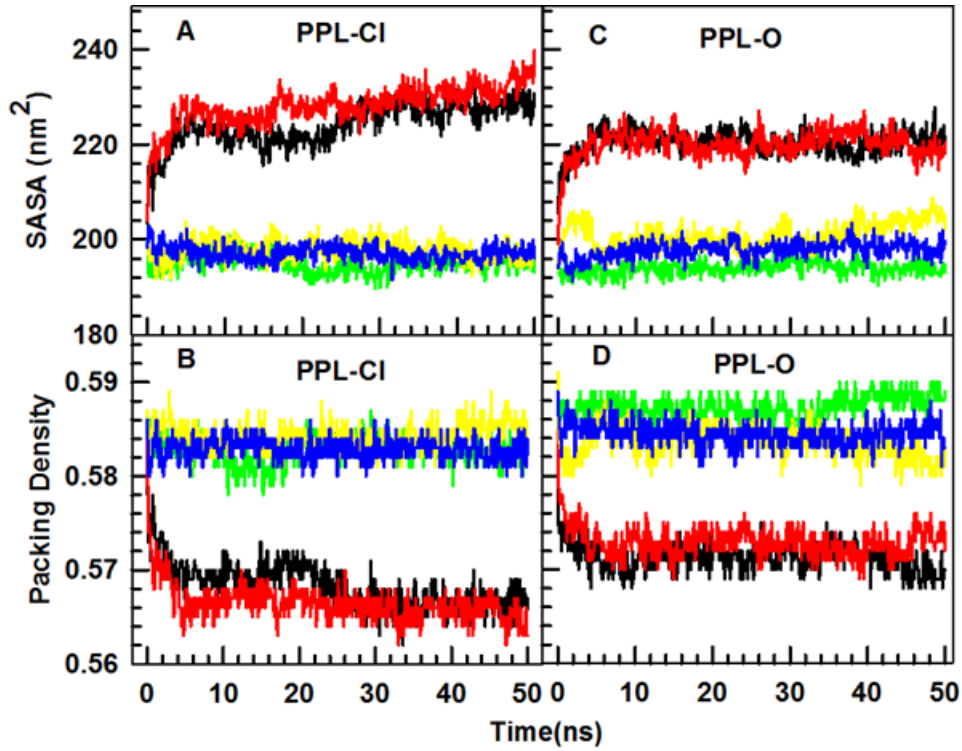


Figure 3. SASA and packing density of PPL-Cl (A and B) and PPL-O (C and D) calculated in different solvents: water (black), ethanol (red), octanol (green), toluene (yellow), and octanol + crystal-water (purple).

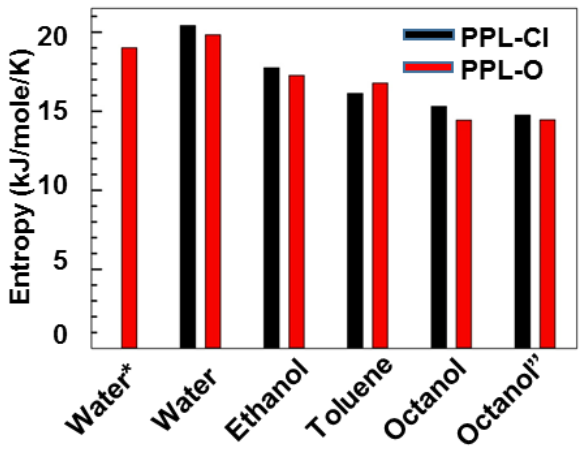


Figure 4. The conformational entropy of PPL-Cl (black bars) and PPL-O (red bars) in different solvent conditions: water, ethanol, toluene, octanol, and octanol + crystal-water (labeled as octanol′′). PPL^{+colip}-O is presented at left extreme for reference (marked with *).

Conformational entropy derived for PPL-Cl and PPL-O (Fig. 4) in different solvents suggested that PPL-Cl had slightly higher entropy than PPL-O except in the case of toluene. Among all the solvents, the protein's conformational entropy was high in the presence of water. As the polarity of solvent decreased, the conformational entropy of PPL was decreased. The addition of crystal waters along with octanol decreased the conformational entropy of PPL-Cl whereas it did not affect PPL-O.

3.4.3 Dynamics of PPL$_{mut}$-Cl in Non-aqueous Solvents

Understanding the conformational changes involved during the opening of lid in different solvents would assist to engineer the protein for higher activity and substrate specificity. However, the simulations carried out here did not show any lid opening

movement in any of the solvent conditions. Therefore, we performed the simulations at slightly increased temperatures which can enhance the conformational search space [218, 219]. The higher temperature simulations carried out at 333 K and 353 K suggested that there was a marginal increase in RMSF values (Fig. 5), however, it could not initiate lid opening (Fig. 6). Comparison of the interactions formed between the residues in *N*-terminal domain and in lid region showed that D250 forms H-bond with Y115 which was stable in most of the solvent conditions even at 353 K (Fig. 7). Also, E254 formed H-bond interaction with D250 in the closed conformation of PPL which was not found in the open conformation. Therefore, we constructed an *in silico* double mutant of PPL with D250V and E254L (PPL_{mut}-Cl) to remove these interactions and to initiate the lid opening. The molecular dynamics of PPL_{mut}-Cl in closed state in different solvents were carried out for 50 ns each. The preliminary analysis of the resultant MD trajectories is presented in Fig. 8. The RMSD changes suggested that the overall fluctuations of residues in PPL_{mut}-Cl were less except in the presence of ethanol. The R_g values also corroborated with this. Analysis of RMSF of individual residues suggested that the major fluctuations observed in PPL_{mut}-Cl were around domain 2 of the protein in water, toluene and octanol. In ethanol, the lid region of the protein also showed relatively larger fluctuations. However, comparison of initial and final structures extracted from MD simulations (Fig. 9) suggested that these conformational changes did not initiate the opening of lid in PPL_{mut}-Cl.

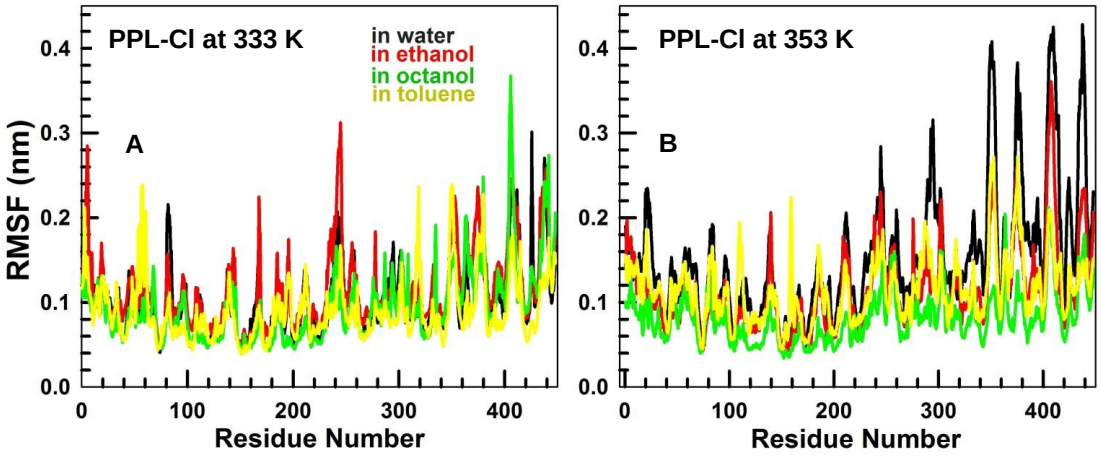


Figure 5. RMSF value of residues in PPL-Cl calculated from the MD simulations carried out at 333 K (A) and 353 K (B) in water (black), ethanol (red), octanol (green) and toluene (yellow).

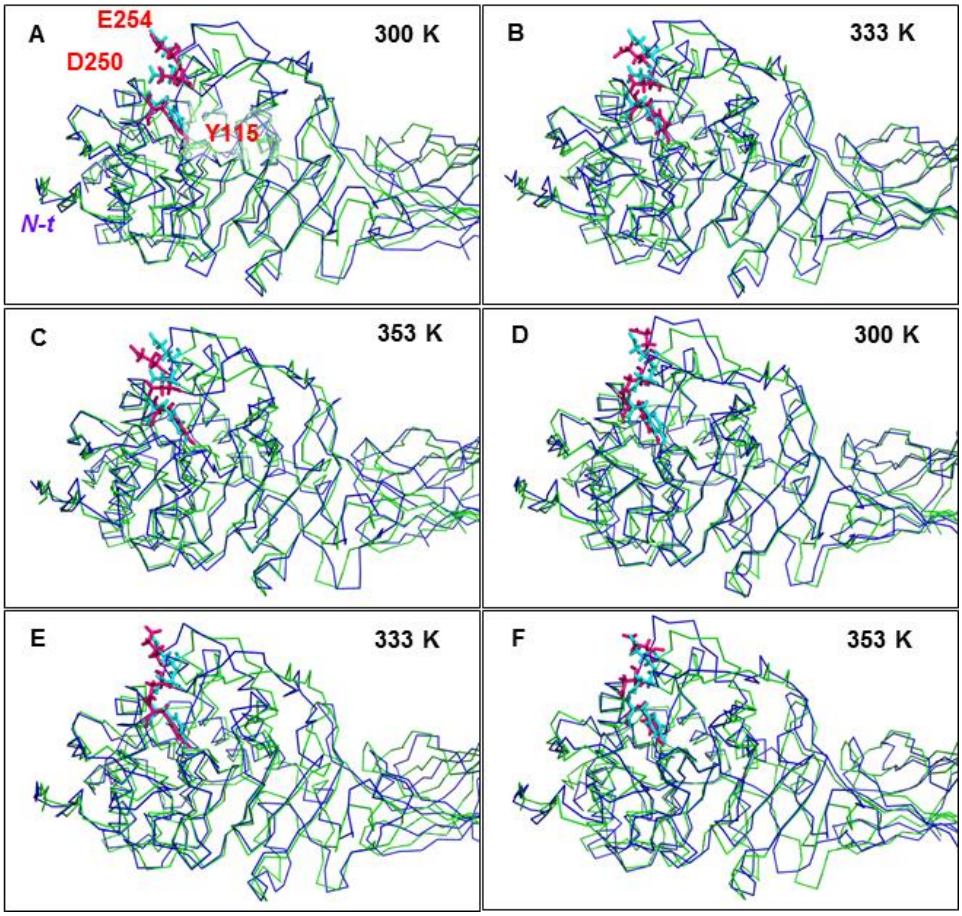


Figure 6. The initial (green) and final (blue) structures of PPL-Cl as obtained from the MD simulations carried out in water (A-C), and in octanol (D-F) at 300 K (A & D), 333 K (B & E) and 353 K (C & F). The initial and final conformations of the residues D250, E254, and Y115 is presented in cyan and red colors, respectively as stick models. *N*-terminal of the protein chain is marked as *N-t* for reference.

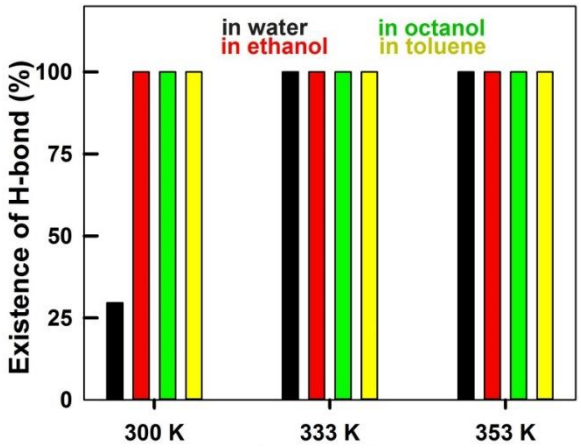


Figure 7. Existence of hydrogen bond between D250 and Y115 during the simulation of PPL-Cl in different solvent and temperature conditions.

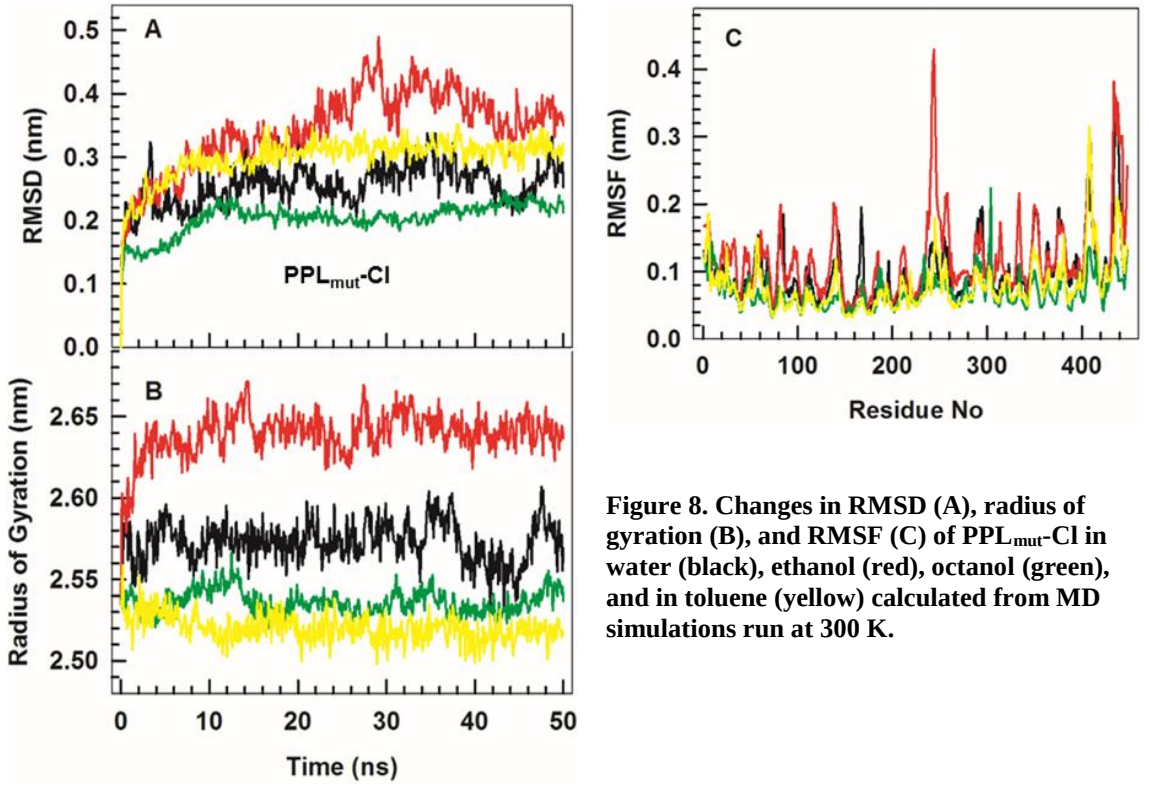


Figure 8. Changes in RMSD (A), radius of gyration (B), and RMSF (C) of PPL$_{mut}$-Cl in water (black), ethanol (red), octanol (green), and in toluene (yellow) calculated from MD simulations run at 300 K.

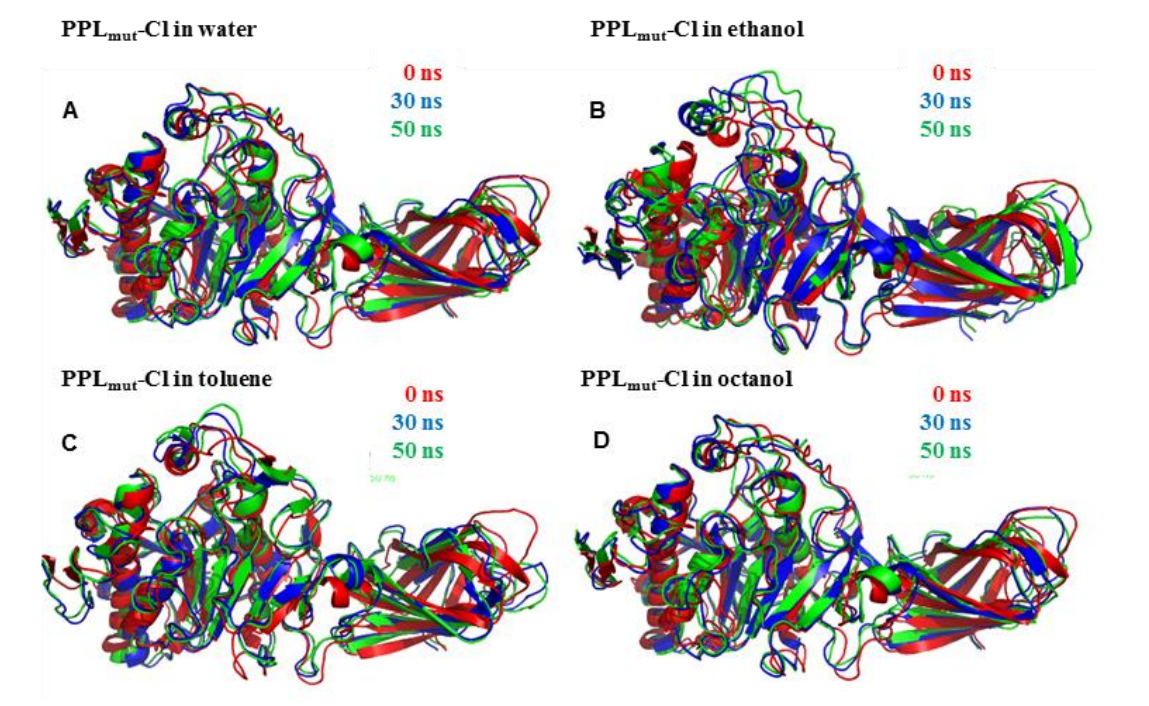
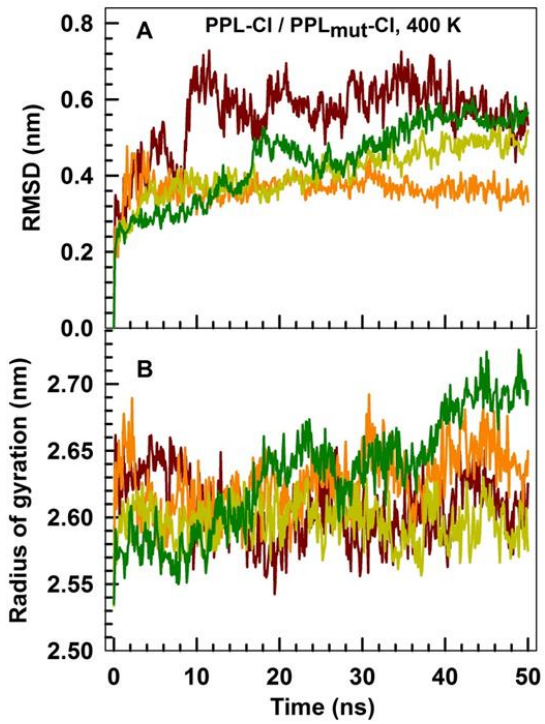


Figure 9. Structures of PPL$_{mut}$-Cl extracted at different time intervals, 0 ns (red), 30 ns (blue), and 50 ns (green) during MD simulations run at 300 K in water (A), ethanol (B), toluene (C) and octanol (D).



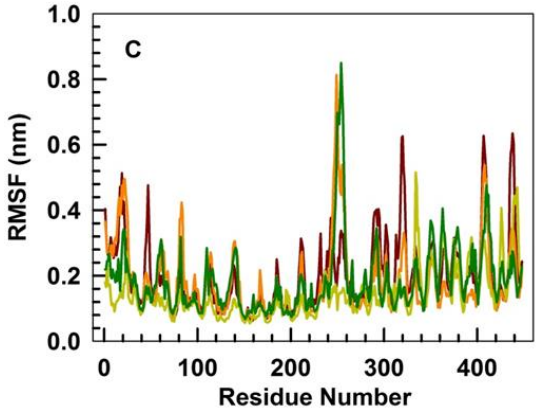


Figure 10. Changes in RMSD, radius of gyration, and RMSF values calculated from MD simulations run at 400 K for wild type PPL-Cl in water (brown) & octanol (olive) and PPL_{mut}-Cl in water (orange) & octanol (green).

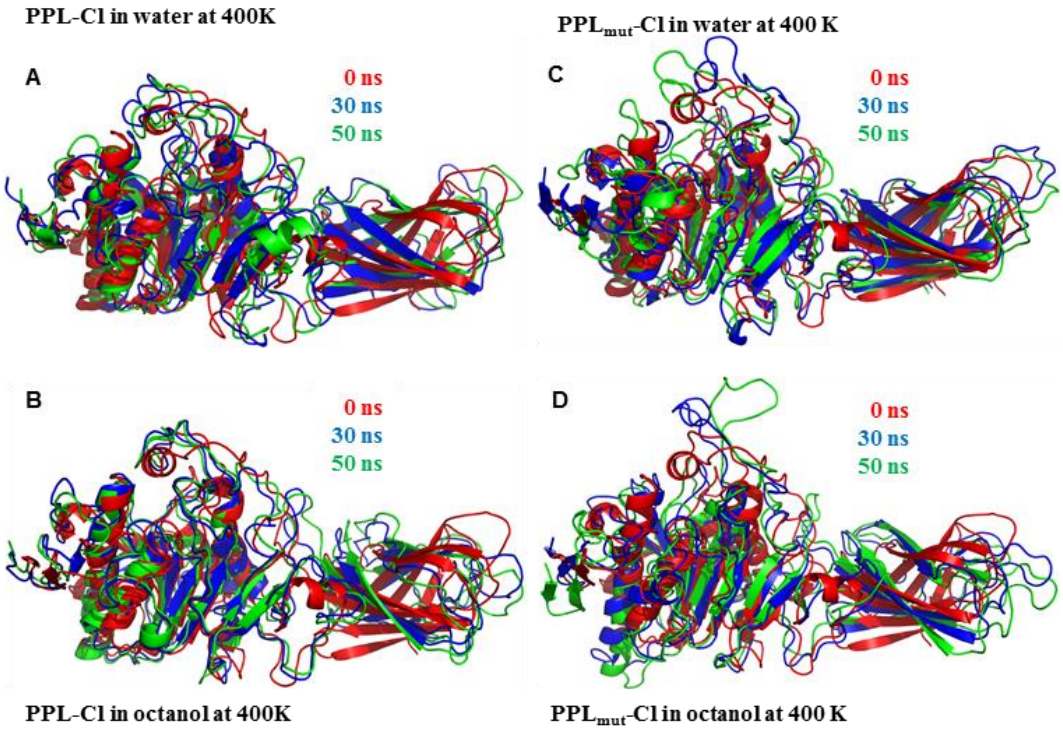


Figure 11. Structures extracted from MD simulations carried out at 400 K for PPL-Cl in water (A) & octanol (B), and PPL_{mut}-Cl in water (C) & octanol (D) at different time intervals, 0 ns (red), 30 ns (blue), and 50 ns (green).

The lid in closed conformation, even after the removal of stabilizing interactions, suggested that there could be still some local energy barriers. Hence, we carried out the simulation of PPL_{mut}-Cl at elevated temperature to enhance the conformational search space [219]. These simulations were performed on PPL-Cl and PPl_{mut}-Cl in water and in

a non-aqueous solvent (octanol) at 400 K to identify major conformational changes. The RMSD of PPL-Cl in water and octanol showed larger changes at the initial period of the dynamics which later reached relatively stable states (Fig. 10). However, the R_g value of PPL-Cl was not significantly altered in these conditions suggesting constrained global changes in the protein (Fig. 10). The noise in the data could be attributed to the increase in simulation temperature. At the same time, PPL_{mut}-Cl in water and octanol showed a gradual increase in their RMSD values along with slight increase in their R_g values as well. Comparing these changes with individual residue fluctuations suggested that the major fluctuations observed in PPL-Cl correspond to the atomic motions around domain 2 whereas in PPL_{mut}-Cl the major fluctuations could be found around the lid region (Fig. 10). Further, the structures of PPL-Cl and PPL_{mut}-Cl extracted from different time points during the simulation clearly indicated that PPL_{mut}-Cl at elevated temperature initiated the opening of lid, in the presence of both water and octanol, whereas such changes in the lid region were not observed in PPL-Cl (Fig. 11).

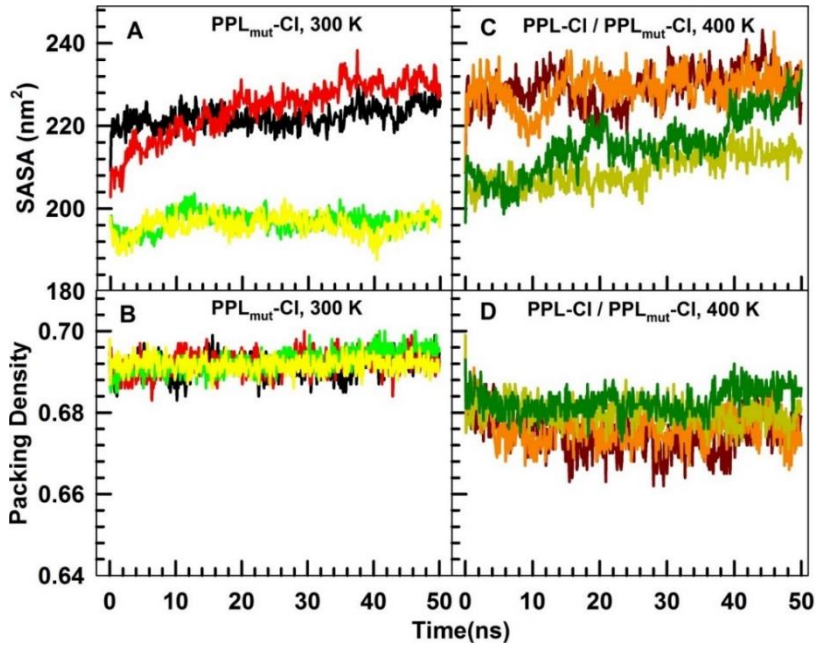


Figure 12. (A&B) SASA and packing density calculated from MD simulations run at 300 K for PPLmut-Cl in water (black), ethanol (red), octanol (green), and toluene (yellow). (C&D) SASA and packing density calculated from MD simulations run at 400 K for wild type PPL-Cl in water (brown) & octanol (olive) and PPLmut-Cl in water (orange) & octanol (green).

SASA values were calculated for PPL_{mut}-Cl in different solvents at room temperature and at 400 K. PPL_{mut}-Cl in water and ethanol showed higher SASA values than in toluene and octanol as it was observed in PPL-Cl (Fig. 12). Moreover, at higher temperature PPL-Cl (Fig. 12) showed increase in SASA values compared to room temperature. This might be due to the larger fluctuations observed around domain 2 of the protein at 400 K which was evident from increased RMSF values of the residues in domain 2. The SASA value of PPL-Cl was similar to PPL_{mut}-Cl in water at 400K, though the lid opening dynamics was observed in the case of PPL_{mut}-Cl. However, in the

presence of octanol PPL_{mut}-Cl showed increase in SASA value than PPL-Cl at 400 K which could be attributed to the lid opening movements in PPL_{mut}-Cl (Fig. 11).

3.4.4 Comparison of All the Simulations

In order to verify whether the simulations run under different solvent and temperature conditions (Table 1) sufficiently represent the possible fluctuations and the variations, the first ten eigenvectors were calculated from Cartesian-PCA for each simulation. The results clearly indicated that first few eigenvectors could cover at least 60-80% of the overall dynamics (Fig. 13). Further, to ensure that the time period of simulation covered the essential subspace of large amplitude fluctuations, the root mean square inner product (RMSIP) for each simulation were calculated using the first ten eigenvectors (Table 3) obtained from both Cartesian- (cPCA) and dihedral- (dPCA) PCAs. In almost all the cases, the values of RMSIP were greater than 0.5 this indicated that the simulation time was sufficient enough to reflect the major fluctuations both in Cartesian and dihedral subspaces (Table 3).

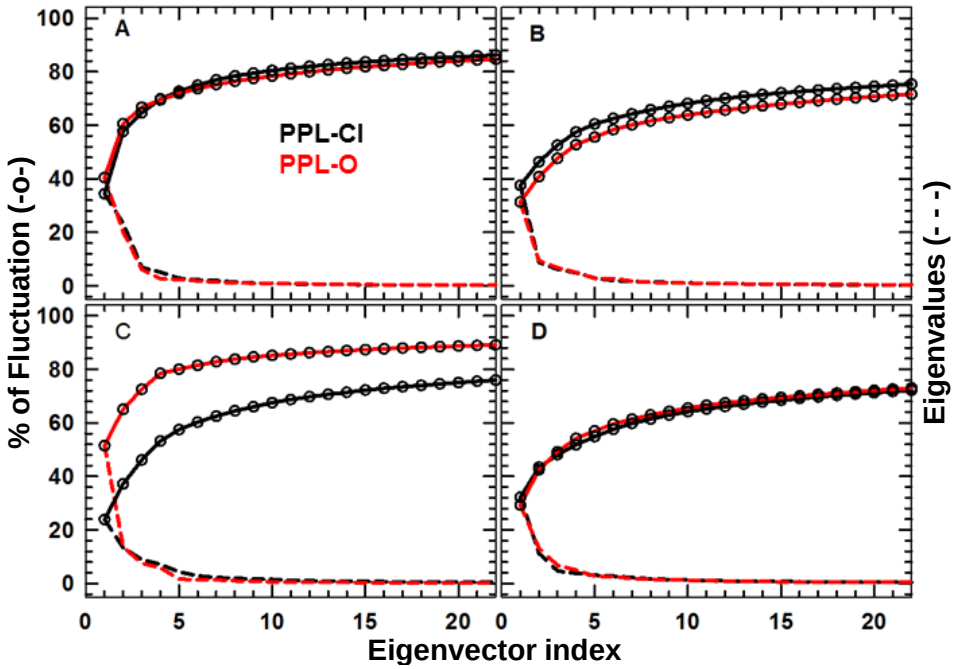


Figure 13. The cPCA of PPL in organic solvents: The percentage of fluctuation (-o-) and eigenvalues (---) of PPL-Cl (black) and PPL-O (red) in ethanol (A), octanol (B), toluene (C) and octanol + crystal-water (D).

Table 3. RMSIP calculated using the first ten eigenvectors

Variants	Cartesian space	Dihedral space
PPL^{+colip} (water)	0.612	0.590
PPL-O (water)	0.523	0.509
PPL-Cl (water)	0.463	0.537
PPL-O (ethanol)	0.493	0.550
PPL-Cl (ethanol)	0.521	0.555
PPL-O (toluene)	0.698	0.616
PPL-Cl (toluene)	0.688	0.564
PPL-O (octanol)	0.594	0.637
PPL-O (ocatnol + crystal-water)	0.569	0.619
PPL-Cl (octanol)	0.546	0.593
PPL-Cl (octanol +crystal-water)	0.595	0.613

Moreover, the RMSIP were calculated between the simulations carried under different conditions using the first ten eigenvectors obtained from cPCA and dPCA as well (Table 4). The lower values (≤ 0.4) suggested that the dynamics of PPL under different solvent conditions are significantly different. In other words, the major changes under these solvent conditions could be identified from the time-space used for the simulations. It was also noted that the RMSIP of dPCA showed higher dissimilarity among the simulations.

Table 4. RMSIP between the simulations calculated using the first ten eigenvectors[¶] of PPL in different solvents[§]

	A	B	C	D	E	F	G	H	I	J	K
A	-	0.299	0.186	0.413	0.196	0.350	0.213	0.358	0.371	0.192	0.197
B	0.485	-	0.158	0.299	0.152	0.237	0.143	0.194	0.197	0.119	0.111
C	0.420	0.397	-	0.210	0.437	0.169	0.301	0.152	0.150	0.242	0.244
D	0.427	0.411	0.386	-	0.221	0.354	0.203	0.330	0.315	0.177	0.174
E	0.406	0.377	0.402	0.382	-	0.184	0.345	0.174	0.171	0.280	0.282
F	0.369	0.356	0.338	0.353	0.342	-	0.168	0.259	0.248	0.147	0.144
G	0.358	0.348	0.347	0.353	0.350	0.329	-	0.228	0.223	0.338	0.355
H	0.436	0.402	0.378	0.401	0.389	0.367	0.354	-	0.392	0.217	0.235
I	0.433	0.406	0.373	0.410	0.375	0.349	0.338	0.418	-	0.215	0.231
J	0.335	0.325	0.334	0.328	0.352	0.303	0.325	0.332	0.318	-	0.344
K	0.383	0.362	0.381	0.358	0.399	0.333	0.346	0.385	0.376	0.339	-

[¶] Plain and grey shaded values were calculated from eigenvectors obtained from Cartesian and dihedral spaces, respectively.

[§] A, PPL+colip (water); B, PPL-O (water); C, PPL-Cl (water); D, PPL-O (ethanol); E, PPL-Cl (ethanol); F, PPL-O (toluene); G, PPL-Cl (toluene); H, PPL-O (octanol); I, PPL-O (octanol + crystal-water); J, PPL-Cl (octanol); and K, PPL-Cl (octanol +crystal-water).

3.4.5 Free Energy Surface of PPL_{mut}-Cl and Lid Opening Dynamics

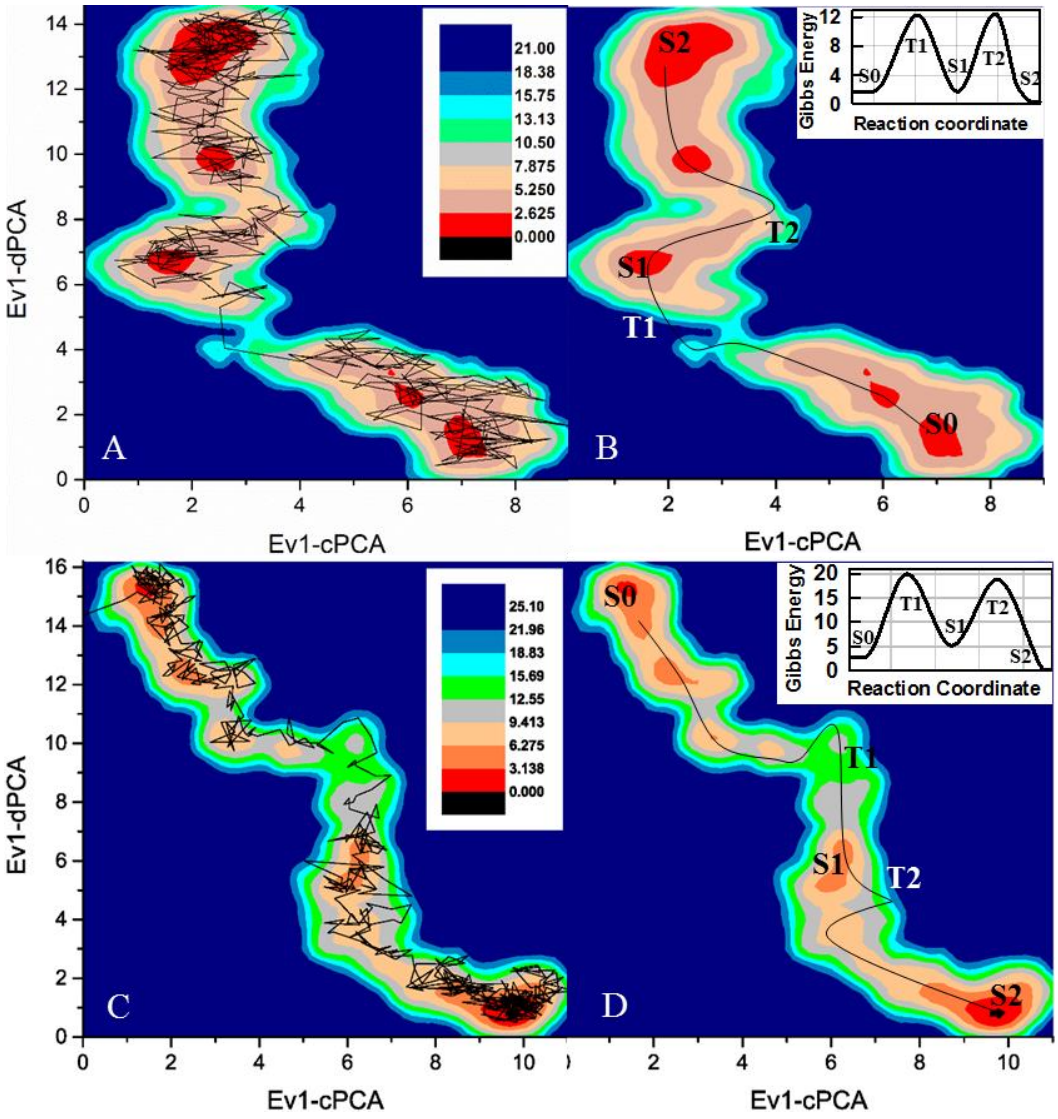


Figure 14. Free energy landscape of PPL_{mut}-Cl in water (A) and octanol (C) constructed by projection of corresponding first eigenvectors of dPCA against cPCA with the trajectory of movement in black lines. The most probable pathway with major transitions in water (B) and octanol (D) are identified. The Gibbs energy (in kJ/mol) changes along the reaction coordinates calculated in water and octanol are shown as insets in the respective panels. The color codes represent the energy levels.

The major conformational changes which led to opening of the lid in PPL_{mut}-Cl at 400 K in water and octanol were investigated by construction of free energy surface (FES) from their respective MD trajectory. The FES in both the solvents were derived by projection of first eigenvectors of cPCA and dPCA as the coordinates (Fig. 14). The structures representing energy minima were identified around 8.5, 21.9, and 48.0 ns (S0, S1 and S2, respectively) and energy maxima were identified around 18.4 and 35.0 ns (T1 and T2, respectively) along the simulation trajectory in water. Similarly, minima (S0, S1, and S2 at 3.5, 23.5, 46.2 ns, respectively) and maxima (T1 and T2 at 17.8 and 25.9 ns,

respectively) were identified in octanol also. In order to get more insight on the interactions, the existence of hydrogen bonds between the lid and rest of the protein were also analyzed (Fig. 15-17).

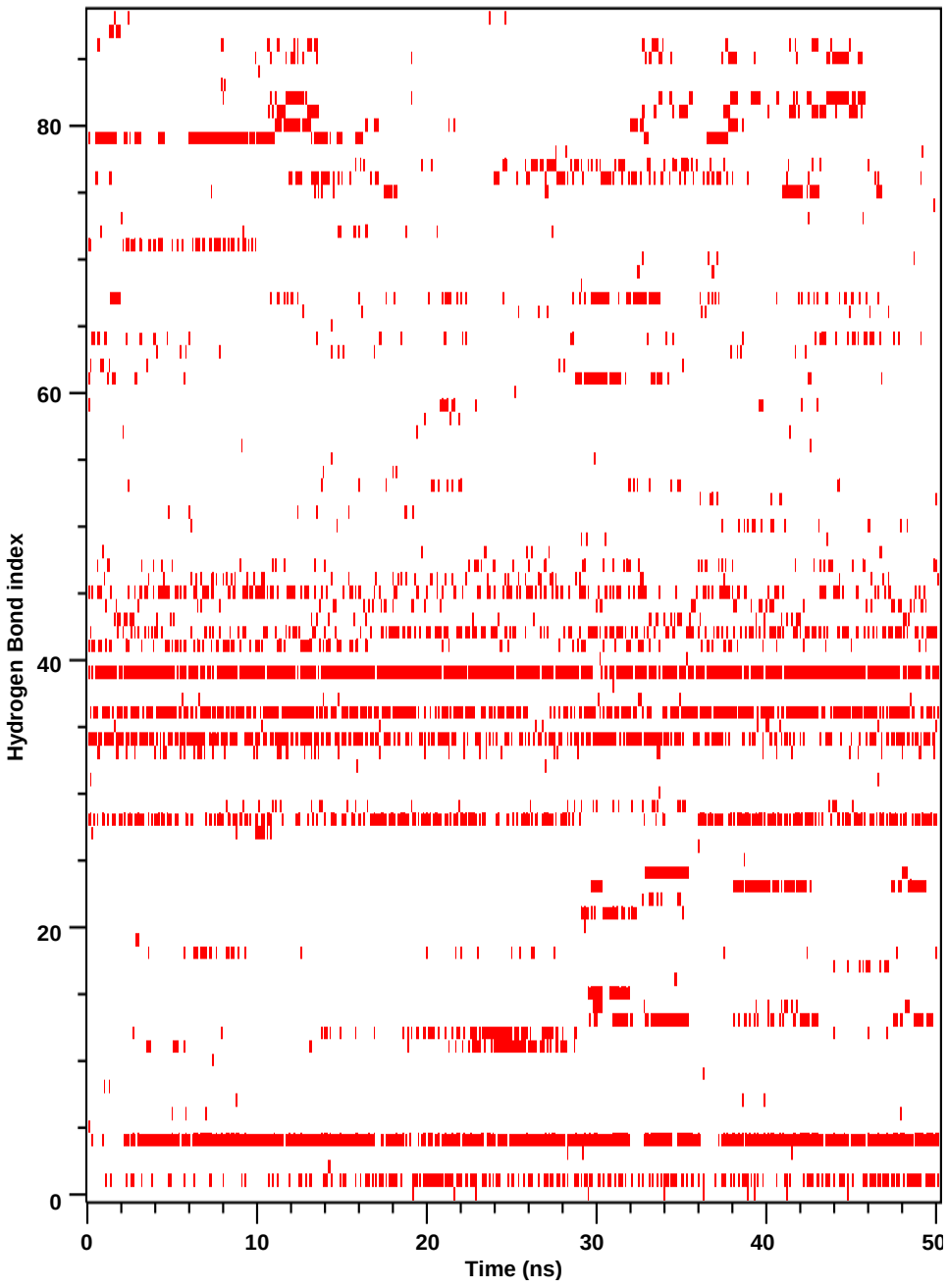


Figure 15. Existence map of H-bond interactions of lid residues with other residues in the protein during the simulation of PPL-Cl in water at 300 K. The presence of H-bond interactions are shown in red and the blank spaces represent the absence. The list of interacting partners are given in Appendix A1.

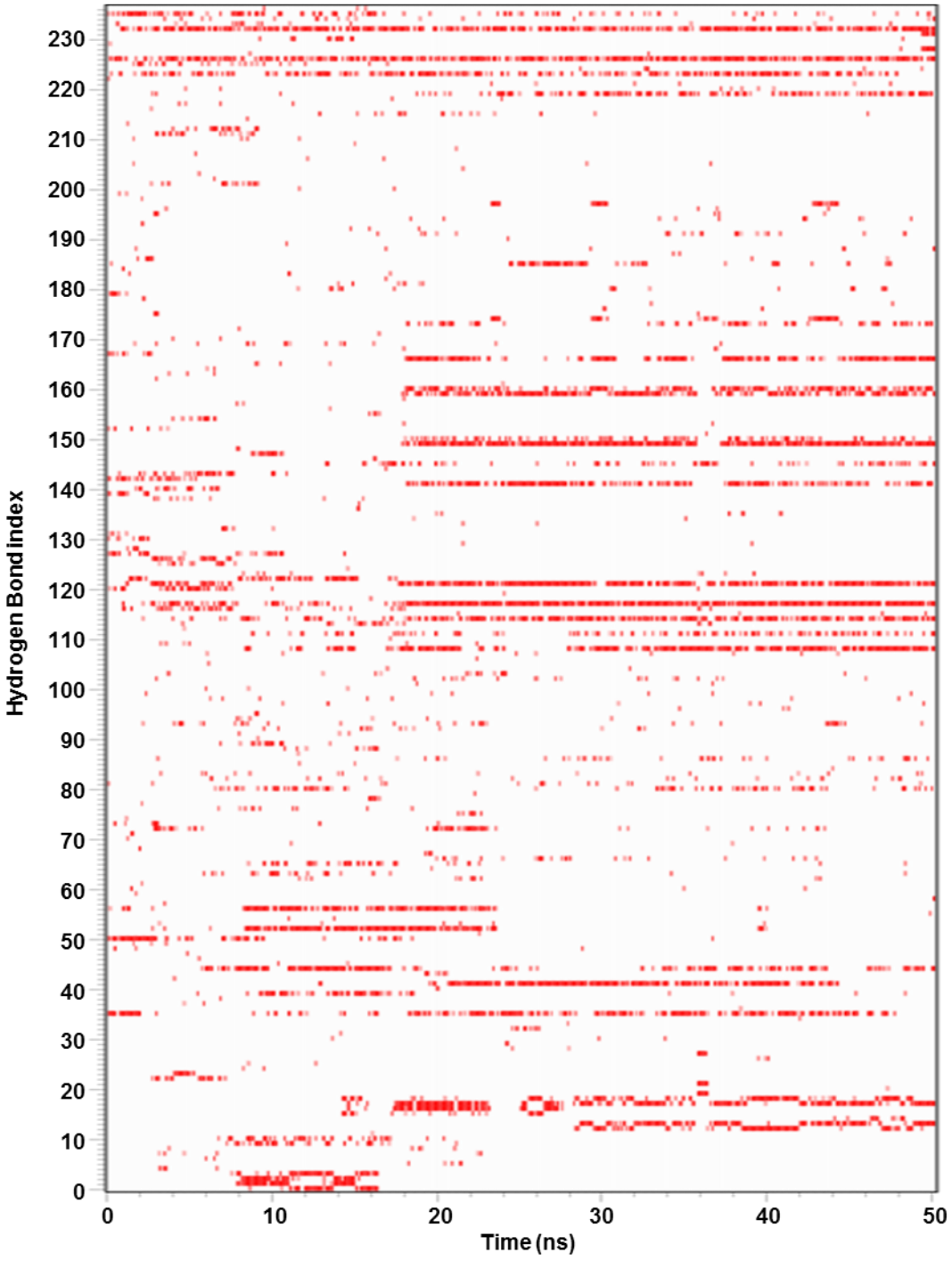


Figure 16. Existence map of H-bond interactions of lid residues with other residues in the protein during the simulation of PPL_{mut}-Cl in water at 400 K. The presence of H-bond interactions are shown in red and the blank spaces represent the absence. The list of interacting partners are given in Appendix A2.

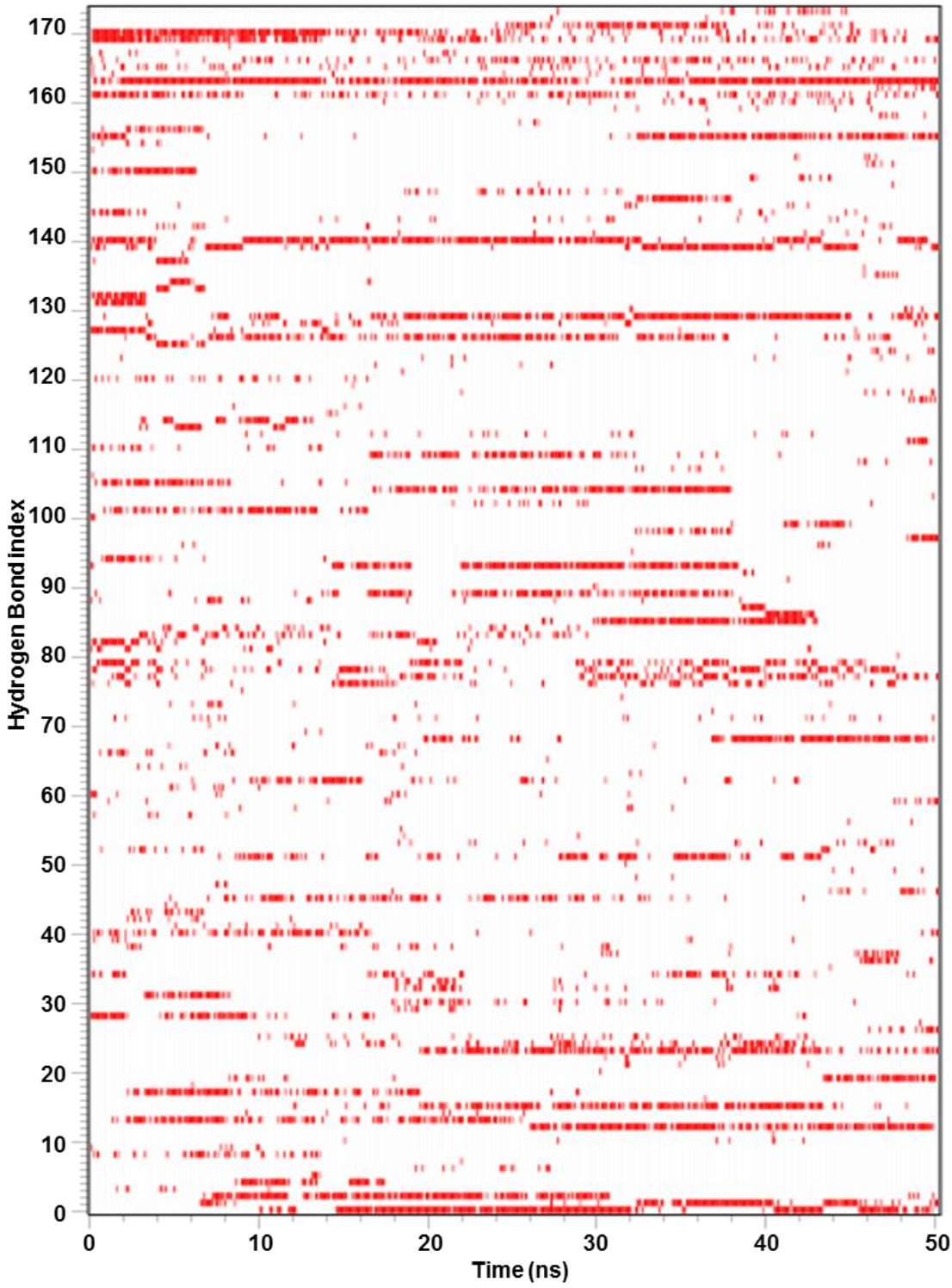


Figure 17. Existence map of H-bond interactions of lid residues with other residues in the protein during the simulation of PPL_{mut}-Cl in octanol at 400 K. The presence of H-bond interactions are shown in red and the blank spaces represent the absence. The list of interacting partners are given in Appendix A3.

3.5 Discussion

3.5.1 Basic Trajectory Analysis

The dynamics of PPL was followed in water and three non-aqueous solvent conditions in both closed and open states to probe solvent-induced conformational changes at atomistic-detail. The results show that PPL in water and ethanol has more fluctuations in both closed and open conformations. Except in the case of PPL-O in water (previous chapter), in all other conditions the major fluctuations are observed in domain 2, particularly around the surface-exposed β5' loop. The distance between two domains of PPL calculated in these conditions clearly indicates that the increased RMSD could be attributed to the domain movements rather than any major conformational changes around the lid region (Fig. 18). Thus, these conformational changes could not transfer PPL-Cl to PPL-O or *vice versa*. Also, in the case of PPL-O in toluene, the domain movements contributed for the increase in fluctuations around domain 2. This was evident from the changes observed in R_g and SASA values for PPL-O in toluene (Fig. 1 and 3). In the presence of octanol, the dynamics of both PPL-Cl and PPL-O are reduced; even inclusion of the water molecules (octanol + crystal-water) could not increase the flexibility. These observations are supported by the changes observed in R_g and SASA values. The packing density of the protein in different solvents also complements these observations. PPL in the solvents with higher SASA has lower packing density (Fig. 3C-D). Moreover, the conformational entropy of PPL predicts that the open state has constrained motions compared to the closed state in all the solvents, except in the case of PPL-O in toluene where domains are moving closer to each other during the simulation time. It is also notable that the conformational entropy decreases with decreasing polarity of the solvent and even the inclusion of crystal water molecules in the presence of octanol did not increase the entropy. These results suggest that the protein flexibility is globally restrained, as polarity of the solvent decreases.

Though PPL-O could show lid closing movement in water, the lid opening movements are not observed for PPL in any of the solvents at ambient conditions. To further explore the conformational changes involved during the lid opening, dynamics of a double mutant of PPL (D250V and E254L) in closed conformation was studied in all the four solvents at ambient conditions. Comparison of RMSD and R_g values of PPL_{mut}-Cl (Fig. 8) with PPL-Cl suggests that the overall fluctuations (RMSD) are not considerably altered by the mutation in water, toluene and octanol. Only in ethanol, increase in RMSD and radius of gyration could be noticed. The RMSF values of PPL_{mut}-Cl propose that the observed increase in fluctuations might be due to the changes around the lid region of the protein in ethanol. However, these changes could not initiate lid opening of PPL_{mut}-Cl (Fig. 9).

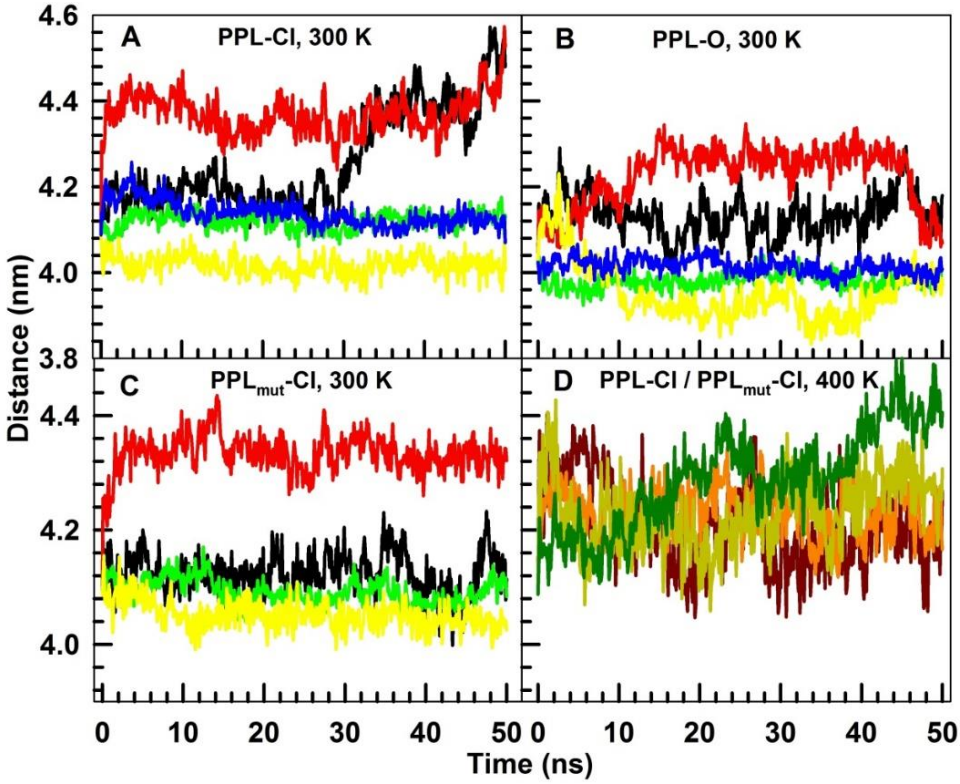


Figure 18. Distance between domain 1 and domain 2 during the simulations run at 300K for (A) PPL-Cl, (B) PPL-O, and (C) PPL$_{mut}$-Cl in water (black), ethanol (red), octanol (green), toluene (yellow) and octanol with crystal water (blue). (D) Domain-domain distance during the simulations run at 400K for PPL-Cl in water (brown) & octanol (olive) and for PPL$_{mut}$-Cl at 400K in water (orange) & octanol (green).

In order to facilitate the lid opening movements of PPL-Cl, MD simulation was performed at elevated temperature (400 K) in two extreme solvent conditions, i.e., in water and in octanol. The RMSD of PPL-Cl shows increase in water, but not in octanol (Fig. 10). The increased RMSD in water might be due to the fluctuations around lid region and domain 2 which is evident from RMSF values of corresponding amino acids (Fig. 10). However, these conformational changes could not lead to lid opening (Fig. 11 A-B). These results suggest that even at higher temperature, the interactions between D250-Y115, and D250-E254 do not allow the lid opening in PPL-Cl (Fig. 6 and 7). Therefore, the dynamics of double mutant of PPL-Cl (PPL$_{mut}$-Cl) was carried out at 400K in water and octanol which showed movement of lid from its closed conformation towards open conformation (Fig. 11C-D). At higher temperature, PPL$_{mut}$-Cl shows slight decrease in RMSD in water whereas in octanol it marginally increases. In water, R_g and SASA values are not altered by mutation whereas R_g increases in octanol (Fig. 10). Comparing these results with RMSF values suggest that PPL-Cl in water possesses more flexibility around domain 2 and at the region connecting the lid with domain 2, which are highly constrained in the open conformation of the protein. Major fluctuations are found in the regions around helix 12 and the loop connecting helix 13 with sheet 11 which are at the bottom of the lid along with the loop at the end of *C*-terminal domain of the protein.

These fluctuations could increase RMSD and R_g values of PPL-Cl. At the same time, the dynamics of domain 2 is highly restrained in octanol.

In case of PPL_{mut}-Cl, most of these fluctuations are reduced and R_g and SASA values are comparable to PPL-O (Fig. 10). However, the mutations led to increased fluctuations around G114 (*N*-terminal cap of helix 4) which interacts with the side chain of D250 and the region around the *C*-terminal cap of helix 4 probably due to loss of helicity. In octanol, PPL_{mut}-Cl shows maximum fluctuations around the loop connecting the domains and the residues at the end of *C*-terminal. At higher temperature, the major fluctuations are noted around helices 3 and 4, particularly on Gly residues (19, 82, 110 and 141) along with the lid region. Further, the mutation caused larger residue fluctuations on PPL_{mut}-Cl in octanol compared to water which resulted in increase in R_g and SASA values of PPL_{mut}-Cl in octanol at higher temperature.

3.5.2 Lid opening Dynamics of PPL_{mut}-Cl in Water

From the projection of first eigenvector of dPCA against cPCA obtained from MD trajectory of PPL_{mut}-Cl at 400 K, at least three minimal energy conformations were identified. The conformational changes between these transitions were analyzed (Fig. 14 A and B).

3.5.2.1 Step 1 - SO –T1- S1: (8.5 ns – 17.1 ns – 18.3 ns)

The initial conformation of PPL_{mut}-Cl shows significant changes at the earlier stages of the simulation. This could be attributed to increase in simulation temperature introduced during production simulation. The further changes are examined from the first minimal energy state, S0 which is around 8.4 ns. The initiation of lid movement probably starts with the loss of hydrogen bond between Asp248 on the lid and Arg23 on domain 1 which could be the major energy barrier between S0 and T1. This leads to movement of the strands at *N*-terminal end away from the lid which is supported by movements of (cap) residues in helix 3 and helix 4. However, β5-loop simultaneously moves along with the lid during the first step. This could be due to the strong interactions between Lys81 on the loop and Asp258 on the lid and the possible interaction between Asp258 with Arg112 which is spatially nearer to β5-loop and lid. During these movements at the top of the lid, conformational changes at the neck of the lid are also observed, particularly loss of hydrogen bonds between Pro212 and S244 of lid. These changes are associated with small conformational changes at the loop connecting two domains and other loop regions on both the domains.

3.5.2.2 Step 2- S1 to S2: (18.3 ns – 26 ns – 35 ns)

The further lid movement is probably facilitated by breaking the salt bridge between K81 and D258. As this energy barrier is crossed (T2), the lid moves further away

from domain 1 and reaches more opened conformation. These changes are accompanied by conformational transitions around the helices 3 and 4 and the loop region connecting these helices. It is also observed that these conformational transitions are accompanied by loss of hydrogen bonding interaction between Asn295 and Ser384 at the interface of the domains.

3.5.3 Lid opening Dynamics of PPL_{mut}-Cl in Octanol

Using the first eigenvectors of dPCA and cPCA, derived from MD trajectory of PPL_{mut}-Cl in octanol at 400 K, FES was constructed. From the FES, at least three minimal energy conformations were identified and the transitions between these states were analyzed (Fig. 14C-D).

3.5.3.1 Step 1- S0 to S1- SO –T1- S1: (3.5 ns – 18 ns – 23.5 ns)

The first minimal energy conformation of PPL_{mut}-Cl is identified around 3.5 ns. In the first major transition, the polar interaction of Asp85 with Arg92 is lost and it establishes an interaction with Arg265 which is nearer to the terminal of lid. Moreover, Lys81 which is part of β5-loop forms polar interaction with Asp258 on the lid. These interactions moved the lid away from domain 1 along with the loss of interaction between Glu188 and Arg191 in the β-turn surrounding the active site. This is associated with the movement of β2-loop and helix 1 at *N*-terminal region. A large amplitude fluctuation is also found on domain 2 and the interfacial residues between the domains.

3.5.3.2 Step 2- S1 to S2: (23.5 ns – 26 ns – 46 ns)

During this transition, the lid moves further away from domain 1 which is facilitated by loss of polar interaction between Glu83 on β5-loop and K269 at helix 11 which is part of the lid domain. However, the interaction between Lys81 and Asp258 is found to be intact during these conformational changes. These observations suggest that the rearrangements of interactions between β5-loop and the terminal region of the lid plays vital role in lid opening dynamics of PPL_{mut}-Cl.

3.5.4 Comparison of Lid Opening Pathways in Water and Octanol.

Mutation and the raise in temperature altered the conformation of PPL-Cl in water and octanol. In water, major conformational changes are observed around the lid region including the loss of helical structures and movement of lid and β2-loop closer to each other. In case of octanol, major conformational changes occur around the lid and domain 2. Also, the lid opening pathway is found to be different in both the cases. Disruption of polar interaction between R23 and D248 was the major initial conformational barrier in water; however, the interaction between D85 and R92 is identified to be the major barrier in octanol. Further, K81 loses its interaction with D258 in water during the lid opening whereas in octanol this interaction is relatively stable during the dynamics. Notably, β10-

loop connected to the *N*-terminal of lid does not show much conformational changes during the lid opening whereas the helix connected to the *C*-terminal of lid shows large fluctuations in octanol, but not in water. The extent of lid opening, measured as a distance between tip of the lid (residues 250-254) and the active site triad, suggests that the lid opening movement is larger in octanol than in water (Fig. 19).

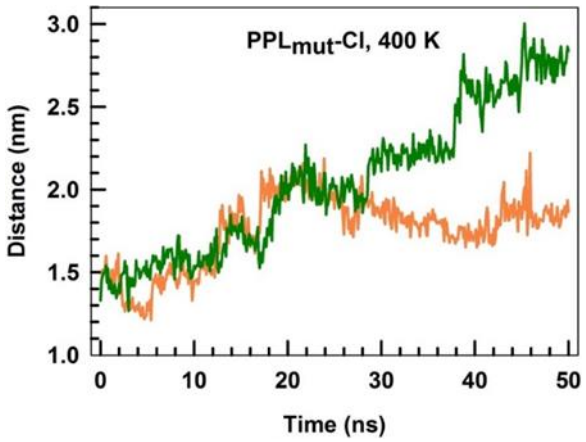


Figure 19. Distance of the tip of lid (residues 250-254) from the active site triad during the simulation of PPL_{mut}-Cl at 400 K in water (orange) and octanol (green).

3.5.5 Active Site Conformation in Different Solvent Conditions.

The conformational changes of residues belong to the active site triad, in PPL during different simulations, were examined by extracting their side chain conformations at the interval of every 10 ns (Fig. 20) and the distribution of backbone and side chain dihedral angles (Fig. 21). The dihedral angles of active site residues are altered in the cases of PPL_{mut}-Cl at 400 K and PPL^{+colip}-O, otherwise in all the other conditions they are not affected. In case of PPL^{+colip}-O, only the side chain dihedral angles of Ser153 and His264 are different from other conditions whereas in PPL_{mut}-Cl at 400 K the backbone dihedrals of Ser153 are altered. However, the difference in conformations of active site residues in closed and open conformations are minimal. This suggests that during lid opening movement backbone dihedrals of active site residues may be slightly altered; however, as PPL reaches completely open state, they would attain the similar conformations as in the closed state.

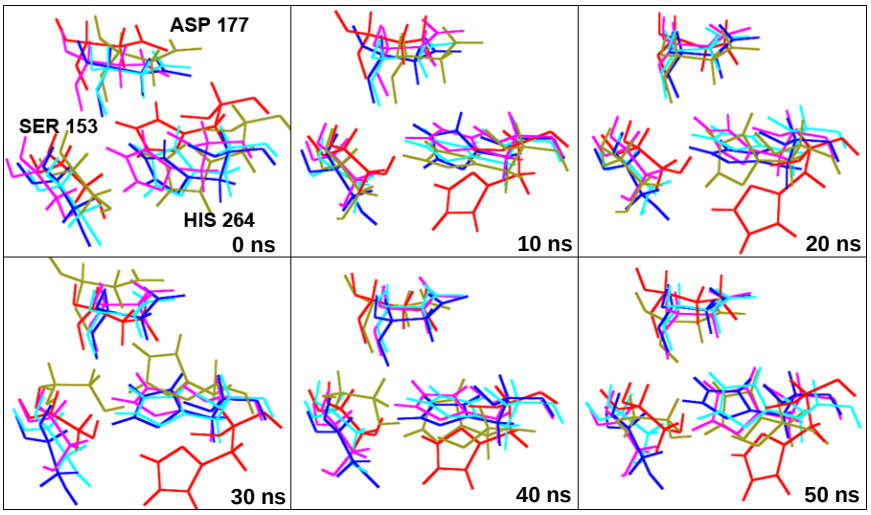


Figure 20. Conformational changes of active site residues Ser153, Asp177, and His264 during the simulation of PPL-O monitored at the interval of each 10 ns in ethanol (blue), octanol (magenta), and octanol + crystal-water (cyan). Also, for PPL_{mut}-Cl in octanol at 400 K (olive) and PPL+colip-O in water at 300 K (red) are presented.

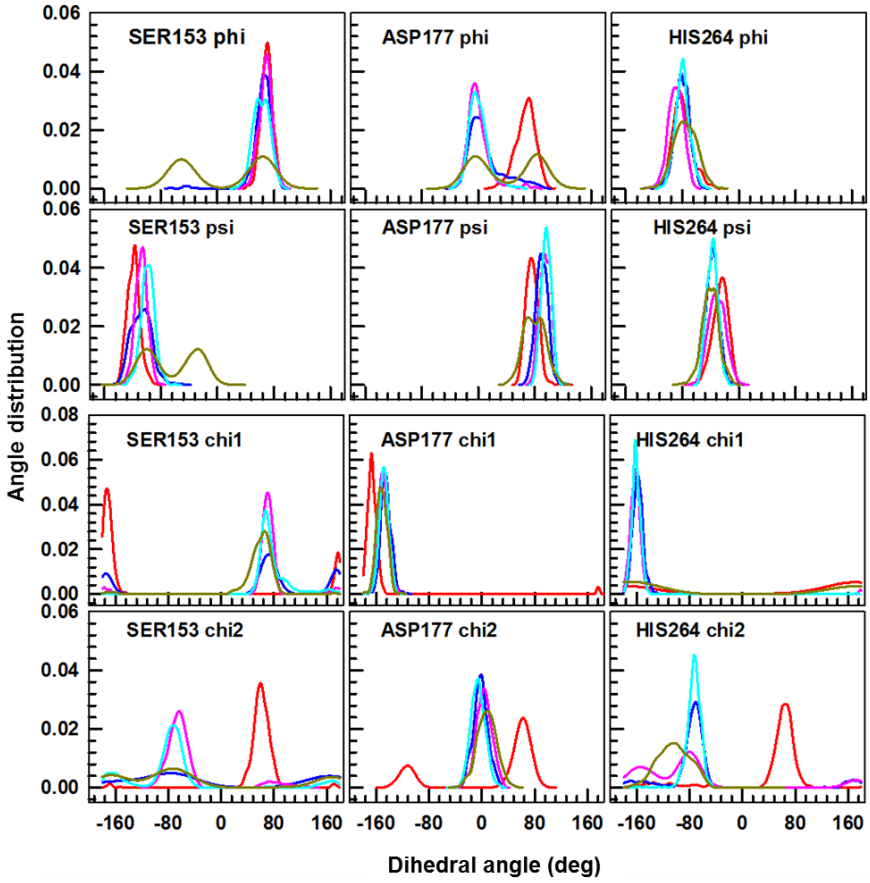


Figure 21. Distribution of backbone (phi and psi) and side chain (chi1 and chi2) dihedral angles of active site residues Ser153, Asp177 and His264 during MD simulation of PPL^{+colip}-O in water (red), PPL-O in ethanol (blue), PPL-O in octanol (magenta), PPL-O in octanol + crystal-water (cyan), and PPL_{mut}-Cl in octanol at 400 K (olive).

3.5.6 Mechanism of Lid Closure in Different Lipases.

To the best of our knowledge, this is the first atomistic-level study on lid opening mechanism of a pancreatic lipase. However, the lid opening movements of microbial lipases which also have lid conformation-dependent activity and same structural features as *N*-terminal domain of PPL have been studied using MD simulation. In general, the studies have shown that the lid moves towards open conformation in the presence of non-polar solvents and at oil/water interfaces [176, 208, 209]. In some lipases, such as from *Candida Rugosa,* slight opening of lid in water is also observed [176]. It is also found from both experimental and computational studies that the extent of opening and the stability of lid vary with solvent polarity [176, 207-210]. Further, disruption of salt bridges between the charged residues are found to be the major factor at the initial movements of lid opening in most of the lipases [175, 210, 211], as observed here in the case of PPL as well. Though some of the lipases, such as *Rhizomucor miehei*, show rigid-body movement [176], *Yarrowia lipolytica* [208], *Candida rugosa* [176], and *Burkholderia cepacia* [175] lipases undergo large conformational changes and formation of new interactions during lid opening which are not found in either of the extreme conformations. These results suggest that even the lipases with very close sequence and structural similarity might adapt different lid opening pathways.

3.6 Conclusion

MD simulation of PPL was carried in water, ethanol, octanol, toluene and octanol + crystal-water to monitor the lid movement. In all the cases, the fluctuations observed around the lid region were minimal. Though major fluctuations were noted on PPL-Cl in the presence of water and ethanol which was associated with increase in R_g and SASA values, it could not induce the lid opening movement. Also, larger conformational changes were noted on PPL-O in toluene mostly due to domain-domain movement. To further explore the lid opening conformation of PPL-Cl, simulations were carried out at higher temperature (400 K), however, it could not result in lid opening movement. Having broken the interactions between D250-Y115, and D250-E254 by a double mutant (PPL_{mut}-Cl) and raising the simulation temperature to 400 K, significant conformational changes were observed around the lid region which was moving away from domain 1 in water and octanol. The analysis of FES of these simulations suggested that the lid opening pathway in water and octanol are not similar. Also, the lid was found to be in more open conformation in octanol compared to water suggesting that non-polar environment could facilitate lid opening dynamics compare to aqueous environment.

Chapter 4

Role of Bile Salt in the Formation and Stabilization of its Ternary Complex with Pancreatic Lipase and Colipase

4.1 Abstract

Mammalian pancreatic lipases require bile salt along with colipase for optimum catalysis. Bile salt is also known to activate the pancreatic lipases in the absence of colipase at very low concentration. Near and above the micellar concentrations, bile salt inhibits the activity which can be reverted back by colipase. All atom molecular dynamics simulation has been implemented to study the role of monomeric and micellar sodium taurocholate (NaTC, bile salt) in the stabilization of the active state of porcine pancreatic lipase (PPL). An exposure of hydrophilic residues was noted at sub and supra-micellar concentration of NaTC. At supra-micellar concentration of NaTC, colipase showed considerable decrease in radius of gyration and the hydrophilic solvent accessible surface area. Monomeric and oligomeric NaTC were present throughout the surface of the protein and the complex. Micellar aggregates were either in solvent or interacting to rear side of the colipase. Lid was prevented from denaturation at all concentration of NaTC and maximum stabilization was observed in lipase-colipase complex at supra-micellar concentration. By superposing the structures along with NaTC with respect to each other, a plausible model was generated. Lipase-colipase complex makes a hydrophobic plateau, which includes hydrophobic tip of colipase fingers, lid in the open conformation and a few residues of the *N*-terminal domain of PPL. The hydrophobic plateau might be efficient substrate binding site. The growth of the micellar aggregate at the rear side of the fourth finger of the colipase might provide additional stability to the complex. The *C*-terminal was not found close enough to the micelle, hence its role in stabilizing the complex could not be affirmed with NaTC as bile salt.

Abbreviations used in the chapter:

NaTC, Sodium taurocholate; TcL, NaTC at low concentration; TcH, NaTC at high concentration; PPL, Porcine pancreatic lipase in the absence of colipase and NaTC; PPL^{+colip}, Lipase-colipase complex in the absence of NaTC; PPL^{+TcL}, Lipase in the presence of TcL; PPL^{+TcH}, Lipase in the presence of TcH; Colip, Colipase in the absence of NaTC and lipase; $Colip^{+TcL}$, Colipase in the presence of TcL; $Colip^{+TcH}$, Colipase in the presence of TcH; $PPL^{+colip+TcL}$, Lipase-colipase complex in the presence of TcL; $PPL^{+colip+TcH}$, Lipase-colipase complex in the presence of TcH; PPL^{-colip}, Lipase chain of lipase-colipase complex in the absence of NaTC; $PPL^{-colip+TcL}$, Lipase of lipase-colipase complex in the presence of TcL; $PPL^{-colip+TcH}$, Lipase chain of lipase-colipase complex in the presence of TcH; PPL-CL, Lipase in closed conformation in the absence of colipase and NaTC.

4.2 Introduction

Proteins in cells work in association with other biological macromolecules or biomolecular assemblies, unlike the proteins in a test tube. All the essential processes in cells are precisely controlled by protein-protein and/or protein-nucleic acid interactions [220-222]. Protein-protein interaction (PPI) has a major role to play in cell signaling and cell cycle regulation [223-225]. The biological processes involved in cell functioning and their regulations have so efficiently evolved such that almost all the activities and their feedback controls are governed by PPIs [220-225]. It has been always an intriguing question how the PPIs are able to tune the high signal to noise ratio, in the intra- and inter-cellular communications [226-228]. Any disruption or destabilization of PPIs and their networks may cause significant consequences to the cellular processes like apoptosis, metabolism, cell growth and cell division. Disturbances in these events might cause diseases like Marfan-like connective tissue disorder [229], Retinitis pigmentosa [230], Nephronophthisis [231], Alzheimer's disease, AIDS, diabetes [232], Branchiootic syndrome [233], Polycystic kidney disease [234] and cancer [235, 236] etc.

The disruptions and fine-tuned control in the PPIs can be achieved by numerous ways. In schizophrenia, mutation in DISC1 gene, sequesters it from interaction with crucial protein [237] and the mutation at the interface of PPI has shown to have significant role in many cancer [236]. The synaptic transmission mediated through syntaxin based PPI is inhibited by the use of botulinum toxin [238] and many small molecules act to influence the PPI [239-241]. In the normal cell functioning these interactions are controlled by the processes like phosphorylation and dephosphorylation with the help of several kinases and phosphatases, respectively [242-244] and sumoylation, acetylation, sulfonation, glycosylation, neddylation, [245-249] etc.

In this report, we have investigated a special type of protein-protein interaction which controls the activity of porcine pancreatic lipase (PPL) using molecular dynamic (MD) simulation. PPL non-covalently binds to its co-enzyme called colipase. This PPI gains its uniqueness by the requirement of surfactant for efficient catalysis of PPL in the presence of colipase. PPL has a short helical segment in the *N*-terminal region called lid which controls the accessibility of the active site by substrates for catalysis. *C*-terminal of the PPL and lid region in the *N*-terminal assists the binding of colipase in proper orientation. Lipases are ubiquitous across bacteria, fungi and mammals. They all share certain common features like the presence of lid, interfacial activation and stability in organic solvents [250-252]. Lipases in lower organisms do not have *C*-terminal homologous to PPL and also do not require co-enzyme to stabilize the lid in its open conformation. However, in higher organisms the conformational state of the lid is controlled by binding of colipase and non-peptide molecules [253, 254].

PPL-colipase interaction is not strong and tend to dissociate [132, 255] in the absence of micelle of bile salt. Bile salts are surfactant-like amphipathic molecules which play crucial role in stabilizing PPL-colipase interaction. Exact role and mode of bile salt interaction with PPL-colipase complex is yet unclear. However, it has been reported from dialysis experiments that a single micelle of bile salt binds to one colipase in the absence of PPL [255-258]. At very low concentration (far below CMC) of bile salt, in the absence of colipase, slight activation of PPL is observed [259-261]. However, further increase in concentration (near CMC), bile salt inhibits the catalytic reaction. The activity is completely inhibited at above the CMC of the bile salt. This inhibitory effect is reverted by the addition of colipase at all the concentrations of bile salt [129, 132, 259-261]. The absence of inhibition at higher concentration of bile salt could be due to the formation of 1:1:1 ternary complex of PPL-colipase-bile salt micelle [132, 262]. With the help of small-angle neutron scattering (SANS) experiments a model has been proposed [132, 262-264] for the ternary complex with the bile salt micelle at the concave junction made by the colipase and C-terminal domain of the PPL. Nevertheless, atomic resolution model and residue-level interactions of bile salt with lipase-colipase is still elusive. Here, all atom molecular dynamic (MD) simulation has been used to reconstruct the ternary complex with the help of experimental results available about the each component of the enzyme complex. The results are well in agreement with the models proposed by Pignol *et al* [262] and Hermoso *et al* [263], with only slight differences in the position of the bile salt micelle.

4.3 Experimental Methods

4.3.1 Structures and Parameters

X-ray crystal structure of PPL-colipase (PPL^{+colip}) complex was obtained from protein database (PDB Id: 1ETH) [132] after removing co-crystalized water and other ligands. PPL in open conformation was created by removing colipase from PPL^{+colip}. NMR-solution structure (PDB Id: 1PCN) of colipase was also used for MD simulation [265]. Conjugated trihydroxy bile salt, sodium taurocholate (NaTC) was used in all the simulations in the physiological concentration of sodium chloride (0.15 M). Parameter for NaTC was obtained from Automated Topology Builder v2.0 [266]. Two different concentrations of NaTC were used, 12 molecules of NaTC for lower concentration and 60 molecules for higher concentration, in a simulation box of 1000 nm^3.

4.3.2 Molecular Dynamic (MD) Simulation and Data Analysis

All MD simulations were performed on GROMACS 4.6.3 [52, 149-151], using the GROMOS 54A7 force field and SPC/E water model. Cut-off and simulation schemes were used as mentioned elsewhere or in Chapter 2 [253]. NaTC was simulated in the presence and absence of proteins at lower and higher concentration. Lipase, lipase-colipase complex and colipase were simulated in the presence (both low and high

concentration) and absence of NaTC. All the simulations were performed for 50 ns. The simulations are listed in Table 1. The data analyses were performed using GROMACS tools, UNIX shell script and in-house R-codes. PyMOL[155] was used for structure visualization and obtaining trajectory snapshot and image capture.

Table 1. List of simulation carried out

No	System name	NaTC conc	NaTC number	Time (ns)
1	PPL+colip	-	0	50
2	PPL+colip+TcL	Low	12	50
3	PPL+colip+TcH	High	60	50
4	PPL	-	0	50
5	PPL+TcL	Low	12	50
6	PPL+colip+TcH	High	60	50
7	Colip	-	0	50
8	Colip+TcL	Low	12	50
9	Colip+TcH	High	60	50
10	Colip-NMR+TcL	Low	12	50
11	Colip-NMR+TcH	High	60	50
12	TcL	Low	12	50
13	TcH	High	60	50

4.4 Results

4.4.1 Trajectory Analysis

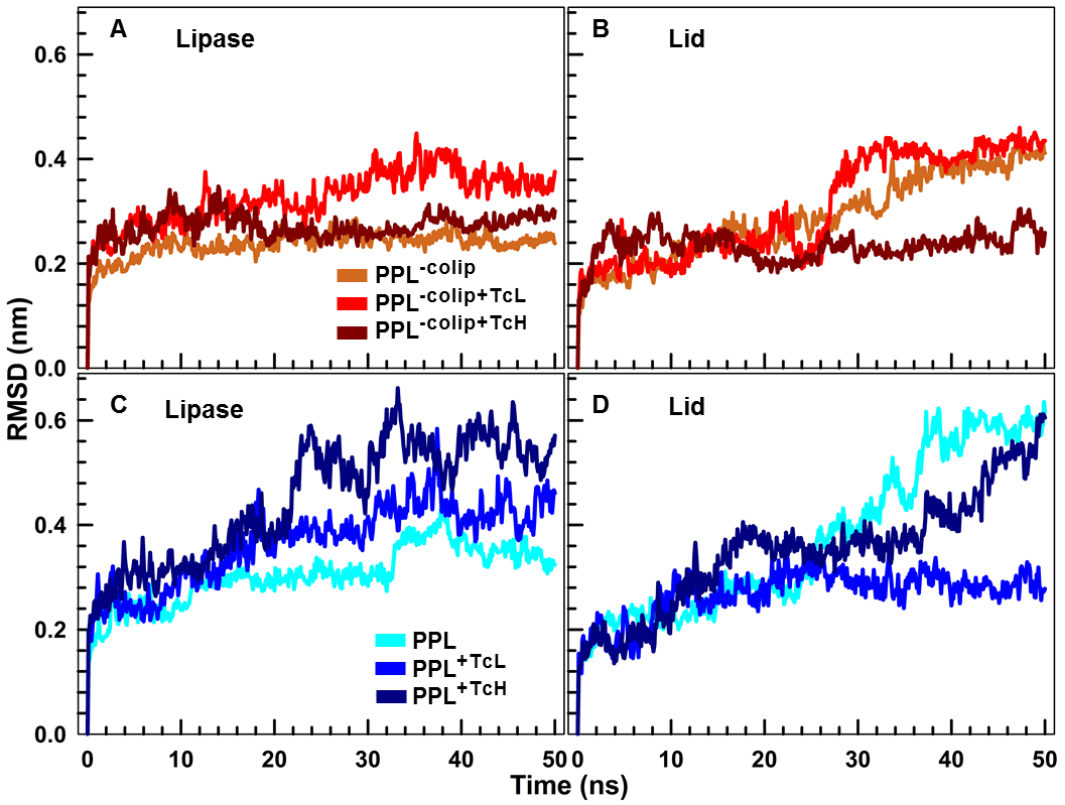


Figure 1. RMSD of all the C_α-atoms of the lipase chain (A & C) and only the lid region (B & D) in the presence (PPL^{-colip}) and the absence (PPL) of colipase and in the presence of low (TcL) or high (TcH) concentration of bile salt.

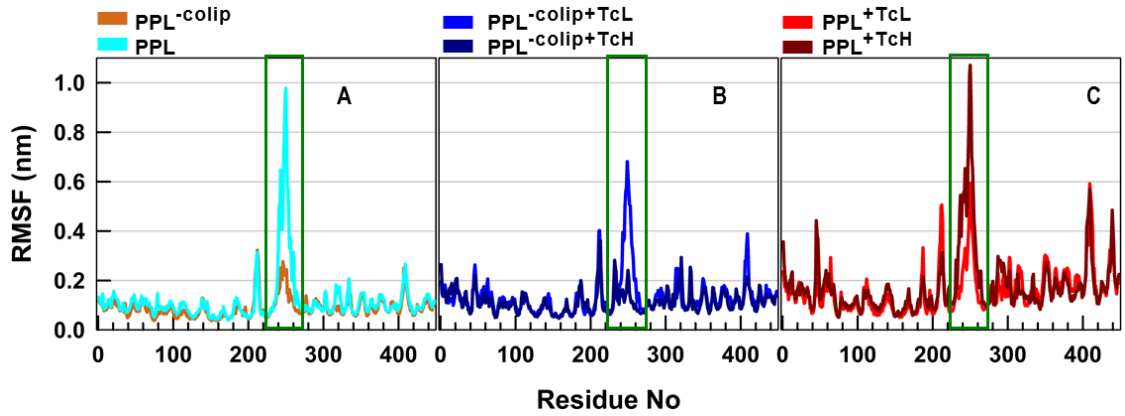


Figure 2. RMSF of C_α-atoms of (A) lipase (PPL) and lipase-colipase complex (PPL^{-colip}), (B) lipase-colipase complex in low ($PPL^{-colip+TcL}$) and high ($PPL^{-colip+TcH}$) concentrations of NaTC, and (C) lipase, in the absence of colipase, in low (PPL^{+TcL}) and high (PPL^{+TcH}) concentration of NaTC. The lid region is marked by green rectangular boxes in all the panels.

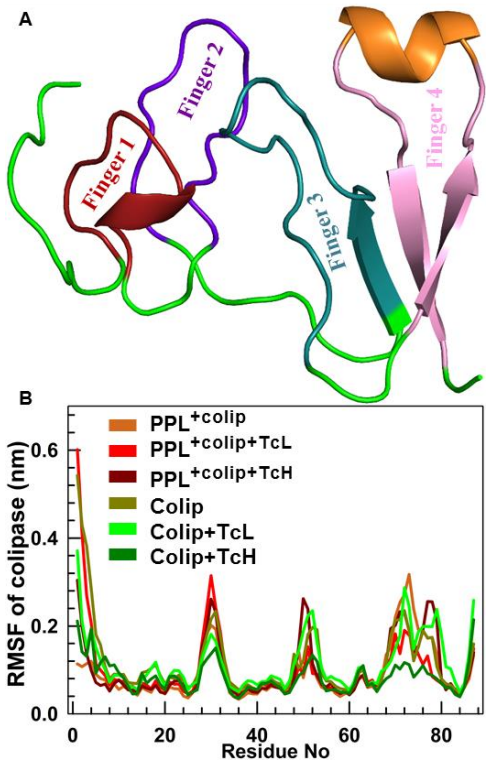


Figure 3. (A) Cartoon diagram of colipase obtained from the crystal structure of lipase-colipase complex (PDB id:1ETH). Four protruding fingers are marked as finger 1, finger 2, finger 3 and finger 4 corresponding to the amino acid residues 14 to 23, 27 to 39, 47 to 64, and 68 to 88, respectively. (B) RMSF of colipase calculated from the simulations performed in the presence and the absence of lipase chain (PPL) and in low (TcL) or high (TcH) concentrations of NaTC (TcL and TcH).

Root mean square deviation (RMSD) of lipase (Fig. 1) was found to be more in the presence of NaTC (PPL^{+TcL}, PPL^{+TcH}, $PPL^{-colip+TcL}$, $PPL^{-colip+TcH}$) than its absence. RMSD of PPL^{-colip} was found to be the least among the simulated trajectories of lipase whereas the maximum fluctuation in the structure was observed in for the protein in the presence of high concentration of NaTC (PPL^{+TcH}). The RMSD analysis of the lid region alone suggested that in the absence of colipase, lid had larger fluctuations without NaTC followed by that in the presence of high (PPL^{+TcH}) and low (PPL^{+TcL}) concentrations of NaTC. The lid region of the protein, in the presence of colipase, showed more fluctuation in the absence and the presence of low concentrations of NaTC (PPL^{-colip} and $PPL^{-colip+TcL}$) compared to $PPL^{-colip+TcH}$. The root mean square fluctuations (RMSF) of individual residues during the simulation were calculated using C_α-atoms of the residues as reference (Fig. 2). It was observed that the *N*- and *C*-terminals of the protein gain more fluctuations in the presence of NaTC. The fluctuations around the lid region were found to be significantly reduced in the presence of colipase, particularly in high concentration of NaTC ($PPL^{-colip+TcH}$) (Fig. 2B). In the absence of colipase, lid fluctuation was observed to be similar in all three cases (PPL, PPL^{+TcL} and PPL^{+TcH}), with slight decrease in the presence of low concentration of NaTC (PPL^{+TcL}).

Colipase is 95 amino acid long polypeptide chain with five disulphide bonds. It is found to be stable against heat and surface-active agents [267]. The structure of colipase shows four protruding hydrophobic patches projecting in same direction which are known as fingers (Fig. 3A). In order to examine the bile salt binding sites on the colipase, the structure of colipase was extracted from lipase-colipase complex. Colipase alone was used for further MD simulations with low or high concentrations of NaTC. These results were compared with the colipase dynamics obtained from the lipase-colipase complex. In all the cases, the RMSF of colipase (Fig. 3B) was less than 4 Å except at the *N*-terminal where it went up to 6 Å. At the core, residues of the fingers 2, 3 and 4 showed notable

fluctuations. Among the trajectories, the lowest RMSF was observed for the colipase simulated in the presence of high concentration of NaTC (colipase+TcH).

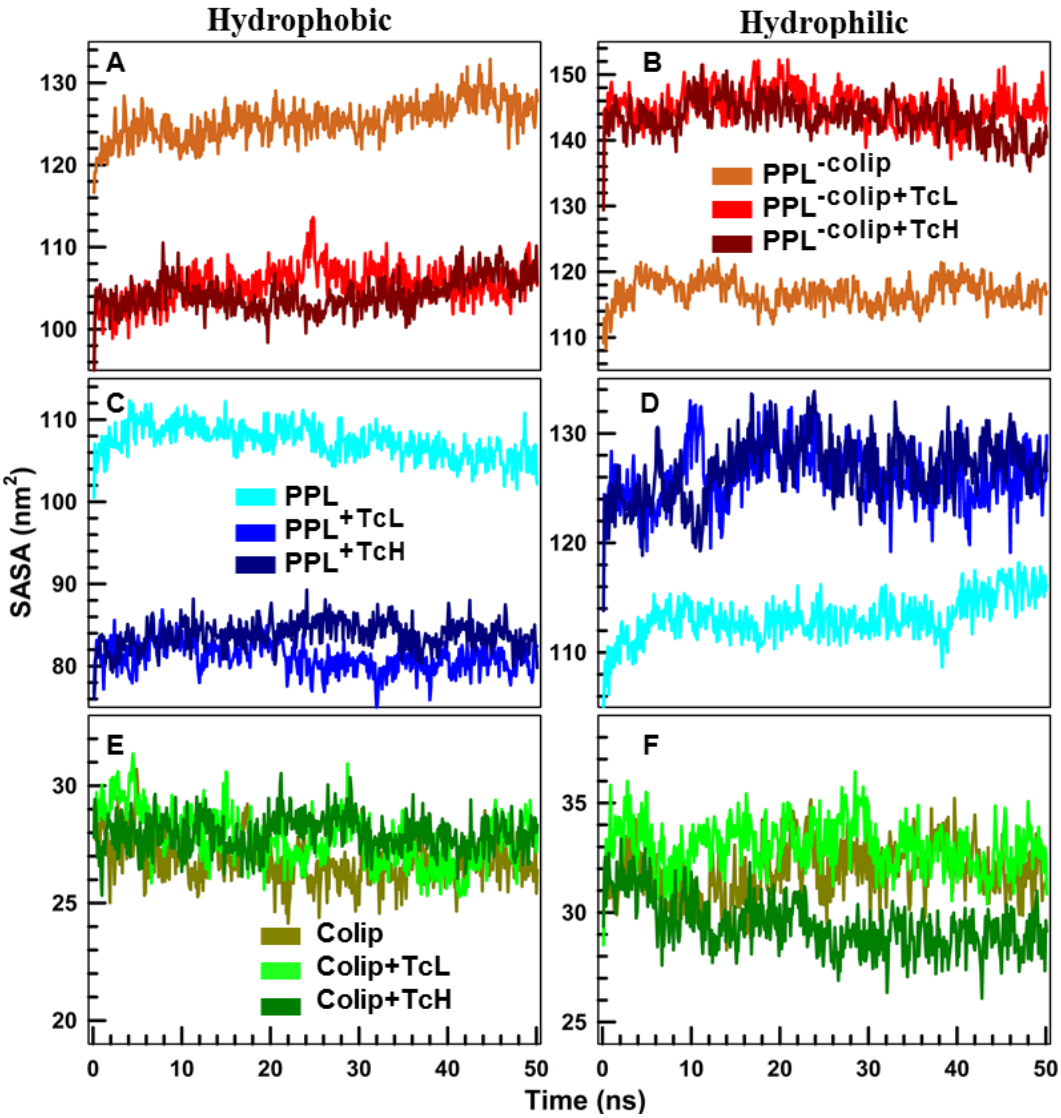


Figure 4. Change in hydrophobic SASA (left) and hydrophilic SASA (right) for the lipase in lipase-colipase complex (PPL^{-colip}) (A and B), lipase in the absence of colipase (PPL) (C and D), and only colipase in the absence of lipase (Colip) (E and F) in low (TcL) and high (TcH) concentrations of NaTC.

Solvent accessible surface area (SASA) of the complexes was divided into two different components: hydrophobic and hydrophilic. Changes in both the components were individually evaluated for lipase and colipase chains in the presence and the absence of NaTC (Fig. 4). Notably, hydrophobic SASA decreased with the addition of NaTC to the lipase-colipase complex and to lipase. In case of colipase, the hydrophobic SASA was

nearly unchanged in the presence and the absence of NaTC. However, hydrophilic SASA was increased by the addition of NaTC to lipase-colipase complex and to lipase which was slightly decreased for colipase in the high concentration of NaTC. Lipase in its open conformation is known to have large secondary structural changes around the lid region (residue 238-262) in the absence of colipase and NaTC [253, 268]**.** The secondary structural analysis of the lipase lid region (Fig. 5) proposed that in the presence of NaTC the helical structure of the lid was almost intact throughout the simulation even in the absence of colipase, though trivial changes in position and considerable fluctuations were observed (as mentioned above).

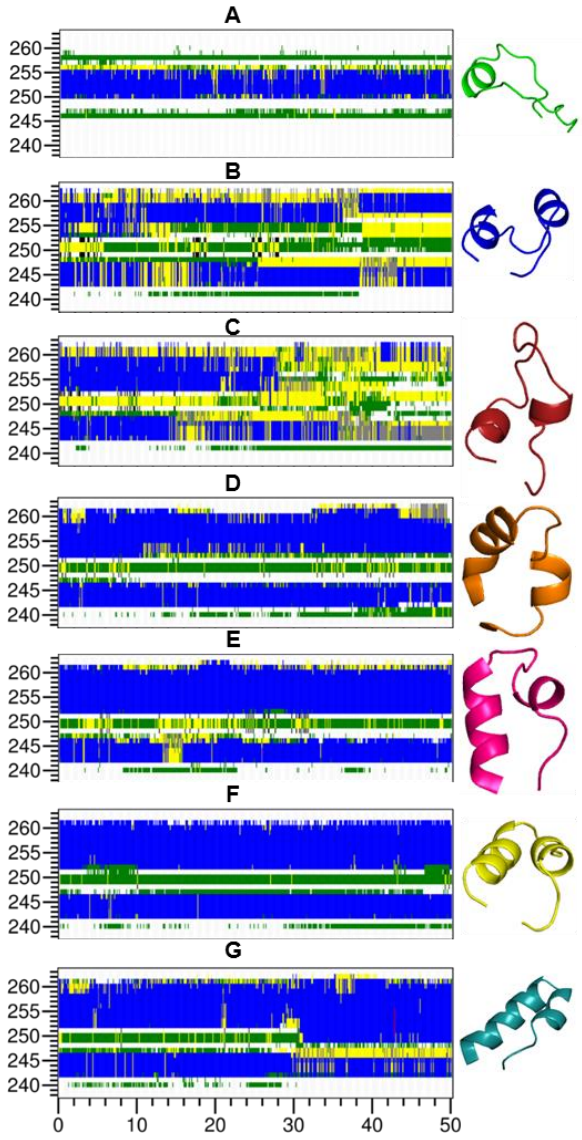


Figure 5. Secondary structural changes in the lid region (residues 238-262) of PPL-CL (A), PPL^{+colip} (B), PPL (C), PPL^{+TcH} (D), PPL^{+TcL} (E), PPL$^{+colip+TcH}$ (F), PPL$^{+colip+TcL}$ (G) during their respective MD simulation. Color codes for secondary structure description are: blue – α-helix, yellow – turn, green – bend, white – coil, grey – 3_{10}-helix, black – β-Bridge and purple – π-helix. Right panels present the cartoon diagram of the lid region at the end of 50 ns simulation.

4.4.2 Aggregation Property of NaTC

NaTC is a conjugated trihydroxy amphiphilic bile salt molecule [269]. NaTC forms micelles at the concentrations above 4 mM in water at room temperature and physiological salt concentration. In order to understand the micellation properties of NaTC and its effect on the dynamics of lipase and colipase chains, NaTC was simulated in water, and in the presence of lipase and colipase. 12 and 60 molecules of NaTC was considered to represent lower (TcL) and higher (TcH) concentrations, respectively. The simulations were initiated by randomly positioning the NaTC molecules and the molecules were allowed to move without any constraint. The MD simulation exhibited the formation of different types of aggregates. For further analysis, the aggregates were classified as oligomers (1-4 molecules), small aggregates (5-8 molecules) and micellar aggregates (9-12 molecules). All the atoms of NaTC were considered to quantitate the fraction of NaTC in their respective aggregates, using in-

house R-code, with a cutoff of 4 Å. Low concentration of NaTC (TcL) did not show any micellar aggregates whereas high concentrations of NaTC had all the three types of aggregates (Fig. 6) when simulated in water. It was also observed that no small aggregates or micellar aggregate were present in the PPL$^{+Colip+TcL}$ and Colip+TcL suggesting that these simulations are comparable to the lower concentrations of NaTC conditions used in experiments. Further, the data suggested that in the presence of protein chains more than half of the NaTC were found to be in oligomers except in the presence of colipase alone (colipase+TcH). In colipase+TcH, NaTC had almost equal fractions oligomers and small aggregate with notable fraction of micellar aggregates as well.

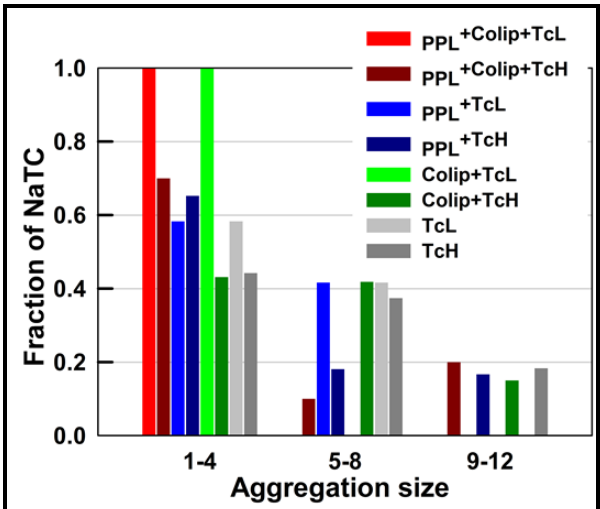


Figure 6. The mean fraction of different types of NaTC aggregations formed during the last 10 ns of MD simulations performed in the presence (PPL^{+colip}) and the absence (PPL) of colipase and in only colipase (Colip). The different aggregates formed by low (TcL) and high (TcH) concentrations of NaTC in water is included for reference (light and dark gray, respectively).

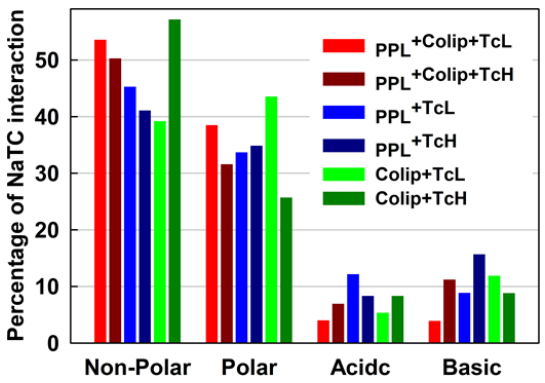


Figure 7. The percentage of interaction of NaTC with each groups of amino acids in the presence (PPL^{+colip}) and the absence of colipase (PPL), and only colipase (Colip) in low (TcL) or high (TcH) concentrations of NaTC.

4.4.3 NaTC Interaction Sites on Lipase and Colipase

The percentage of NaTC interaction with each amino acid of PPL and colipase was calculated. At both the NaTC concentrations, a larger number of interactions were observed to be with non-polar residues, followed by polar residues whereas the interactions with charged residues were less (Fig. 7). In the lipase-colipase complex Leu, Ile and Pro were noted as the most interacting residues with NaTC. In lipase chain Leu, Ile, Gly, Pro and Val, and in the colipase Leu, Ile and Gly were interacting more with the surfactant. Among the polar residues Asn, Gln, Ser and Thr were interacting more with lipase-colipase complex whereas Cys, Ser, Asn, and Thr were interacting in lipase in large number. In lipase and lipase-colipase complex at high concentration of NaTC, Lys was also considerably involved whereas at low concentration of NaTC, there were no considerable interactions with charged residues. On the other hand, lipase at low concentration of NaTC showed considerable fraction of interaction with Asp. In colipase

as well, polar residues showed more interactions with NaTC than the charged residues (Fig. 8).

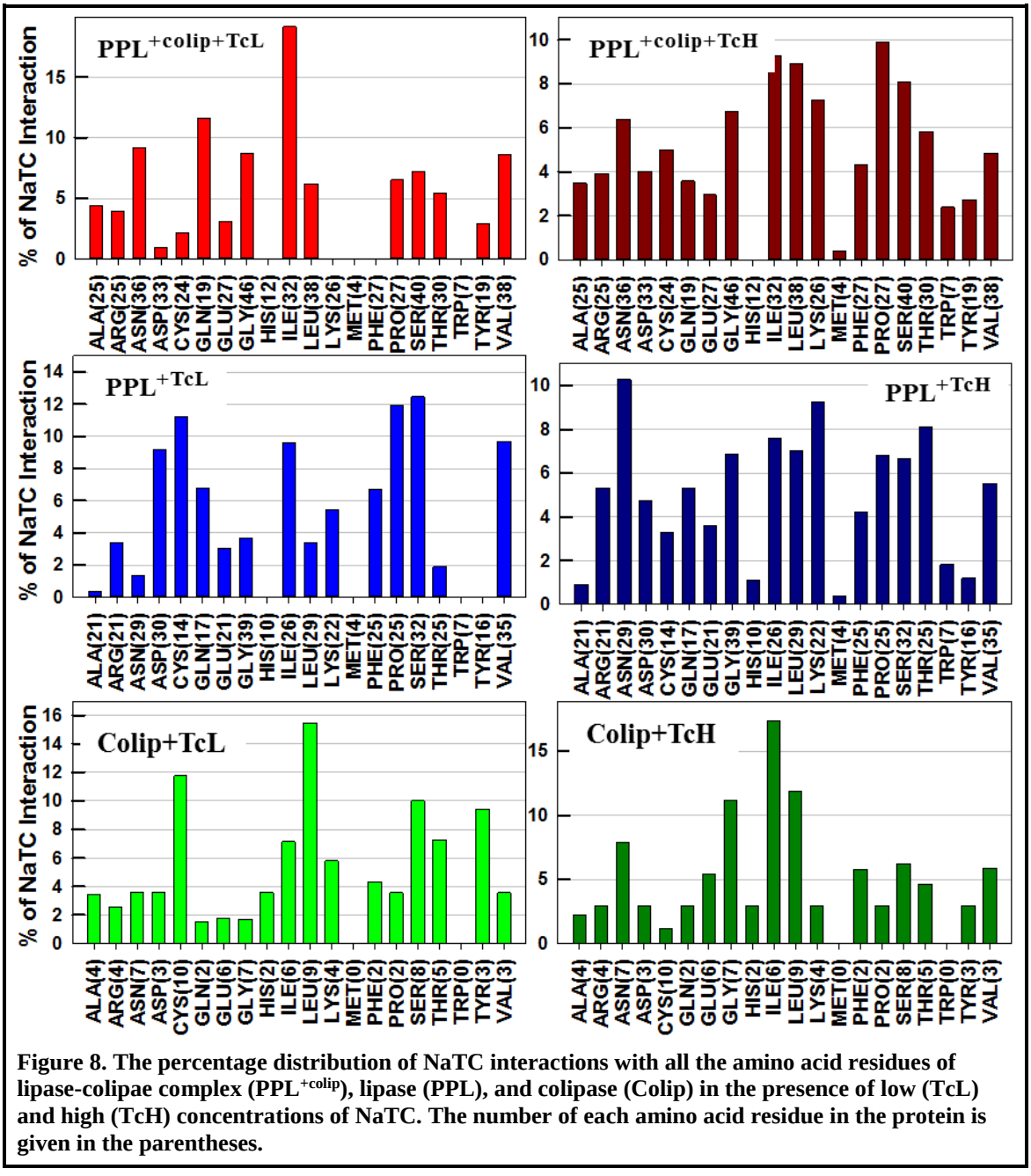


Figure 8. The percentage distribution of NaTC interactions with all the amino acid residues of lipase-colipae complex (PPL^{+colip}), lipase (PPL), and colipase (Colip) in the presence of low (TcL) and high (TcH) concentrations of NaTC. The number of each amino acid residue in the protein is given in the parentheses.

4.4.4 Dynamics of Colipase

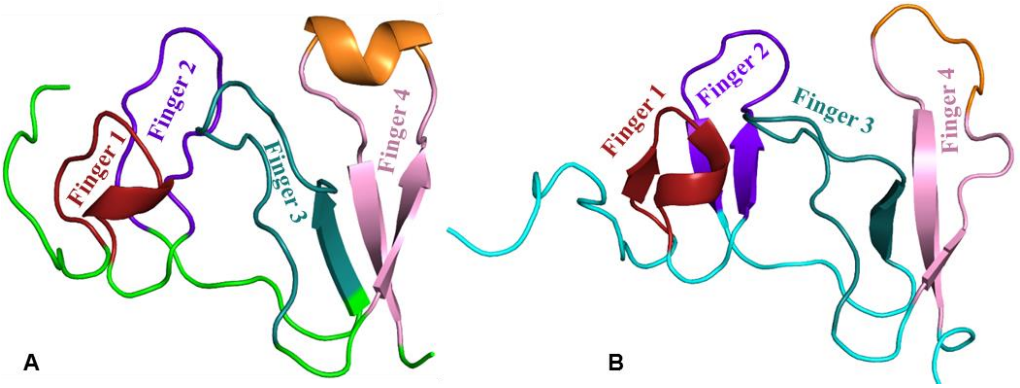


Figure 9. The cartoon diagram of colipase (A) obtained from the crystal structure of lipase-colipase complex (PDB id: 1ETH) and (B) obtained from the solution NMR structure (PDB id: 1PCN). The four protruding fingers are highlighted with different colors.

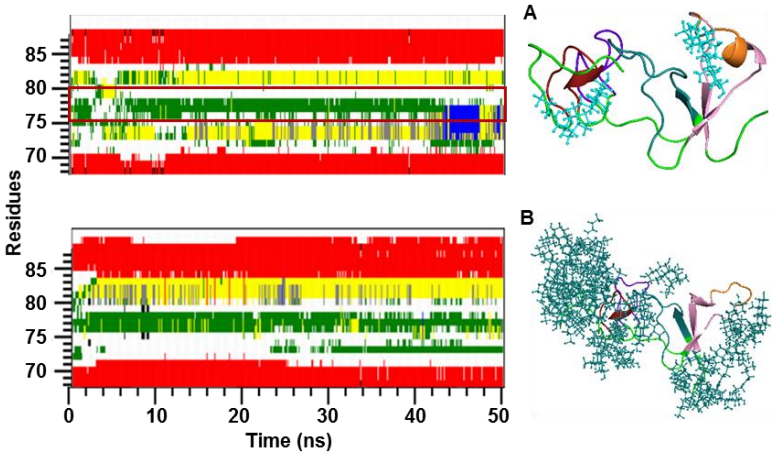


Figure 10. The change in secondary structure of the fourth finger of colipase during MD simulations of only colipase in (A) low concentration or (B) high concentration of NaTC. The initial structure of colipase is from solution NMR (1PCN). Color codes for secondary structure description are: blue – α-helix, yellow – turn, green – bend, white – coil, grey – 3_{10}-helix, black – β-Bridge and purple – π-helix.

The presence of lipase and NaTC were found to significantly influence the dynamics of colipase during MD simulation. For complete insight on the colipase dynamics, three different types of colipase were implemented in MD simulation in the presence and absence of NaTC; colipase in the presence of lipase (PDB: 1ETH), colipase in the absence of lipase (colipase extracted from 1ETH) and NMR structure of colipase (PDB: 1PCN). The fourth finger of colipase has a short helical stretch of 6 residues (residue 75-80) (Fig. 9A) in the crystal structure of the colipase, but it was not found in the NMR structure (Fig. 9B). However, the helix structure was induced by low concentration of NaTC during the simulation of NMR structure of colipase (Fig. 10). It was also observed from the other simulations (using crystal structure of colipase) that the helix remained stable in the presence of NaTC (Fig. 11-12) except in PPL^{+Colip} (Fig. 11A).

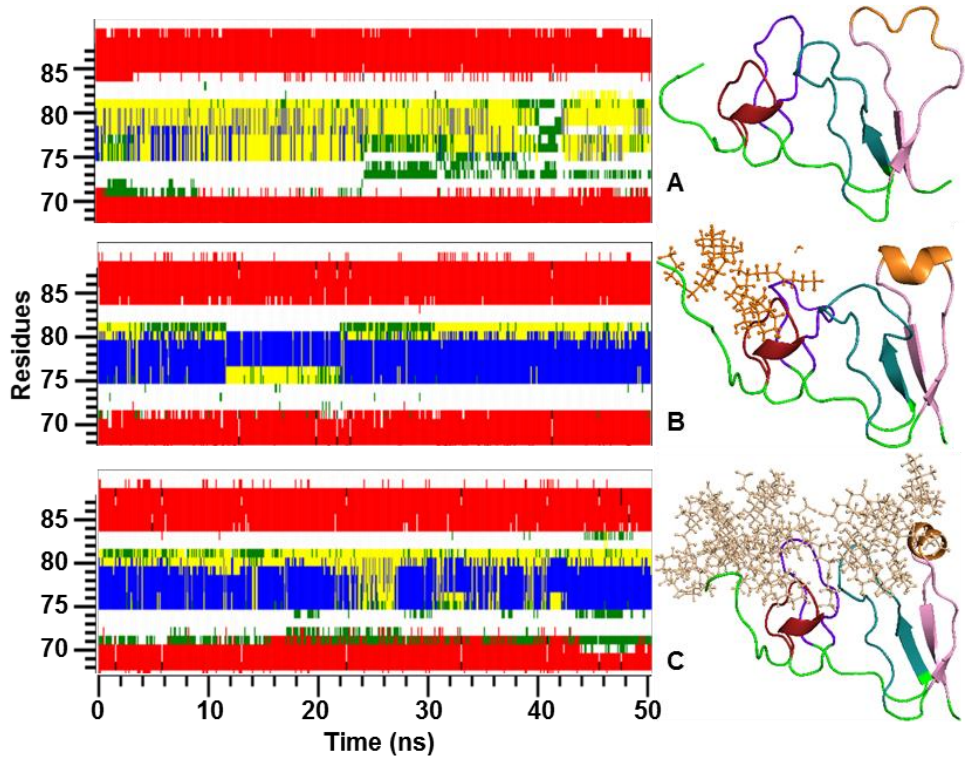


Figure 11. The change in secondary structure of the fourth finger of colipase during MD simulations of lipase-colipase complex (PPL^{+Colip}) in the absence (A), in low concentration (B) and in high concentration of NaTC. Refer Fig. 10 for color code.

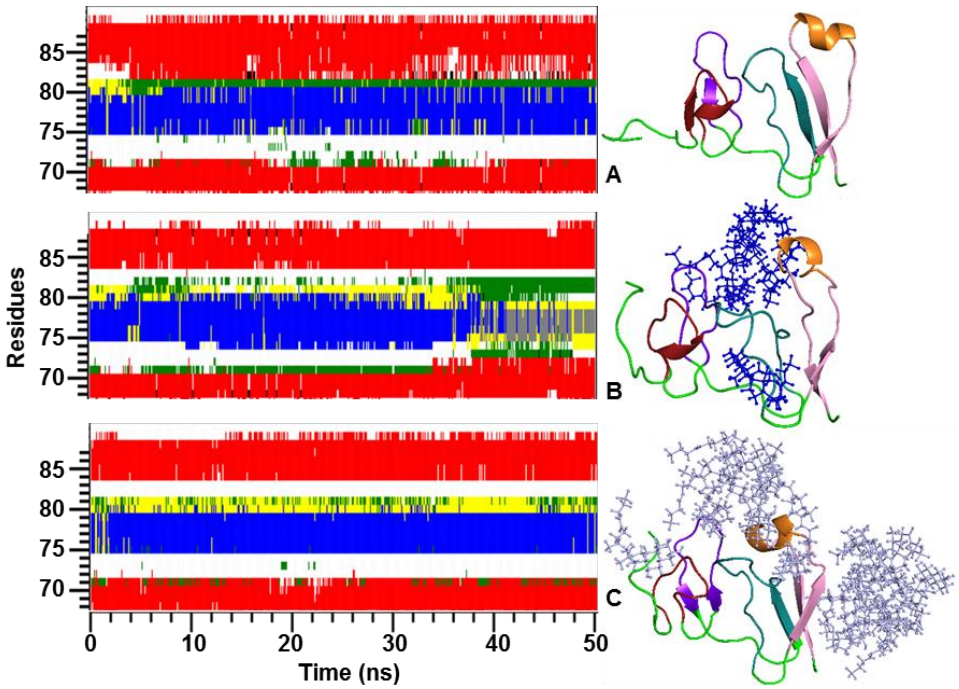


Figure 12. The change in secondary structure of the fourth finger of colipase during MD simulations of only colipase (without lipase) in the absence (A), in low concentration (B) and in high concentration of NaTC. The initial structure of colipase is obtained from the crystal structure of lipase-colipase complex (PDB id: 1ETH). Refer Fig. 10 for color code.

4.5 Discussion

4.5.1 Effect of NaTC on the Structures

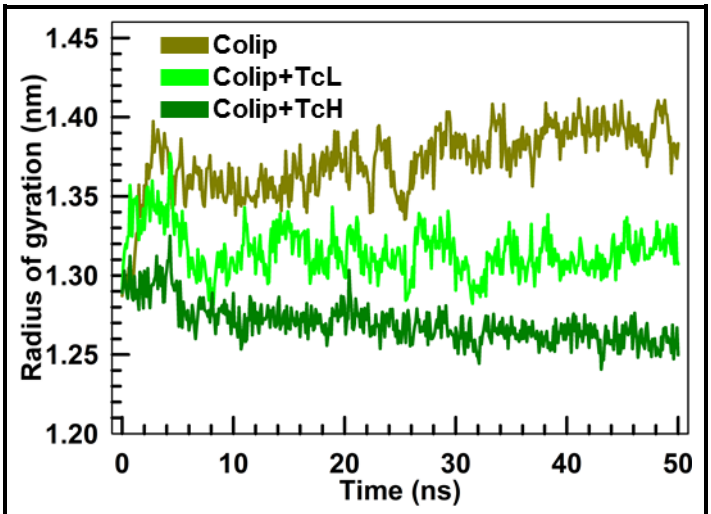


Figure 13. The radius of gyration (R_g) of Colipase when simulated in the absence of lipase. R_g of colipase is plotted as a function of time in the absence and presence of NaTC.

Significant changes in the global dynamics of the protein could be observed in the presence of both lower and higher concentrations of NaTC which is directly evident from the RMSD changes (Fig. 1). In the presence of colipase, RMSD is minimum in the absence of NaTC and maximum RMSD is shown by $PPL^{+Colip+TcL}$. RMSD of $PPL^{+Colip+TcH}$ has intermediate value, but close to PPL^{+Colip}. Similary, the RMSD of PPL^{+TcH} is maximum followed by PPL^{+TcL} and minimum in PPL in the absence of colipase. The mean fluctuation in the structure of lipase in the presence of NaTC correlates well with the increase in the hydrophilic SASA and decrease in hydrophobic SASA (Fig. 4) suggesting that the added NaTC would preferably occupy the hydrophobic surfaces on the lipase both in the presence and absence of colipase. However, when NaTC is added to colipase, only the higher concentrations of NaTC could reduce its hydrophilic SASA whereas hydrophobic SASA almost remains same. This could be attributed to the increased compactness of the colipase at supra-micellar concentration of NaTC which is evident from the radius of gyration (Fig. 13). This reduction of SASA is mostly contributed by hydrophilic residues. Particularly, the formation of polar interactions between bile salt micelle and the fourth finger of colipase (residues Glu70 and Asp72) at the rear side of the colipase might provide strong anchorage for the micelle.

Multiple interaction sites for NaTC could be found on the surfaces of both the lipase and the colipase chains which facilitate the binding of NaTC in monomeric and in oligomeric state as well (Fig. 9). The global changes observed in the protein which led to the increase in RMSD could be attributed to the multiple sites available on the protein which affected the fluctuations of hydrophobic and hydrophilic residues. Among all, $PPL^{+Colip+TcH}$ in which the colipase is held with lipase in the presence of TcH shows relatively less fluctuations. Also, the number of NaTC molecules bind at the junction of the lipase lid and colipase is relatively more which might provide additional stability to the complex, thus minimizing the RMSD. It could be further affirmed by the reduction of RMSD around the lid region (Fig. 1B) of $PPL^{+Colip+TcH}$. Interestingly, RMSD of the lipase

and their respective lid do not follow the same trend. In the absence of the NaTC, the lid RMSD is maximum for PPL in the absence of colipase. This suggests that NaTC induces global rearrangements and exposes the hydrophilic groups on the surface which can further assist the NaTC binding. On the contrary, the lid residues are more stabilized in the presence of NaTC and colipase. It was also apparent from the snapshot of the last frame (Fig. 14F) where more numbers of NaTC molecules are bound to the hydrophobic tips of colipase fingers along with the lid. In the absence of colipase, lid RMSF are almost same (Fig. 2). Probably, the direct interaction of NaTC molecules with the lid could prevent its denaturation in all the cases. Further, the *N*-terminal domain and *C*-terminal domain of the lipase show more RMSF (Fig. 2) in the presence of the NaTC. This could be due to the increased flexibility of the residues around the binding sites of NaTC. In the colipase chain, the RMSF is reduced at high concentration of NaTC (colip+TcH), because of formation of micellar and oligomeric aggregates of NaTC around the colipase (Fig. 3 and Fig. 12C). Two most preferable binding sites are identified on the colipase for NaTC: the micellar aggregate binding site at the rear side of the fourth finger; and oligomeric binding site covering the hydrophobic tips of the fingers.

4.5.2 Conformational Changes on the Lid

Apart from the terminal regions, NaTC has considerable effect on the lid dynamics of the lipase. In all the cases, direct interactions between the lid and NaTC molecules are observed. In the presence of NaTC, secondary structure of the lid is intact (Fig. 5D-G). Though the values of RMSD and RMSF suggest that there have been positional fluctuations and displacement in the lid residues, the topology of the lid is largely maintained. In the absence of the NaTC, distortions in the lid secondary structure depend on the initial conformation of the protein. The structural changes on the lid in closed conformation and in the presence of colipase are little whereas in the open conformation lid shows larger structural and positional distortions (Fig. 5A-C). In fact, the lid moves towards the closed conformation, if the MD simulation is initiated from the open-state without any constrains [253].

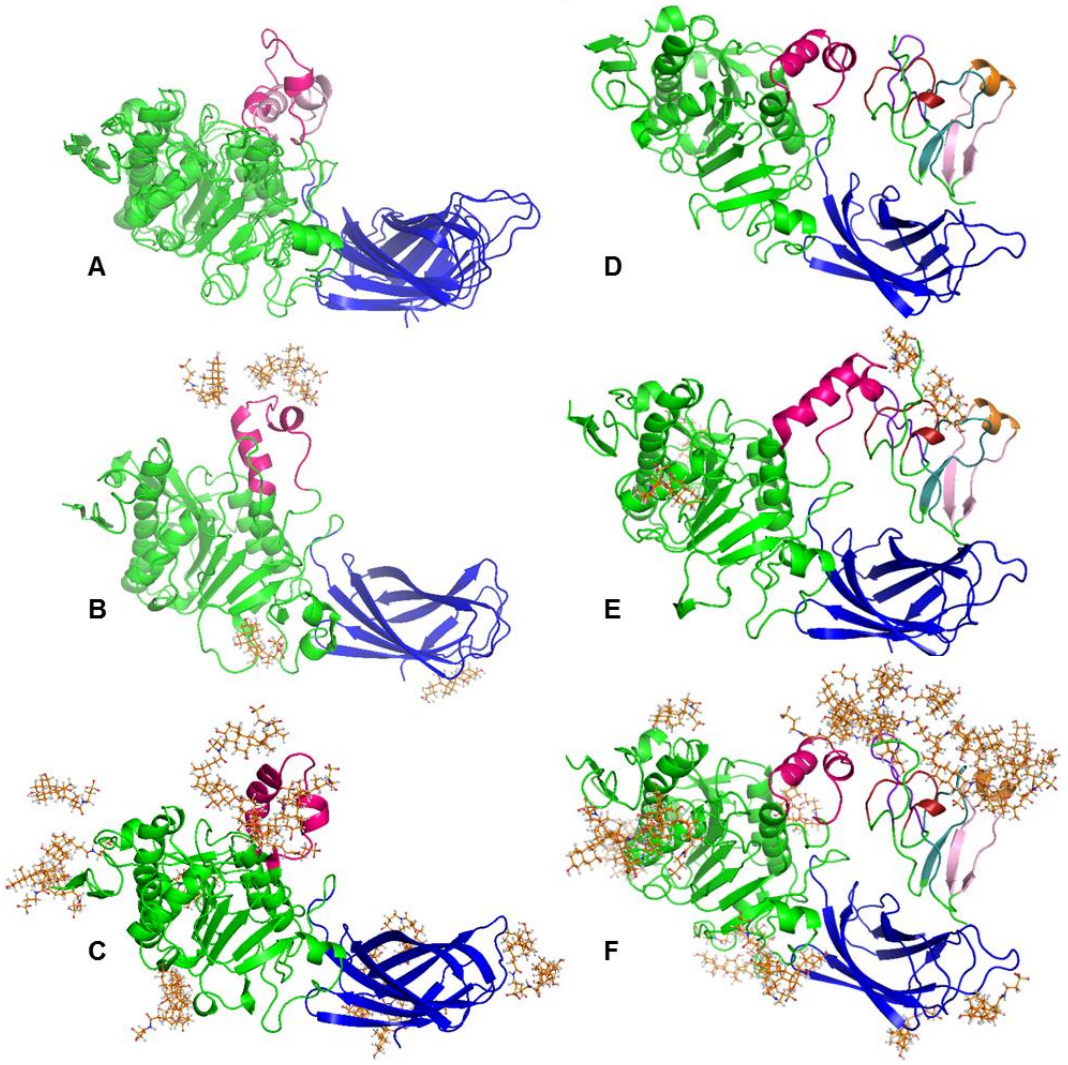


Figure 14. Snapshots of the last frames from the trajectory of 50 ns MD simulations of lipase (A-C) and lipase-colipase complex (D-F) in the absence (A and D), in low concentration (B and E) and in high concentration (C and F) of NaTC. N-terminal, C-terminal and the lid of lipase chain is represented by green, blue and pink colors, respectively. NaTC molecules are presented as stick models.

4.5.3 Self-aggregation of NaTC and its Residue-Specific Interactions

NaTC consists of three hydroxyl groups, one amide and one sulfate group which are polar and rest of the molecular skeleton is hydrophobic in nature. Taurodeoxycholate (derivative of dihydroxy conjugated cholic acid, NaTDC) is known to form larger aggregates than NaTC in a wide range of ionic strength [269, 270]. In the present work, MD simulations of NaTC show that the molecule can form oligomeric aggregates and smaller aggregate at low and high concentrations. Large micellar-like aggregates, with 9-12 monomers are formed only at higher concentration of NaTC [269, 270]. In other words, 12 molecules of NaTC in 1000 nm^3 box behave as a sub-micellar concentration and 60 molecules of NaTC represent supra-micellar concentration. For time and cost effective computing NaTC is chosen against NaTDC. Also, the matured micellar

aggregates of NaTC could be created by MD simulation without applying any constrains [269, 270]. Further, NaTDC inhibits the activity of lipase at the concentrations 1 to 3 mM whereas NaTC enhances the activity in rat and human pancreatic lipase. Only at above 3 mM, NaTC inhibits the lipase activity and it is retained even above the CMC of NaTC in the absence of colipase [129, 254].

NaTC molecules interact with all classes of amino acid residues. However, more preferential interactions could be found with hydrophobic residues followed by polar residues. A small fraction of interaction with the acidic and basic residues is also noted during the simulation (Fig. 7). The presence of three hydroxyl groups act as hydrogen bond donors toward acidic and polar residues, the amide and sulfate groups interacted with basic and polar residues, and the hydrophobic carbon skeleton shows affinity with the non-polar residues.

4.5.4 Bile Salt Interaction Models

Brockerhoff, H. (1973) [271] proposed that for pancreatic lipases the nature of both substrate and product are hydrophobic and they must be reluctant to leave the oil phase of the substrate emulsion. So near the active site of the lipase, the oil-water interface should be in direct contact and the substrate molecules should remain partially in the oil phase for the catalysis. Though the role of surfactants in the stability of the lipase in its active conformation was not discussed in earlier studies, it was shown that bile salt resides on the surface of the substrate emulsion. Chapus *et al.* (1975) [272] proposed a scheme in which the activity of lipase depends on the formation of a complex comprising the lipase, colipase, bile salt and the substrate emulsion. According to this model, bile salt forms a mono layer on the emulsion and free micelles. These forms of bile salts have direct interacting capacity with the colipase which further creates or unmasks the lipase binding site. This scheme, however, did not include the substrate entry or exit sites. Borgström (1977) [273] proposed that the lipase and colipase complex is formed in the bulk phase which has micelle binding site. This ternary complex finally binds to the substrate emulsion.

The knowledge of the pancreatic lipase and its interfacial activation greatly enhanced after Winkler *et al.* (1990) [274] and van Tilbeurgh *et al.* (1993) [161] deciphered the structure of lipase and lipase-colipase complex with micelle, respectively. Hermoso *et al.* (1997) [263] with the help of small-angle neutron scattering (SANS) experiment modeled the orientation of lipase, colipase and bile salt micelle. According to this model, pancreatic lipase activation is not interfacial, but it is mediated by the bile salt micelle and colipase. The micelle could be placed at the concave grove formed between the colipase and *C*-terminal of the lipase. The center of mass separation of colipase and micelle was proposed to be ~27 Å, with 22-24 molecules of the NaTDC. The micelle was estimated to be prolate with 10 Å radius of minor axis and nearly 20 Å for major axis.

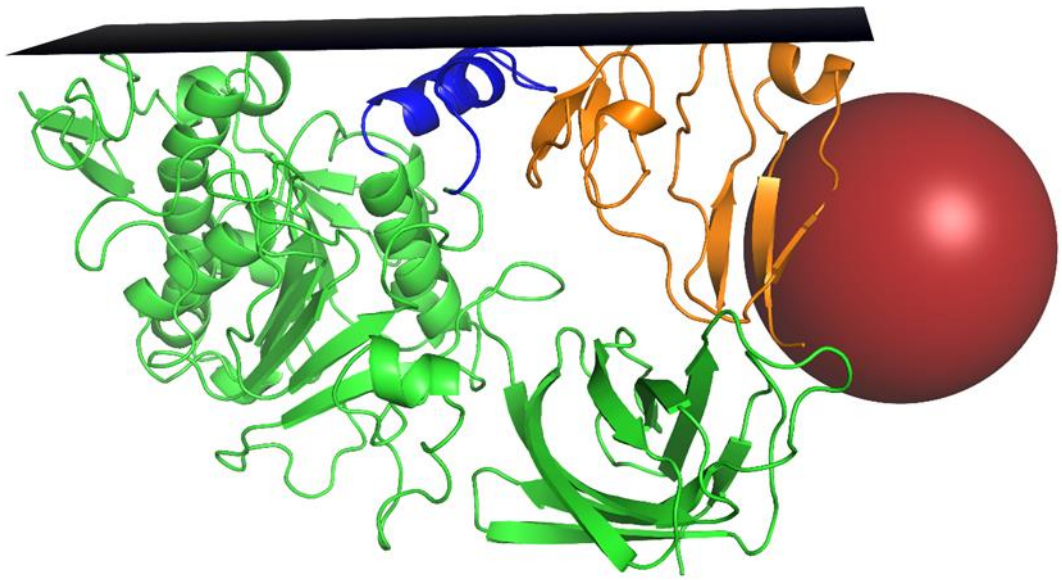


Figure 15. Schematic representation of lipase (green) - colipase (orange) complex with the micelle of NaTC. The red sphere represents the total volume of the micelle consisting of nine molecules of NaTC and the black plane represents the hydrophobic plateau formed by protruding fingers of colipase, lid (blue) and the *N*-terminal residues of the lipase.

In the present study, NaTC at high concentrations form micelle-like aggregates when simulated without any constrain in the presence and the absence of the lipase. The micelle formed in the presence of lipase-colipase complex is not found on the surface of the lipase. However, monomeric NaTC molecules bind on the lipase surface and also small aggregates of NaTC could be identified on the tip of lid and lipase-colipase interface. It is also observed that the micellar form of NaTC (Fig. 15 and video file V4) is created by 9 monomeric molecules with the axial radii of 11 and 15 Å (minor and major axis, respectively). The distance between the center of mass of the colipase and micelle is 23.59 Å as obtained from Colip+TcH simulation. The exact position of micelle on PPL^{+Colip} was evaluated by superimposition of the colipase-micelle complex on PPL^{+Colip}. The probable position of the bile salt micelle is found to be the concave grove formed between the colipase and *C*-terminal of the lipase which is similar to the model proposed by Hermoso *et al.* The positioning of the micelle suggests that the micelle has direct interaction with the fourth finger of colipase. The MD simulation also suggests that hydrophobic tip residues on colipase along with the lid and *N*-terminal residues (Gly109, Gly110, Gly114, Val35, Gly19, Ile20, Val21, Pro24, Leu25, Ile27 and Leu28) create a hydrophobic plateau (black plane in Fig. 15 and video file V4) which could interact with the substrate emulsion. Also, the large micelle (larger than 15 Å) might also interact with the substrate emulsion and provide a special anchorage of the complex.

4.6 Conclusion

The conjugated trihydroxy bile salt, NaTC is a convincing model of bile salt to study the role of bile salt in the lipase-colipase stabilization at the physiological conditions. Twelve and sixty molecules of NaTC in 1000 nm^3 of the simulation box satisfactorily represented sub-micellar and supra-micellar concentration of the NaTC molecules, respectively. From MD simulation, it is found that NaTC can be in monomeric as well as in higher aggregate forms with aggregation number varying from 4-8 and 9-12, in the presence and absence of the protein. A cooperative exposure of the polar residues in the protein was evident from increase in hydrophilic SASA in the presence of NaTC. These rearrangements of polar residues might be helpful in creating the binding site for pancreatic lipase on the bile salt coated emulsion. At supra-molecular NaTC concentration, colipase shows decrease in the exposure of polar residues which probably assists the micelle formation at the rear side of fourth finger and facilitates small aggregate binding at the hydrophobic tips of the other fingers. It is known that at both the concentrations NaTC molecules are efficient in inhibiting the denaturation of the lid. This could be attributed to the direct interaction of NaTC to the lid residues. This also explains the role of NaTC, at very low concentration, which supports the lipase activity in the absence of the colipase by preventing the enzyme denaturation at the oil-water interface. In the 50 ns simulation, lipase-colipase complex was considerably stabilized by supra-micellar concentration of the NaTC keeping the lid more intact in the open conformation. The presence of the micelle at the rear side of fourth finger assists the lipase-colipase stability and the *C*-terminal of lipase has minimal role in the formation of colipase-micelle complex.

Chapter 5

Insights into Protein-TNS (2-p-toluidinylnaphthalene-6-Sulfonate) Interaction using Molecular Dynamics Simulation

5.1 Abstract

Fluorescent dyes are used to investigate the hydrophobic exposure of proteins in different conformational states. Despite of their wider applications, the molecular level interactions of these dyes with different states of proteins are not well understood. In this part of the thesis, we have analyzed the interaction between protein and 2-*p*-toluidinylnaphthalene-6-sulfonate (TNS) using molecular docking and simulation. Three different unfolded conformations of alpha-lactalbumin (α-LA) and ribonuclease A (RNase A) were generated. TNS was docked against these protein states and the molecular simulation of protein-TNS complexes were carried out. These results were compared with TNS in water and TNS bound to the native form of bovine serum albumin (BSA). The results suggest that TNS binds to the proteins through ionic and hydrophobic interactions. TNS bound to the unfolded conformations of α-LA and RNase A and the native-BSA has less solvent accessibility and surrounded by more hydrophobic residues. The orientation of TNS in protein bound states were analyzed using two of its bond angles and dihedrals. The analysis shows that the aromatic rings of TNS move from near planar to non-planar in different protein-TNS complexes and ring rotation is much constrained. Orientation of TNS is influenced by the conformation of the binding pocket and the hydrogen bonding interactions between the protein and TNS. Therefore, we suggest that TNS fluorescence enhancement upon binding to unfolded states of proteins might be due to the hydrophobic environment and reduced solvent accessibility rather than the binding ability and specific orientation of TNS in the bound state.

5.2 Introduction

There has been a continuous research in the field of protein structure in relation to its function. Understanding protein structure and stability would answer various questions regarding the diseases caused by protein malfunctioning. Various states of proteins, from natively folded to random coil have been studied using various biophysical techniques [275-286]. Fluorescence as a spectroscopic tool is tremendously used in the study of biomolecules [287-292]. Most of the proteins have aromatic amino acid residues, tryptophan, phenylalanine, and tyrosine, which provide intrinsic fluorescence to the proteins. This has larger applications in protein structural studies. Fluorescent molecules (such as dansyl chloride and fluorescein) chemically conjugated to amino acid side chains [293-296], and externally added fluorophoric dyes are also used to follow conformational changes in proteins. These dyes are sensitive to their micro-environment such as polarity, ionic strength, and pH [277, 281-283]. The most commonly used fluorophores are pyrene, 8-anilino-1-naphthalene-sulfonate (ANS), 6-(*p*-toluidino)-2-naphthalenesulfonate (TNS), and, 6-propionyl-2-dimethylaminonaphthalene (prodan) [277-285] . The quantum yield of these dyes in a non-polar environment is, in general, quite higher than in a polar environment which is used to probe the conformational changes of proteins.

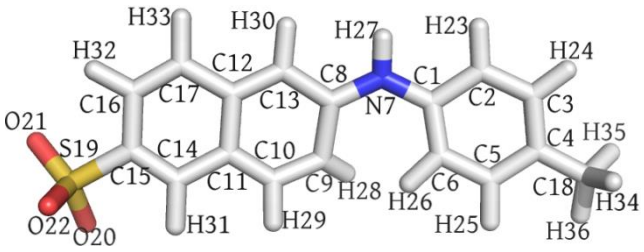


Figure 1. Line representation of TNS with individual atoms labels.

Among the fluorescent dyes, TNS (Fig. 1) is widely used as fluorescence probe to find hydrophobic pockets in proteins [277, 283-286, 297-299]. Fluorescence intensity change of TNS was used to follow the conversion of pepsinogen to pepsin by Wang *et al.* [284]. The transitions between non-charged and charged conditions of axons were characterized using fluorescence changes of TNS-stained axon [297]. In these studies, binding of TNS with proteins was mainly attributed to the hydrophobic interactions [277, 278, 282-286]. McClure and Edelman proposed that variation in the solvation of excited state is the origin for TNS fluorescence changes in different environments [277]. Studies with short peptides and micelles suggested that polar proximity and proper positioning of the aromatic ring are the necessary conditions to increase the quantum yield of TNS [284, 300, 301]. These studies also emphasized on the hydrophobic forces for the stabilization of protein-TNS interaction. However, the exact binding modes and molecular orientations are yet to be investigated to provide more insight on the TNS interaction with polypeptides.

Computer simulations have proved their importance in elucidating the interactions at the molecular and atomic level. They have spanned almost the whole spectrum of biomolecular research, from small molecules [302-304] to large molecules like ribosomes [305], from the processes of aggregation [304], micellization [306] to membrane [307] and membrane vesicles formation [308]. In the present study, we have carried out a series of molecular dynamics simulation experiments to understand the nature of TNS interaction with different conformational states of the proteins, ribonuclease A (RNase A), alpha-lactalbumin (α-LA) and bovine serum albumin (BSA). RNase A and α-LA were unfolded *in silico* and three partially unfolded conformations of the proteins were chosen to analyze TNS binding properties. Since TNS quantum yield increases when it binds to the native form of BSA [3], BSA was considered as a reference to compare the binding orientation and interactions. Molecular dynamic (MD) simulation of the TNS bound proteins were carried out to study the conformational changes of TNS upon binding to different states of proteins and in water. The results were analyzed in terms of ligand orientation, binding interactions, solvent accessibility, and the surrounding protein residues.

5.3 Materials and Methods

5.3.1 Fluorescence Measurements

TNS was purchased from Sigma-Aldrich. All the solvents in HPLC grade were from SRL, India. TNS stock was made in 10% ethanol and the concentrations were measured using absorption coefficient of 18.9 mM^{-1} cm^{-1} for TNS at 317 nm [277]. Fluorescence spectra of 10 μM of TNS in varying solvents, solvent-water mixtures and in the presence of 10 μM of the proteins (BSA, α-LA and RNase A) were measured between 350 nm and 550 nm after exciting the samples at 317 nm. During the measurements, temperature was maintained at 298 K using Peltier attached with the instrument. For the experiments performed in aqueous solutions (with and without protein), pH was maintained at 7 using 20 mM phosphate buffer. All the measurements were carried out in Fluoromax 3 (Horiba) spectrofluorometer.

5.3.2 Molecular Dynamics Simulation

Molecular dynamics simulations were carried out using GROMACS 4.5.4 [52, 149-151]. The force field parameters used were GROMOS 53a6. Protein structures, 1FS3 (RNase A), 1HFZ (α-LA), and 4F5S (BSA) were obtained from the protein data bank (www.rcsb.org). TNS structure was modeled using Automated Topology Builder and Repository ver1.2 [309] using GROMOS53a6 force field and quantum mechanical calculation at the level of HF/STO-3G. All the protein models, with and without bound TNS, were solvated in a dodecahedron simulation box with SPC-E (single point charge extended) water model with a periodic boundary condition. Lennard-Jones and the short-range Columbic interactions were calculated with a cutoff of 1.0 nm. Particle mesh Ewald (PME) [152] was used for treating long range electrostatic interactions. To obtain unfolded conformations of proteins, the constant volume condition was maintained by NVT at a higher temperature (700 K) [310]. For normal protein dynamics in ambient conditions NVT ensembles were maintained at 300 K. Throughout the simulation, the temperature was maintained by stochastic velocity rescaling scheme [311] with time constant of 0.1 ps. Pressure was maintained at 1 bar with time constant of 2.0 ps by Parrinello-Rahman [153] type pressure coupling. LINCS algorithm was used for all the bond constraints [154].

5.3.3 Docking

Docking of TNS against the proteins was done using AutoDock Vina [312] and GOLD [313]. Initial blind docking was done using AutoDock Vina. The search grid was put on the whole protein and the exhaustiveness was increased to 15-20 to compensate for the large search space. GOLD 3.2 was used for further refinement of the docked poses. The best-scored pose was used for further simulation. PyMOL 1.3 [155] was used for visual inspection and pictorial representation of data. DSSP was used to evaluate the

secondary structure contents. LigPlot+ [314] was used to find the pattern of H-bond and hydrophobic interaction of the amino acid residues with TNS.

5.4 Results

5.4.1 Fluorescence Spectra of TNS

Fluorescence emission of TNS was measured in different solvents with varying polarity (Fig. 2A). The quantum yield of TNS was almost null in water, as reported earlier [315]. TNS showed significant increase in fluorescence intensity in non-aqueous solvents. Among all the solvents, TNS had the highest quantum yields in DMF and DMSO whereas TNS fluorescence intensities were reduced in alcoholic solvents. TNS showed poor quantum yield in hexane, toluene, chloroform and acetone. The wavelength of fluorescence emission maxima (WL_{max}) of TNS was plotted against increasing polarity of the solvents according to the $E_T(30)$ polarity scale of Dimroth-Reichardt [316] (Fig. 2B). A monotonous bathochromic shift was observed as the solvent polarity was increased. However, DMF and DMSO showed a larger wavelength shift than expected suggesting that apart from the polarity, chemical nature of the solvents might also influence the WL_{max} as witnessed in other dyes as well [317-319]. In the present case, the solvents with strong H-bond acceptors shifted the emission wavelength towards the longer region. Fluorescence spectra of TNS in different binary mixtures of water, methanol, DMSO and DMF were measured (Fig. 3). As the polarity of the solvent mixture was decreased by the addition of methanol, DMSO or DMF, the fluorescence emission intensity was increased and the WL_{max} showed hypsochromic shift. Further, the fluorescence emission spectra of TNS were followed in the presence of three different model proteins: BSA, α-LA and RNase A (Fig. 4). Among the three proteins, BSA provided the highest quantum yield followed by α-LA. In the presence of RNase A, TNS showed poor quantum yield similar to that of polar solvents.

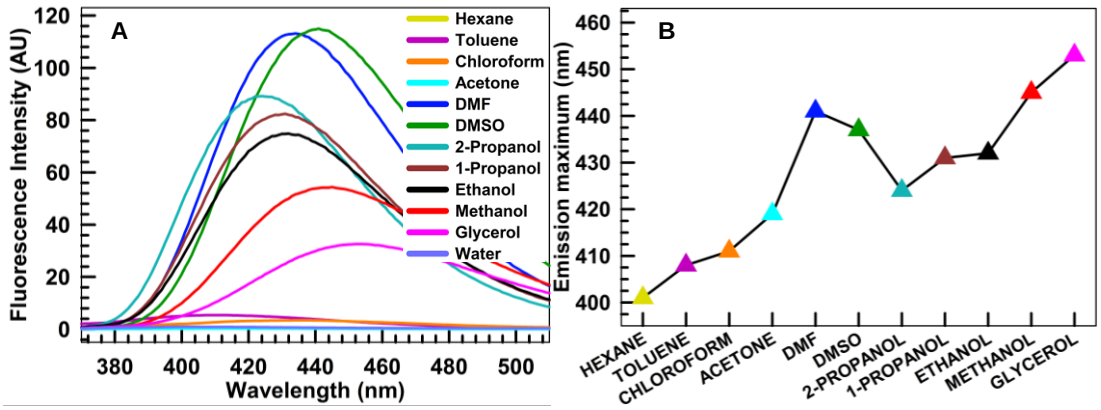


Figure 2. (A) Fluorescence spectra of TNS in different solvents after exciting at 317 nm. (B) Wavelength of emission maximum obtained from fluorescence spectra plotted against the solvents with the increasing order of polarity.

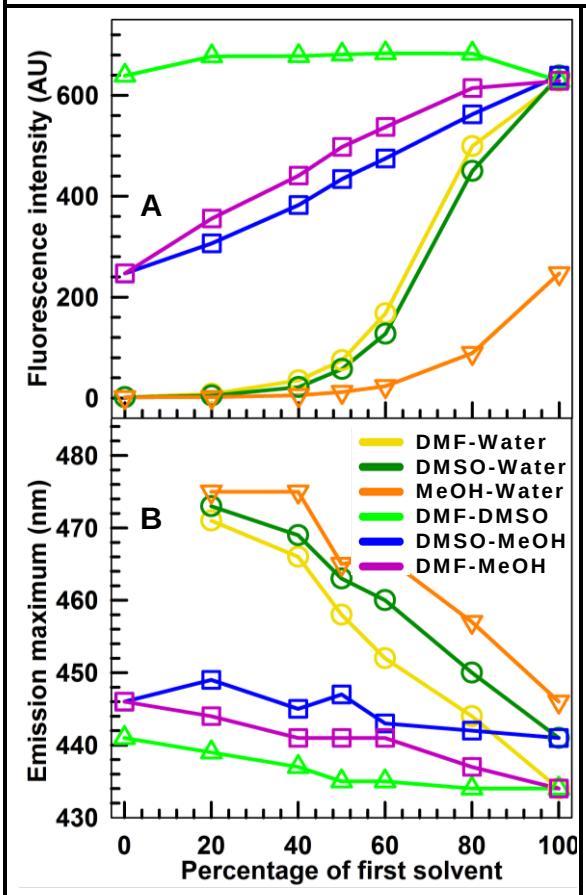


Figure 3. (A) Fluorescence intensity and (B) wavelength of emission maximum of TNS in binary solvent mixtures of water, methanol, DMSO and DMF. The abscissa represents the percentage of the first solvent in the mixure mentioned in the label.

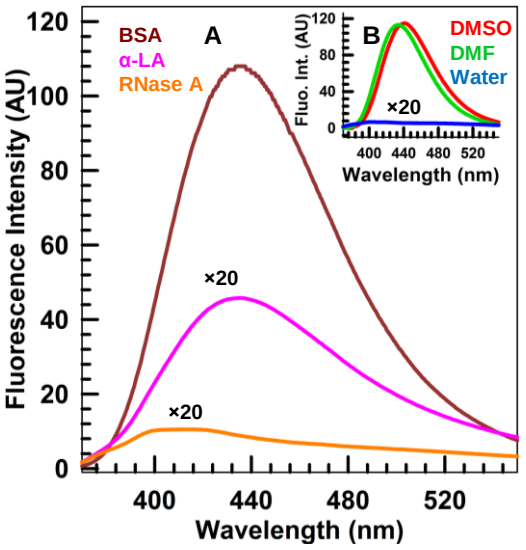


Figure 4. Fluorescence spectra of TNS in the presence of native-states of the proteins: BSA (brown), α-LA (pink) and RNase A (orange) after exciting at 317 nm. The spectra of TNS in the solvents (inset): water (blue), DMF (green) and DMSO (red) are shown for comparison (refer Fig. 2A).

5.4.2 Unfolded States of Proteins

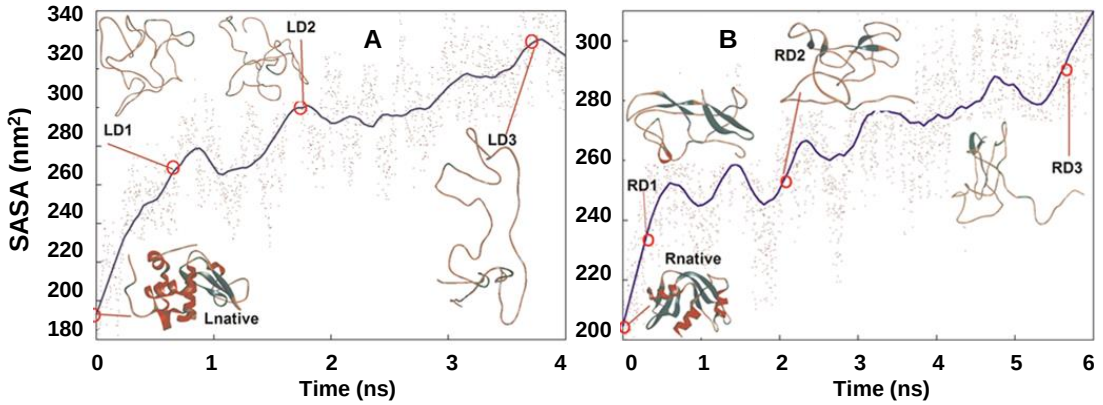


Figure 5. The total solvent accessible surface areas of α-LA (A) and RNase A (B) obtained from high temperature simulations (700 K). The solid lines represents the running average over 100 ps. The ribbon diagrams of selected protein conformations are shown pictorially at the respective time points. Triplicates of high temperature simulations were done for 4 ns (α-LA) and 6 ns (RNase A). α-LA and RNase A attained the maximum SASA and RMSD in 4 and 6 ns, respectively.

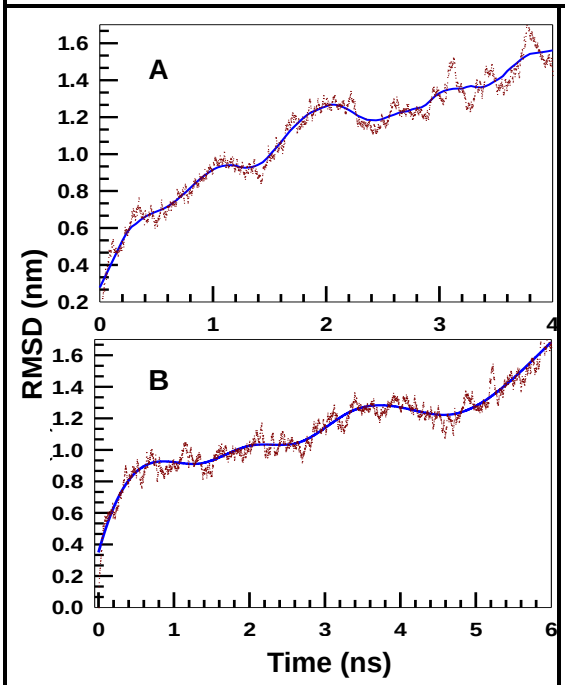


Figure 6. RMSD (A) α-LA and (B) RNase A simulation carried out at 700K. The solid line represents the running average over 100 ps.

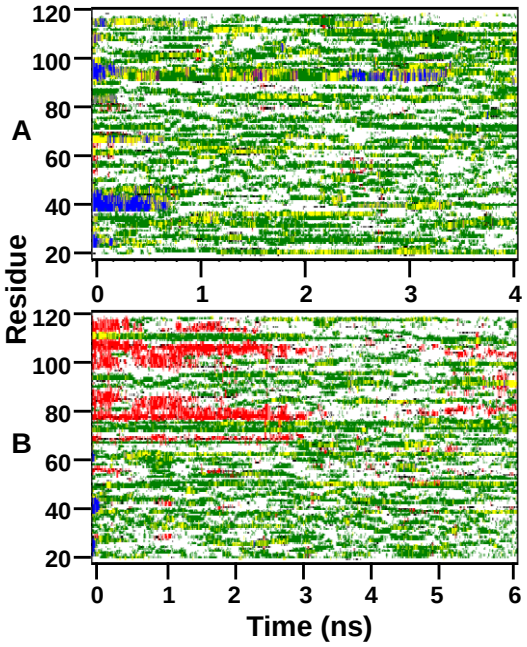


Figure 7. Change in secondary structure of (A) α-LA (B) RNase A during the simulation carried out at 700 K. The color codes represent: blue-α-helix, red-β-sheet, white-coil, black-β-bridge, green-bend, yellow-turn, violet-π-helix, grey-3_{10}helix.

In order to unfold the proteins, dynamic simulations of α-LA and RNase A were carried out at 700 K. At 700 K, the solvent accessible surface area (SASA) of α-LA increased to ~350 nm^2 from 190 nm^2 within 4 ns. However, the SASA of RNase A was around 280 nm^2 (Fig. 5). Therefore, RNase A unfolding simulation was further extended up to 6 ns, where the SASA value reached to 320 nm^2. From the triplicates of 4 ns (α-LA) and 6 ns (RNase A) simulations, it was observed that the maximum unfolding of α-LA

and RNase A was achieved at 700K. The unfolding was also verified with the increase in RMSD and decrease in secondary structure content of the proteins with increasing simulation time (Fig. 6 and 7).

Table 1. Structures selected at different time points from α-LA and RNase A unfolding simulations (Fig. 5)

Structure	Time	SASA	RMSD(nm)
		α-LA	
Lnative	0	165.97	0.00
LD1	1098	255.10	1.01
LD2	1974	322.74	1.32
LD3	3776	364.37	1.74
		RNase A	
Rnative	0	176.74	0.00
RD1	286	236.40	0.71
RD2	1984	264.96	0.99
RD3	5838	327.71	1.56

The secondary structure analysis suggested that α-LA lost its helices and β sheets within the initial 500 ps of the high temperature simulation whereas RNase A retained some of the secondary structures till 6 ns. All of the disulfide bonds remain intact in both the proteins throughout the simulation. From the simulation trajectory, four different conformational structures including the initial native conformation were selected for both α-LA (Lnative, LD1, LD2, and LD3) and RNase A (Rnative, RD1, RD2, and RD3) with increasing order of SASA and RMSD values (Table 1). The selected unfolded structures were quite distinct from their native conformations and also among each other in their secondary structure content (Fig. 8). The native form of α-LA has less β-sheet content than RNase A, but the helical contents of both the proteins were similar. The partially unfolded conformations of α-LA, i.e., LD1, LD2 and LD3, are completely devoid of β sheets, but the unfolded conformations of RNase A, i.e., RD1, RD2, and RD3 showed decreasing order of β-sheet contents. 3_{10} helices were not found in the unfolded conformations of both the proteins except in RD1. A substantial increase in the bend was observed in both the proteins suggesting that the proteins were significantly unfolded with the applied simulation conditions.

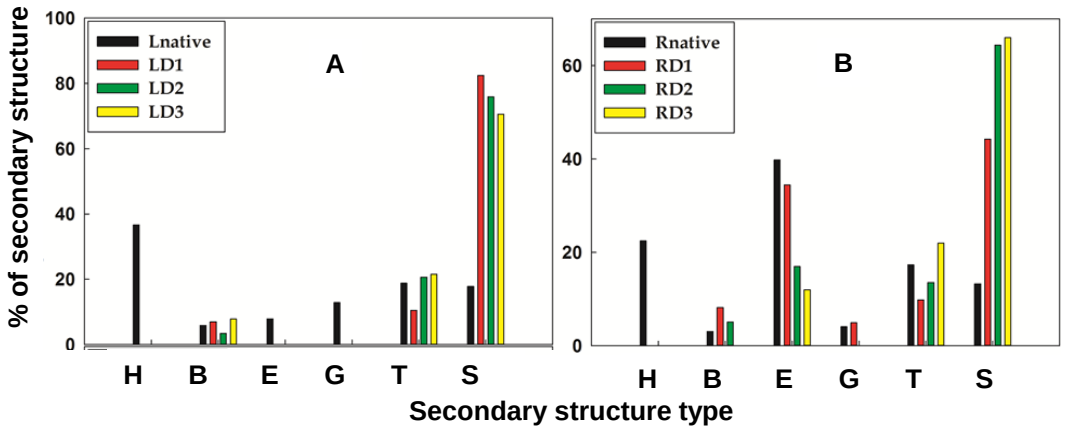


Figure 8. Secondary structure content of the selected conformations of α-LA (A) and RNase A (B) calculated using DSSP algorithm. H, B, E, G, T, and S represents alpha-helix, residues in isolated beta-bridge, extended strand participates in beta ladder, 3_{10} helix, hydrogen bonded turn, and bend respectively.

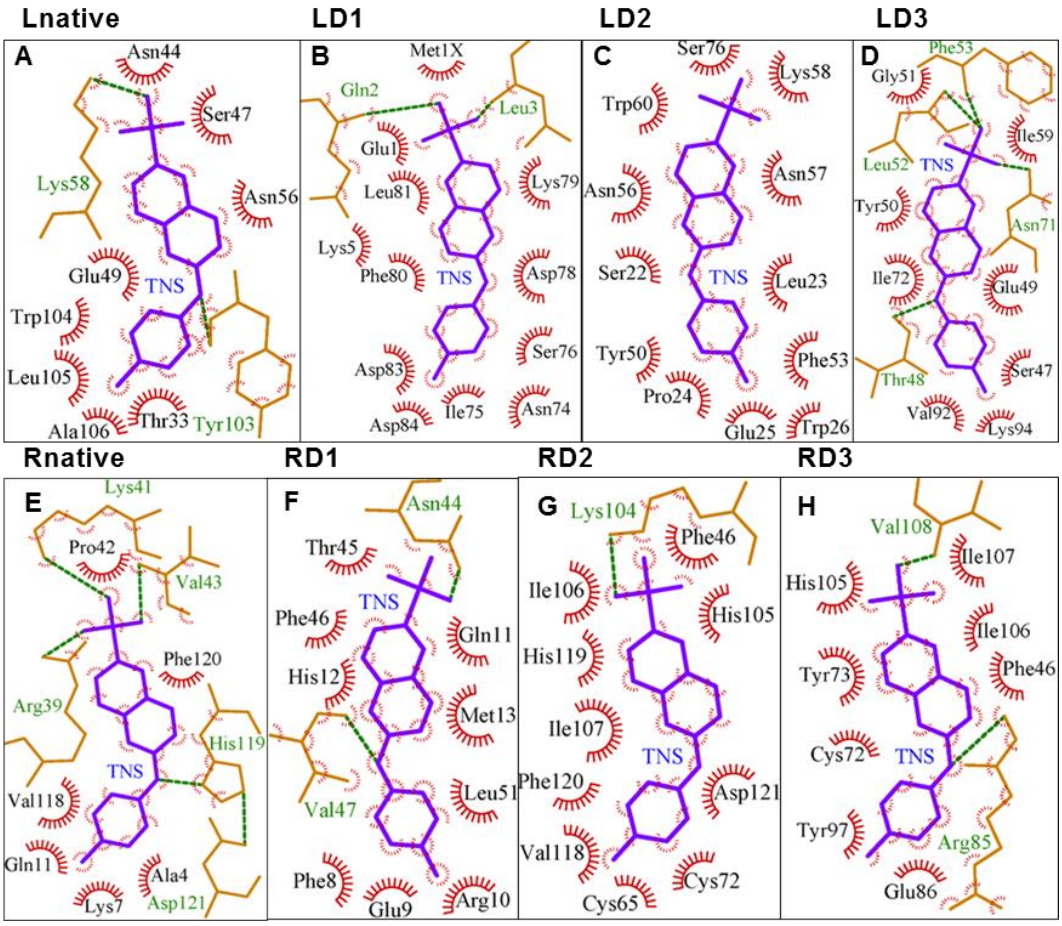


Figure 9. Two-dimensional representations of TNS docked with selected conformations of α-LA (A-D) and RNase A (E-H). The purple sticks represent TNS molecule. The red arcs labeled with amino acid codes are residues of the proteins interacting with the ligand atoms (noted with arcs) through hydrophobic interactions. The yellow sticks represent the residues involved in hydrogen bonding with ligands, and the H-bonds are shown in green dotted lines.

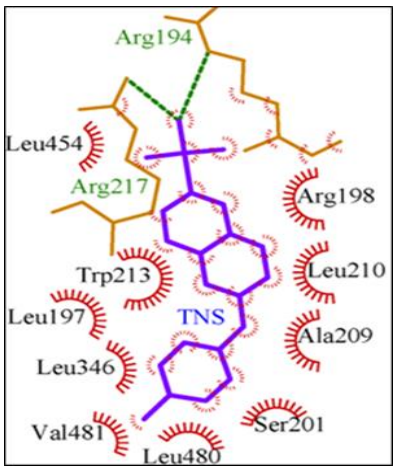


Figure 10. Two-dimensional representations of TNS docked against BSA. The color codes are same as in Fig. 9.

5.4.3 TNS Binding to Proteins

TNS was docked against the selected protein structures of α-LA and RNase A, and the native state of BSA. The docked complexes were analyzed using LigPlot+ (Fig. 9 and 10). From the docking results, it was evident that TNS was surrounded by more hydrophobic residues when it bound to the partially denatured structures of α-LA and RNase A, except LD3 which formed more H-bonding interactions. On the other hand, the native form of BSA showed more hydrophobic interactions than the native states of α-LA and RNase A with TNS. Further, the MD simulations of all the ligand-bound proteins were carried out for 2 ns at room temperature (300 K) to follow the dynamics of protein-TNS complexes. The RMSD vales obtained from molecular simulations are presented in Fig. 11. Table 2 lists all the simulations performed in this study.

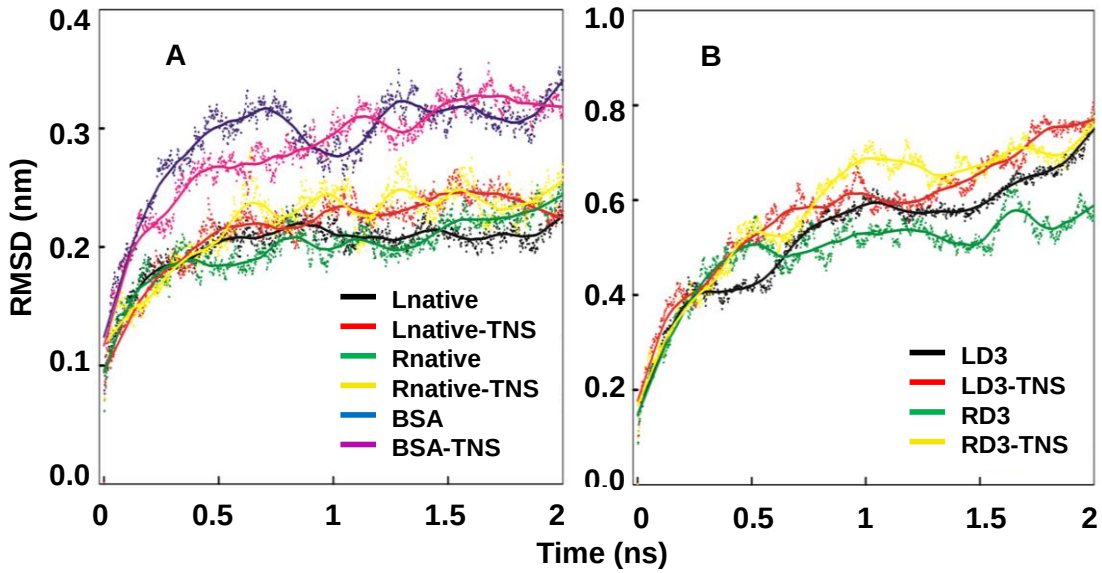


Figure 11. (A) RMSD plots of native forms of α-LA, RNase A and BSA and their TNS bound forms obtained from the MD simulations carried out at 300 K. (B) RMSD plots of representative unfolded conformations of α-LA and RNase A and their TNS bound forms. The solid lines are running average of the RMSD values.

Table 2. Summary of the simulations carried out.

Structural detail	Simulation time (ns)
Unfolding simulation (700 K)	
α-LA x 3	4
RNase A x 3	6
Protein-ligand Complex Simulation (300 K)	
α-LA	
Lnative without TNS	2
LD1 without TNS	2
LD2 without TNS	2
LD3 without TNS	2
Lnative with TNS	2
LD1 with TNS	2
LD2 with TNS	2
LD3 with TNS	2
RNase A	
Rnative without TNS	2
RD1 without TNS	2
RD2 without TNS	2
RD3 without TNS	2
Rnative with TNS	2
RD1 with TNS	2
RD2 with TNS	2
RD3 with TNS	2
BSA native	
BSA without TNS	2
BSA with TNS	2
Free TNS	
TNS	10

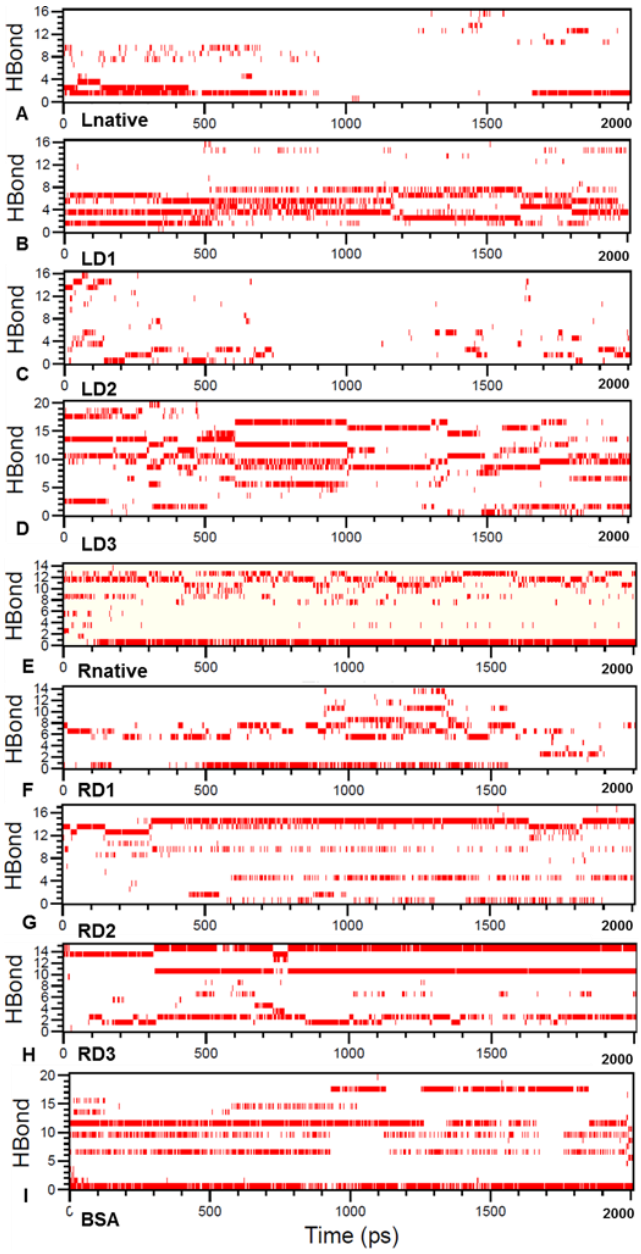


Figure 12. The account of all the H-bonds formed between TNS and the selected conformations of α-LA (A-D) and RNase A (E-H) and the native-BSA (i) during molecular simulation. The numbers on the ordinate correspond to the distinct H-bonds. The details of protein and ligand atoms involved in H-bond is given in Table 3-5.

5.4.4 H-bonding Profile of Protein-TNS Complexes

Since hydrogen bonding is one of the major contributors for protein-ligand interaction, the formation of stable H-bonds between proteins and TNS molecule was analyzed from simulation trajectories (Fig. 12). The atoms O20, O21, O22 of TNS act as

H-bond acceptors and the atom N7 acts as H-bond donor in TNS (Fig. 1). They form H-bonds with the amide backbone and/or the side chains of the proteins. In all the protein-TNS complexes, an average of 15 different H-bonds were observed at different positions during the simulation. Analysis of these bonds indicated that α-LA-TNS complexes did not show any consistent H-bond formation, except in LD1. However, RNase A and BSA complexes showed the presence of H-bonds persisting throughout the simulation, except RD1. Account of each H-bond is mentioned in Table 3-5.

Table 3. Atom wise description of H-bonds obtained from TNS-BSA complexes simulations (Fig. 12)

	BSA		
No	**Donor**	**Hydrogen**	**Acceptor**
1	TNS584N7	TNS584H27	SER201OG
2	ARG194NE	ARG194HE	TNS584O20
3	ARG194NE	ARG194HE	TNS584O22
4	ARG194NH2	ARG194HH21	TNS584O20
5	ARG198NE	ARG198HE	TNS584O20
6	ARG198NE	ARG198HE	TNS584O21
7	ARG198NE	ARG198HE	TNS584O22
8	ARG198NH2	ARG198HH21	TNS584O20
9	ARG198NH2	ARG198HH21	TNS584O21
10	ARG198NH2	ARG198HH21	TNS584O22
11	TRP213NE1	TRP213HE1	TNS584O20
12	TRP213NE1	TRP213HE1	TNS584O21
13	TRP213NE1	TRP213HE1	TNS584O22
14	ARG217NE	ARG217HE	TNS584O21
15	ARG217NH1	ARG217HH11	TNS584O21
16	ARG217NH2	ARG217HH21	TNS584O21
17	ARG217NH2	ARG217HH21	TNS584O22
18	SER453OG	SER453HG	TNS584O20
19	SER453OG	SER453HG	TNS584O21
20	SER453OG	SER453HG	TNS584O22

Table 4. Atom wise description of H-bonds obtained from TNS-α-LA complexes simulations (Fig. 12).

	Lnative			LD1			LD2			LD3		
No	**Donor**	**Hydrogen**	**Acceptor**	**Donor**	**Hydrogen**	**Acceptor**	**Donor**	**Hydrogen**	**Acceptor**	**Donor**	**Hydrogen**	**Acceptor**
1	TNS123N7	TNS123H27	ASN102O	TNS123N7	TNS123H27	SER76OG	SER22N	SER22H	TNS123O20	TNS123N7	TNS123H27	SER47OG
2	TNS123N7	TNS123H27	TYR103O	TNS123N7	TNS123H27	ASP78O	SER22N	SER22H	TNS123O21	TNS123N7	TNS123H27	SER47O
3	SER47OG	SER47HG	TNS123O21	GLN2N	GLN2H	TNS123O20	SER22N	SER22H	TNS123O22	TNS123N7	TNS123H27	THR48O
4	SER47OG	SER47HG	TNS123O22	GLN2N	GLN2H	TNS123O21	SER22OG	SER22HG	TNS123O20	TNS123N7	TNS123H27	ASN71OD1
5	ASN56ND2	ASN56HD21	TNS123O20	GLN2N	GLN2H	TNS123O22	SER22OG	SER22HG	TNS123O21	TNS123N7	TNS123H27	ASN71ND2
6	ASN56ND2	ASN56HD21	TNS123O21	LEU3N	LEU3H	TNS123O20	SER22OG	SER22HG	TNS123O22	GLY51N	GLY51H	TNS123O20
7	ASN56ND2	ASN56HD21	TNS123O22	LEU3N	LEU3H	TNS123O21	ASN56ND2	ASN56HD21	TNS123O20	GLY51N	GLY51H	TNS123O21
8	LYS58NZ	LYS58HZ1	TNS123O20	LEU3N	LEU3H	TNS123O22	ASN56ND2	ASN56HD21	TNS123O21	GLY51N	GLY51H	TNS123O22
9	LYS58NZ	LYS58HZ1	TNS123O21	LYS5NZ	LYS5HZ1	TNS123O20	ASN56ND2	ASN56HD21	TNS123O22	LEU52N	LEU52H	TNS123O20
10	LYS58NZ	LYS58HZ1	TNS123O22	LYS5NZ	LYS5HZ1	TNS123O21	ASN56ND2	ASN56HD21	TNS123N7	LEU52N	LEU52H	TNS123O21
11	TRP60NE1	TRP60HE1	TNS123O20	LYS5NZ	LYS5HZ1	TNS123O22	LYS58NZ	LYS58HZ1	TNS123O20	LEU52N	LEU52H	TNS123O22
12	TRP60NE1	TRP60HE1	TNS123O21	SER76N	SER76H	TNS123N7	LYS58NZ	LYS58HZ1	TNS123O21	PHE53N	PHE53H	TNS123O20
13	TRP60NE1	TRP60HE1	TNS123O22	LYS79NZ	LYS79HZ1	TNS123O20	LYS58NZ	LYS58HZ1	TNS123O22	PHE53N	PHE53H	TNS123O21
14	LYS98NZ	LYS98HZ1	TNS123O20	LYS79NZ	LYS79HZ1	TNS123O21	SER76OG	SER76HG	TNS123O20	PHE53N	PHE53H	TNS123O22
15	LYS98NZ	LYS98HZ1	TNS123O21	LYS79NZ	LYS79HZ1	TNS123O22	SER76OG	SER76HG	TNS123O21	TRP60N	TRP60H	TNS123O20
16	LYS98NZ	LYS98HZ1	TNS123O22	PHE80N	PHE80H	TNS123N7	SER76OG	SER76HG	TNS123O22	TRP60N	TRP60H	TNS123O21
17										TRP60N	TRP60H	TNS123O22
18										ASN71ND2	ASN71HD21	TNS123O20
19										ASN71ND2	ASN71HD21	TNS123O21
20										ASN71ND2	ASN71HD21	TNS123O22

Table 5. Atom wise description of H-bonds obtained from TNS-RNase A complexes simulations (Fig. 12).

	Rnative			RD1			RD2			RD3		
No	**Donor**	**Hydrogen**	**Acceptor**	**Donor**	**Hydrogen**	**Acceptor**	**Donor**	**Hydrogen**	**Acceptor**	**Donor**	**Hydrogen**	**Acceptor**
1	TNSN7	TNSH27	HIS12NE2	TNSN7	TNSH27	VAL47O	TNSN7	TNSH27	PRO42O	TNSN7	TNSH27	ARG85O
2	TNSN7	TNSH27	VAL118O	GLN11N	GLN11H	TNSN7	TNSN7	TNSH27	HIS119O	TNSN7	TNSH27	GLU86OE1
3.	TNSN7	TNSH27	HIS119ND1	GLN11NE2	GLN11HE21	TNSO20	TNSN7	TNSH27	ASP121OD1	TNSN7	TNSH27	GLU86OE2
4	GLN11NE2	GLN11HE21	TNSN7	GLN11NE2	GLN11HE21	TNSO21	TNSN7	TNSH27	ASP121OD2	VAL47N	VAL47H	TNSO21
5	ARG39NH1	ARG39HH11	TNSO20	GLN11NE2	GLN11HE21	TNSO22	TNSN7	TNSH27	ASP121O	VAL47N	VAL47H	TNSO22
6	ARG39NH2	ARG39HH21	TNSO20	ASN44ND2	ASN44HD21	TNSO20	ARG85NE	ARG85HE	TNSN7	HIS48NE2	HIS48HE2	TNSO22
7	ARG39NH2	ARG39HH21	TNSO21	ASN44ND2	ASN44HD21	TNSO21	ARG85NH1	ARG85HH11	TNSO21	LYS66NZ	LYS66HZ1	TNSO20
8	LYS41NZ	LYS41HZ1	TNSO20	ASN44ND2	ASN44HD21	TNSO22	ARG85NH1	ARG85HH11	TNSN7	LYS66NZ	LYS66HZ1	TNSO21
9	LYS41NZ	LYS41HZ1	TNSO21	THR45N	THR45H	TNSO20	ARG85NH2	ARG85HH21	TNSN7	LYS66NZ	LYS66HZ1	TNSO22
10	LYS41NZ	LYS41HZ1	TNSO22	THR45N	THR45H	TNSO21	LYS104NZ	LYS104HZ1	TNSO20	TYR73N	TYR73H	TNSO20
11	VAL43N	VAL43H	TNSO20	THR45N	THR45H	TNSO22	LYS104NZ	LYS104HZ1	TNSO21	TYR73OH	TYR73HH	TNSO20
12	VAL43N	VAL43H	TNSO21	THR45OG1	THR45HG1	TNSO20	LYS104NZ	LYS104HZ1	TNSO22	TYR73OH	TYR73HH	TNSO21
13	VAL43N	VAL43H	TNSO22	THR45OG1	THR45HG1	TNSO21	HIS105N	HIS105H	TNSO20	TYR73OH	TYR73HH	TNSO22
14	PHE120N	PHE120H	TNSN7	THR45OG1	THR45HG1	TNSO22	HIS105N	HIS105H	TNSO21	VAL108N	VAL108H	TNSO20
15							HIS105N	HIS105H	TNSO22	VAL108N	VAL108H	TNSO21
16							HIS119NE2	HIS119HE2	TNSO21			
17							ASP121N	ASP121H	TNSN7			

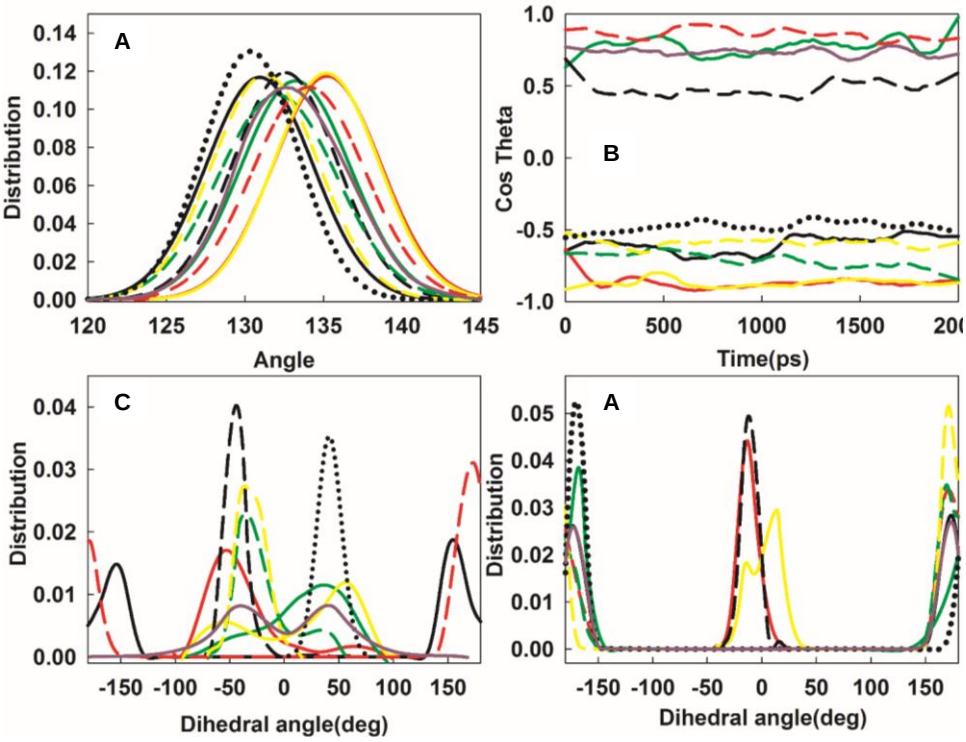


Figure 13. Change in the angles and dihedrals of TNS during the simulation. (A) Distribution plot of the angle between C8-N7-C1 which depicts the bending movement of TNS. (B) Cosine value of the angle between the perpendiculars drawn from naphthalene and toluene rings of TNS representing the degree of planarity between the rings. (C) and (D) Distribution plots of the dihedrals representing C13-C8-N7-C1 and C8-N7-C1-C2 which infers the rotation of naphthalene and toluene rings, respectively. Solid black, red, green and yellow lines represent TNS bound to Lnative, LD1, LD2 and LD3, respectively. Dashed black, red, green and yellow lines represent TNS bound to Rnative, RD1, RD2 and RD3 respectively. Solid purple and dotted black lines represent free TNS and TNS bound to BSA, respectively.

5.4.5 Bond Angles and Dihedrals of TNS

To probe the conformational freedom of protein bound TNS in terms of bending, rotation, and co-planarity of the aromatic rings, two major bond angles and dihedrals of TNS were analyzed. The bond angle formed by C8-N7-C1 atoms was defined as angle-1. It describes the bending movement of the aromatic rings in TNS. The cosine value of the angle between the orthogonal vectors drawn from the planes of naphthalene and toluene rings was defined as angle-2 which represents the rotational movement of the ring systems. The dihedral angles-1 and -2 were defined by the dihedrals made by C13-C8-N7-C1 atoms and C8-N7-C1-C2 atoms, respectively. Dihedral-1 represents the naphthalene ring rotation around C8-N7 bond and the dihedral-2 maps the rotation of toluene ring around N7-C1 bond. The change in angles and dihedral values of TNS in free-state and in all the protein-TNS complexes during the simulation were smoothed for further analysis. The results are presented in Fig. 13. The distribution of bending angle (angle-1) of TNS was similar in both the free (132.5°) and protein-bound states (132.5 ± 2.5°). From the angle-2 values of TNS in solvent, it is clear that both the rings are out of plane nearly by 40°. TNS bound to the native states of α-LA, RNase A and BSA showed

the angular values between 40° and 60°. When bound to LD1, LD3 and RD1, the angle-2 in TNS was fairly less than 30°, whereas in LD2, RD2 and RD3 the angle was around 40°. The most preferable conformation of free TNS has the dihedral-1 either at 40° or at -40° with equal probability in the absence of any bound molecules. The dihedral-1 of protein bound TNS states has less deviation from either of the most probable angles of free TNS, except Lnative and RD1 bound TNS. In the case of Lnative, dihedral-1 was more probable at ±150° and in RD1 a fairly sharp peak was obtained around ±170°. However, the peak width of the protein bound TNS is less compared to the free-TNS, except in the cases of LD2 and LD3. Also, BSA shows more restricted rotation of naphthalene ring with a narrow peak width having maximum at 40°. Further, most of the protein-bound TNS states and free TNS dihedral-2 was restricted at ±170°, except that the Rnative and LD1 were at -13° and LD3 was at 13°.

5.4.6 Solvent Count, Rotational Autocorrelation, and Radius of Gyration

The number of solvent molecules available within 0.35 nm radius around TNS in the free and protein-bound states were calculated (Fig. 14A). As expected, free TNS was surrounded with more number of water molecules than the protein-bound state. In general, TNS bound to the native proteins were surrounded with more solvent molecules than the TNS bound to the partially unfolded proteins. In the case of RD2 and RD3 the solvent counts were comparable to their native state, Rnative. In order to evaluate the rotational dynamics of TNS in various states, rotational autocorrelation decay rates $C(t)$ were derived from the simulation trajectories (Fig.14B). Though $C(t)$ of the free TNS rapidly decayed to zero as expected, in the protein-bound states the decay rates were decreased significantly. The least $C(t)$ decay was observed in RD3 bound TNS and the most rapid decay was observed in LD3 and LD2 bound TNS. In the other complexes, the $C(t)$ rate showed intermediate values. The radius of gyration (R_g) of the proteins and the protein-TNS complexes were calculated from their corresponding simulation trajectories. The mean values of R_g calculated from the trajectory showed that the R_g values of the completely unfolded proteins (LD3 and RD3) and native conformation of BSA were decreased when TNS bound to them. TNS bound to the other conformational states of proteins slightly increased the R_g values compared to their ligand-free states (Fig. 15).

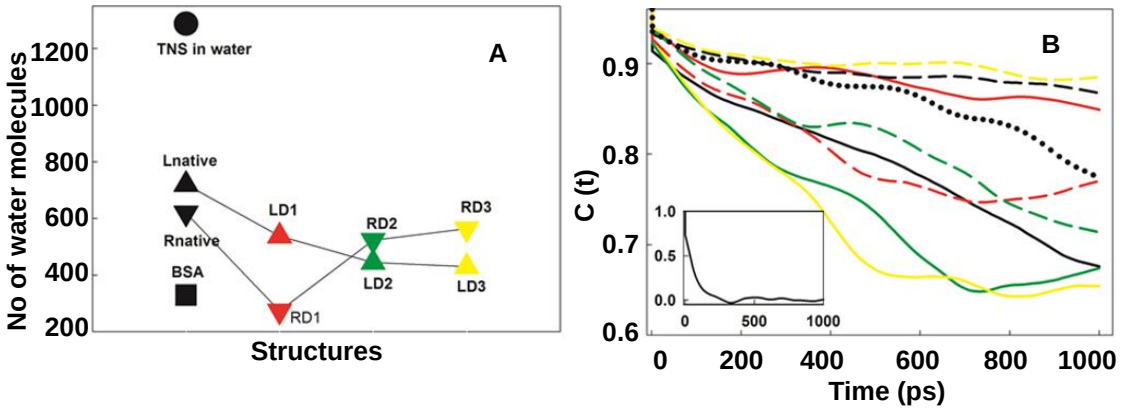


Figure 14. (A)The solvent count around the free and protein-bound TNS, within a radius of 0.35 nm. (B) The rotational autocorrelation function, *C*(*t*) of TNS molecule bound to different proteins. Solid black, red, green and yellow lines represent TNS bound to Lnative, LD1, LD2 and LD3, respectively. Dashed black, red, green and yellow lines represent TNS bound to Rnative, RD1, RD2 and RD3 respectively. Dotted black lines represent TNS bound to BSA. The inset shows C(t) of TNS in water. Here, the rotation is correlated with the vector change of perpendicular on the plane of naphthalene ring of TNS.

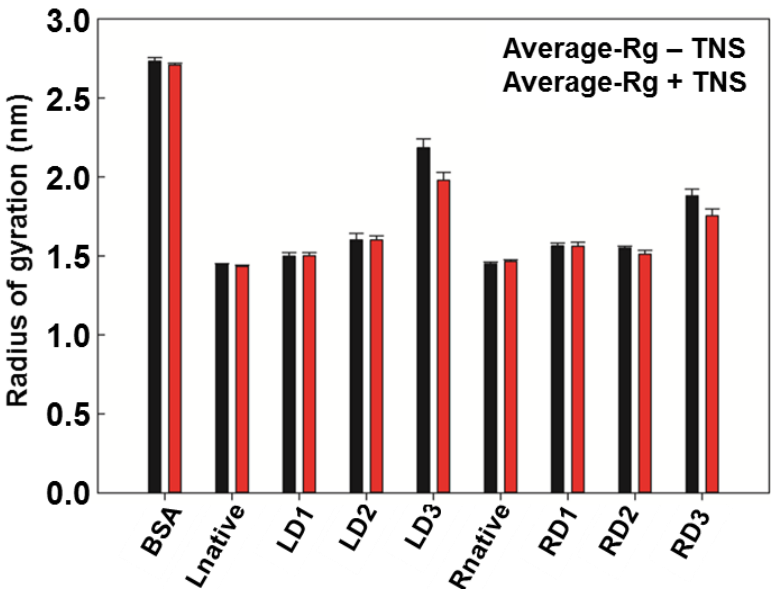


Figure 15. The mean values of radius of gyrations of the proteins and their TNS bound complexes obtained from the MD simulation performed at 300 K for 2 ns. Black bars represent the R_g values of protein alone and the red bars represent the R_g values of protein-TNS complexes.

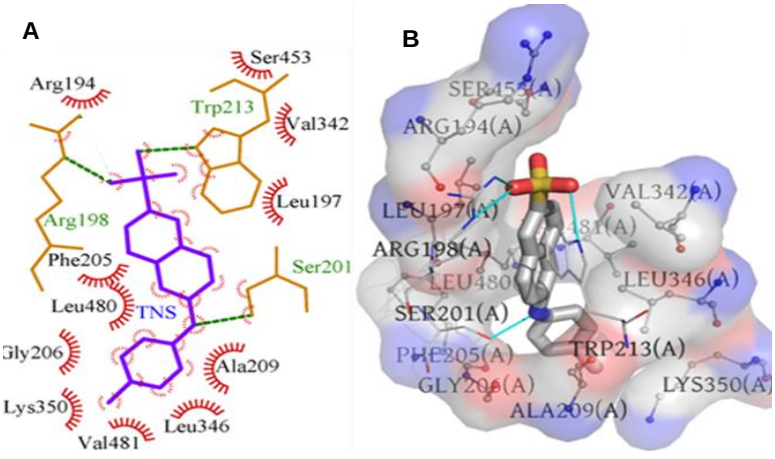


Figure 16. (A) 2D representation of TNS docked with BSA obtained from the average of 2 ns simulated conformations of BSA-TNS complex. (B) 3-D surface representation of TNS binding pocket in BSA obtained from the average of 2 ns simulated conformations. (Color codes are as describrd in Fig. 6)

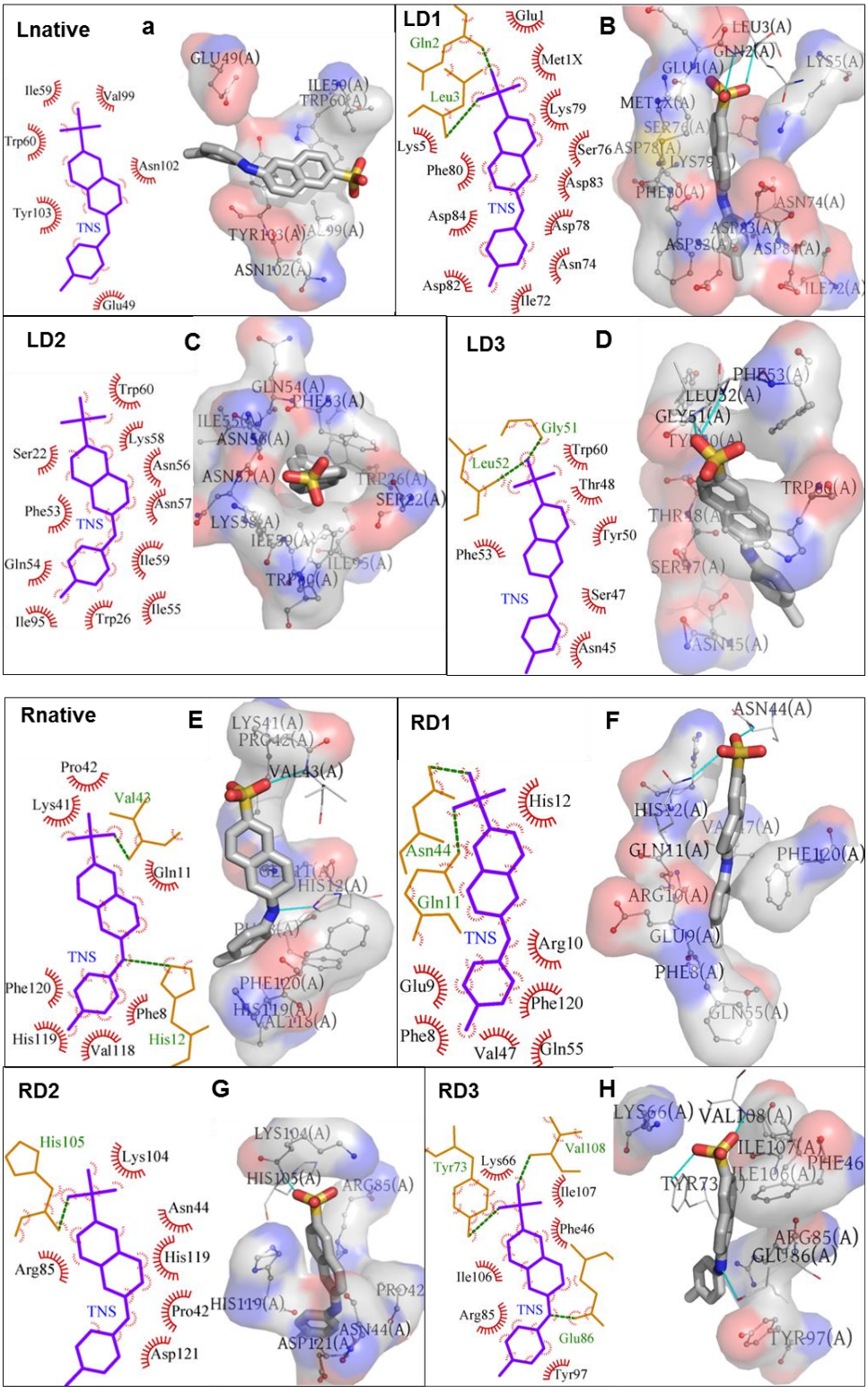


Figure 17. 2D and 3D representations of TNS bound to different conformations of α-LA (A-D) and RNase A (E-H) The purple sticks represent TNS molecule. The red arcs labeled with amino acid codes are residues of the proteins interacting with the ligand atoms (noted with arcs) through hydrophobic interactions. The yellow sticks represent the residues involved in hydrogen bonding with ligands, and the H-bonds are shown in green dotted lines.

5.4.7 Hydrophobic Interactions between Protein and TNS

In order to create an overall hydrophobic interaction profile, an average of the dynamic trajectory was considered as a representative of the most probable conformational state. Therefore, the average structures for each TNS bound native and partially unfolded conformations were derived and analyzed using LigPlot+ (Fig. 16 and 17). From these representative structures it could be inferred that TNS was surrounded by relatively less hydrophobic residues when it was bound to the native α-LA than the unfolded states of α-LA. But in RNase A, the native and unfolded states showed similar number of hydrophobic contacts. TNS bound to BSA showed considerably high number of hydrophobic contacts than Lnative and Rnative.

5.5 Discussion

5.5.1 Conformational States of Unfolded Proteins

The extent of unfolding at 700 K for α-LA and RNase A is not similar. RNase A retains some of the secondary structures. Two out of the four disulphide bonds in RNase A are formed between helix and sheets whereas in the case of α-LA all of the disulfide bonds are formed between flexible loops and helices. This might provide a higher stability to RNase A and some of the secondary structures, particularly part of the β-sheets, are retained even at longer time of the simulation at higher temperature (Fig. 7). Considering these factors, we have carefully chosen different unfolded conformations of the proteins based on their increasing SASA and RMSD values (Fig. 5-6). Further, the analysis of secondary structure content of the chosen conformations showed that the structure is gradually lost in both the proteins, though RNase A retains some amount of β-sheets (Fig. 8).

5.5.2 Protein-TNS Binding Interactions

From the docking results (Fig. 9-10), we could identify that TNS is preferably surrounded with hydrophobic residues when it is bound to partially unfolded proteins and in the case of native BSA. In order to understand the dynamical changes of these interactions, the mean of the simulation trajectory of the proteins in ligand bound forms are also analyzed (Fig. 16-17). This reinstates the fact that TNS is surrounded by more number of hydrophobic residues in the case of partially unfolded conformations of α-LA whereas in RNase A the number is not significantly increased compared to its native state. At the same time, the hydrogen bonds formed between unfolded conformations of RNase A and TNS are shown to be more stable during the simulation compared to the more fluctuations observed in all the conformations of α-LA and native RNase A over time. The unfolded conformations of RNase A retain some of the secondary structural contents. This might hinder the accessibility of hydrophobic interior of the protein to the TNS molecule. Thus, it contains less hydrophobic surrounding and the molecular binding

interactions are mostly stabilized by hydrogen bonding interactions. Comparing these interactions with the observed interactions in BSA suggest that TNS prefers to be in a hydrophobic environment. If the hydrophobic core is not accessible, the TNS may still form a stable interaction with the protein through hydrogen bonds.

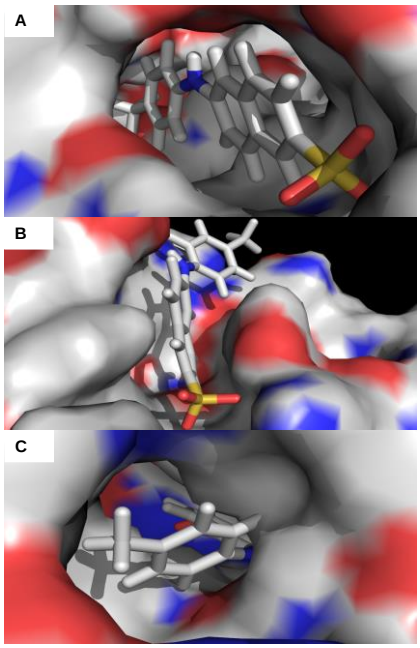


Figure 87. Orientation of aromatic rings of TNS in the binding cavities of the proteins (A) LD3, (B) RD3 and (C) BSA.

5.5.3 Conformational Dynamics and Solvent Accessibility of TNS

From the analysis of dynamics of angle-1 in TNS, it is evident that the bending angle around the nitrogen atom is highly constrained. In the protein bound state, TNS could show the bending of the ring systems to the maximum extent of 2.5° from the free TNS bending angle of 132.5°. However, the twisting of the rings measured through angle-2 suggests that the aromatic rings are more non-planar with an angle above 50°, when TNS is bound to the native forms of the proteins and in the case of RD3. Interestingly, when TNS is bound to the unfolded conformations of the proteins (except RD3) the twisting angle is reduced. In other words, the rings are moved towards more coplanar conformation. It is also observed that this shift towards co-planarity is also facilitated by the orientation of the binding pocket in the protein (Fig. 17-18). In order to facilitate the hydrogen bonding interaction with the protein through its sulfate and/or amino group, TNS might reorient itself. This could energetically favor the binding orientation of TNS towards the protein.

Dihedral angle-1 changes suggest that the accessible dihedrals of naphthalene ring are reduced in the protein-bound TNS compared to the free-TNS except in LD2 and LD3. The reduced conformational change is due to the hydrogen bonding interaction of TNS with proteins through the sulfate group attached to the naphthalene ring. The exceptions for TNS bound to LD2 and LD3 conformational states could be due to the fragile hydrogen bonding interactions (Fig. 12). In the case of BSA, the conformational rotation of naphthalene rings is highly restricted. This might be due to the binding of TNS into the hydrophobic cavity of the protein. Dihedral angle-2 which corresponds to the toluene ring rotation suggests that the rotation of toluene ring system is also highly restricted except in the case of LD3. The toluene ring is always found to be in a hydrophobic pocket, if not found on the surface, probably due to its smaller volume. This could reduce the conformational rotation of the ring system. Further, the rotational autocorrelation decay

also suggests that the TNS has restricted global rotation in the protein-bound forms, except in LD2 and LD3.

The number of accessible water molecules around TNS comes down, when it binds to unfolded conformations of the proteins, except RD3. This suggests that TNS preferably binds to the hydrophobic cavity of the proteins, thus having reduced water accessibility. In the case of BSA where TNS is buried in a hydrophobic pocket, the solvent accessibility comes down significantly compared to the other proteins.

5.5.4 Molecular Orientation *vs.* Hydrophobic Environment

The binding and MD simulation results suggest that TNS binding to the unfolded conformations of proteins has less water accessibility and mostly surrounded by hydrophobic residues. Though the hydrogen bond formation between TNS and the protein stabilizes the binding interaction to certain extent, it might not have considerable influence on the fluorescence changes. It is evident from the fact that TNS forms significant number of hydrogen bonds with the native states of RNase A and α-LA which do not increase the fluorescence quantum yield of TNS. Earlier studies on TNS interacting with peptides also suggested that initial ionic interaction might encourage the binding, but not sufficient enough to increase the fluorescence quantum yield [278, 301]. Further, the orientation of the ring systems may also have less contribution to the fluorescence changes observed for TNS. In the case of native-BSA bound TNS the ring orientation is non-coplanar as compared to the TNS bound to other unfolded conformations of the proteins, but still possess high fluorescence quantum yield. The angles and dihedral accessible for protein-bound TNS is much constrained and close to free-TNS conformational angles. Moreover, the toluene ring of TNS is found to be preferably surrounded by the aromatic residues of the proteins than the naphthalene ring which in most of the cases interacts with the protein via hydrogen bonding through the sulfate group. These binding preferences of TNS and the surface availability and orientation of protein binding pockets could influence the conformational freedom of TNS molecules. Therefore, we conclude that the surrounded hydrophobic pocket and the reduced polar solvent accessibility could be the major contributing factors for the increase in fluorescence quantum yield of TNS rather than the ring planarity and the other interactions.

From the fluorescence experiments it is also evident that the hydrophobic microenvironment was very crucial for the fluorescent behaviour of the TNS. The quantum yield of the TNS is found to decrease as the polarity of the medium is increased (Fig. 2). Similar results are obtained in the mixture of binary solvents where the content of the polar solvent is decreased, the quantum yield is increased with blue shit in the emission maximum (Fig. 3).

5.6 Summary

To probe the TNS interaction with proteins, binding of TNS with native and three different unfolded conformations of α-LA and RNase A were analyzed *in silico*. The unfolded conformations of α-LA and RNase A were obtained from the molecular simulations carried out at 700 K. Analysis of TNS conformational changes upon binding to different conformational states of the proteins suggest that TNS finds more restricted movements in the bound state and the conformational angles and dihedrals accessible by TNS is constrained upon binding. The orientation and planarity of aromatic rings are influenced by the structure of binding pockets which varies even among the conformations providing higher fluorescence to TNS. Though, ionic and hydrogen bonding interactions are seen almost similarly in all the conformational states of proteins, TNS bound to unfolded conformations has less solvent accessibility and mostly surrounded by hydrophobic residues. Therefore, the results suggest that the reduced solvent accessibility and increased hydrophobic packing around the TNS might enhance the quantum yield rather than binding capacity and ring orientation.

Chapter 6

Binding Modes of Fluorescence Probe TNS with Different Conformational States of Proteins: A Computational Study

6.1 Abstract

Characterization of different conformational states of proteins is essential to understand their stability and activity. Biophysical techniques aid in analysing these conformational states and molecular fluorescence is one of the most reliable and quickly accessible methods. Apart from the intrinsic fluorescence of proteins, external fluorescence dyes such as TNS, ANS, nile red and thioflavin are also used to characterize partially unfolded, aggregated and fibrillar states of proteins, though their exact molecular-level interactions with proteins are yet to be completely unraveled. The present study attempts to investigate the binding of TNS molecule on different conformational states of proteins. Unconstrained molecular dynamic simulation of 50 molecules of TNS with the native state of BSA, native and two partially unfolded states of RNase A and α-lactalbumin in water was carried out. Dynamic simulation of TNS alone in different solvents such as water, ethanol, DMF and DMSO was also performed. The binding environments in all the proteins and the solvents were analysed in terms of H-bonding interactions, order of contacts, amino acid specificity and conformational changes of TNS, and correlated with experimentally observed fluorescence changes of the dye. The results suggest that TNS forms aggregates in water whereas in non-aqueous solvents the order of aggregates are less which might result in enhancement in its fluorescence intensity. Further, TNS preferably binds with basic and aromatic amino acid residues of the proteins. In RNase A and α-lactalbumin, most of the TNS molecules tend to form aggregates even with the unfolded conformations of the proteins. However in BSA, the number of aggregated TNS molecules is less and TNS molecules in monomeric form are found in the hydrophobic crevices of the protein. This might result in enhancement of the fluorescence in BSA compared to the other proteins. The angles and dihedrals of TNS in different environments suggest that the bending movement between the naphthyl and tolyl rings is constrained whereas significant planar rotations could be observed both in solvents and in protein-bound states.

6.2 Introduction

Fluorescence spectroscopic methods are widely used in the field of biology to investigate the structure, dynamics and activity of macromolecules owing to their high sensitivity and reproducibility [318, 320-323]. The intrinsic fluorescence of proteins by its naturally occurring aromatic amino acid residues offers valid information on the conformational states and the activity of proteins [324, 325]. Covalently attached fluorescent moieties or externally added fluorescence dyes extend the applications by following the spectral and temporal changes of these fluorochromes in understanding the dynamics of bound or interacting proteins [281, 324, 326, 327]. According to the classical definition of Edelman and McClure, an external fluorescence probe is a small molecule, non-covalently interacting with the protein and showing difference in one or more of its fluorescence properties due to the change in its microenvironment upon binding [328]. This environment-sensitive fluorescence dyes has been widely used to monitor the

structural changes in biological membranes, to identify the intermediates during protein folding, to characterize the protein-protein and protein-ligand interactions, to follow protein aggregation and fibril formation and to estimate the activities of enzymes over the past several years [322, 324, 329-333]. For instance, the intensity ratio between third and first vibronic bands of pyrene is used to assess the polarity of the environment [334]. Nile red is used to identify large protein aggregates through spectroscopy and microscopy [335, 336]. Congo red and thioflavin T are used as histochemical stains to characterize amyloid fibrils [337]. Fluorescent molecular rotors such as dicyanovinyl-julolidine and coumarin derivatives are used to probe the local microviscosity of cellular membranes and to follow bio-polymerization processes [338, 339]. PRODAN, NBD, DMAP and their derivatives are other popularly known probes in biological, environmental and material sciences [340-342]. These external fluorescence probes along with their energy transfer properties with the intrinsic protein fluorochromes could provide detailed information about kinetic events during folding and aggregation of proteins [327, 343-345].

Among the various fluorescence dyes, *N*-arylaminonaphthalene sulfonate derivatives occupy a significant place in probing the conformational changes of proteins [330]. l-anilinonaphthalene-8-sulphonate (ANS) and its covalent dimer bis-ANS are most popular fluorescence probes of this class adapted for biophysical analysis of proteins. The photophysical properties and the binding interactions of these dyes with different proteins have been well studied using experimental and computation methods [346-348]. 2-*p*-toluidinylnaphthalene-6-sulfonate (TNS, Chap. 5 Fig. 1) is another widely recognized aminonaphthalene-based probe characterized for its solvatochromic effect. TNS is found to be almost fluorescence silent in aqueous medium whereas shows a high quantum yield in organic solvents with a hypsochromic shift as the polarity of the medium decreases [277, 328, 349]. This property has been exploited to probe the micellation properties of surfactants [350], membrane fluidity [331], the hydrophobic surfaces on the proteins [277, 328], conformational changes upon ligand binding [299] , ligand binding events [286, 351, 352], protein-surfactant interactions [285], multimeric protein assembly [353], protein aggregation, fibril formation [327, 343] and so on. Despite of their wider applications as biophysical probes, their interaction with biological molecules is yet to be clearly understood.

The binding of fluorescence probes such as ANS and TNS on the proteins' surface is assumed to be mostly hydrophobic and non-specific [354]. However, different experimental and computational studies [301, 347, 355, 356] along with crystallographic structures of protein-dye complexes [357] have emphasized on the specific role of electrostatic interactions as well in the binding. Further, these fluorescence dyes also tend to form aggregates at the concentrations which are most commonly used in biological investigations [319]. However, most of the studies have not considered the aggregation of fluorescence probes in water while analyzing the interactions between the dye and protein

molecules including our earlier studies (reported in Chapter 5). Fluorescence probes may bind to multiple sites on proteins with slightly differing environments and interactions [347, 358, 359]. These discrete multiple binding sites are characterized to be more for anionic probes such as TNS than cationic probes [360]. Though the fluorescence enhancement of the probes is induced by unfolded states of the proteins, the studies on binding of these probes with unfolded conformations of proteins are very limited [347, 361]. Further, addition of cosolvents such as urea and guanidine hydrochloride or increasing temperature which are commonly employed in biophysical methods to populate denatured states might directly influence the fluorescence property of the probes rather than exploring their interactions with the exposed hydrophobic surfaces of proteins during unfolding [358, 362]. Considering these factors, the present study attempts to investigate the probable binding interactions between TNS and proteins using molecular dynamic (MD) simulation method by including multiple number of TNS molecules and by using native and *in silico* denatured conformations of proteins. The results suggests that ionic interaction between TNS and positively charged amino acid residues initiate the binding and the positioning of aromatic rings in appropriate hydrophobic environments would further favour the binding. TNS aggregation and its exposure to solvent upon binding vary among the proteins and between the native and denatured conformations of the proteins. The most probable angles and dihedrals of TNS in solvents and in different proteins were also investigated. The results predict similarity in correlated, but constrained, movements of the two aromatic rings of TNS in solvents and upon binding to the proteins.

6.3 Experimental Methods

6.3.1 Structure of TNS and Proteins

The structure of TNS was drawn and energy minimized in Avogadro 1.1.1 using UFF[363]. Three model proteins bovine α-lactalbumin (α-LA), bovine ribonuclease A (RNase A) and bovine serum albumin (BSA) were selected for the study as they are functionally different and representing different fold classes and surface charges. Their native state conformations were obtained from protein data bank (PDB): 1FS3 (RNase A), 1HFZ (α-LA), and 4F5S (BSA). The *in silico* unfolded conformations of RNase A and α-LA were obtained using the protocol reported in the previous Chapter [361].

6.3.2 Molecular Dynamic Simulation

Molecular dynamic (MD) simulation of TNS molecules was carried out in the presence of three different proteins and four different solvents. The proteins were: native state of bovine serum albumin (BSA), native and two partially denatured states (A0, A2 and A3) of α-lactalbumin (α-LA), and native and two partially denatured states (R0, R2 and R3) of RNase A. 50 molecules of TNS were randomly placed around the protein, kept in a cubic box and then filled with water molecules. Also, 64 molecules of TNS

randomly placed in a cubic box was simulated in four pure solvent conditions: water, ethanol, DMF and DMSO. The systems were energy minimized and equilibrated as NVT and then as NPT ensembles. The final production simulations were run without any constrains at 300 K for 100 ns coupled with leap-frog integrator [364]. Short-range van der walls and short-range electrostatic cut-off was fixed at 1 nm. Long range electrostatics was maintained by particle mesh Ewald (PME) [152] scheme with order of 4 and spacing 0.16. Temperature coupling was treated by V-rescale[311] with time constant 0.1 ps. Parrinello-Rahman algorithm [153] was used to maintain the pressure at 1 bar with the time constant of 2 ps. LINCS algorithm [154] was used for all bond constraints. All the MD simulations were run in GROMACS 4.6.3 using GROMOS53a6 force field on a Unix based cluster of 42 cores, 3.33GHZ speed (per processor), Intel Xeon CPU. The list of simulations carried out is tabulated in Table 1.

Table 1. List of the simulations carried out.

Name	Protein	Solvent	Number of TNS	Simulation Time (ns)
BSA	+	Water	50	100
A0	+	Water	50	100
A2	+	Water	50	100
A3	+	Water	50	100
R0	+	Water	50	100
R2	+	Water	50	100
R3	+	Water	50	100
Water	-	Water	64	100
Ethanol	-	Ethanol	64	100
DMF	-	DMF	64	100
DMSO	-	DMSO	64	100

6.3.3 Analysis

From 100 ns MD simulations, the last 10 ns data were used for further analysis where the systems were stable in all the conditions. Types and number of aggregates, H-bonding interactions, order of contacts, amino acid specificity, electrostatic surface analysis, and angles and dihedrals of TNS were calculated by combining Gromacs tools with in-house codes written in R (for details refer Appendices B1-B3). Electrostatic surface was constructed with the help of PDB2PQR [158, 159] and APBS [160]. PyMOL [155] was used for visualization and to create structural graphics.

6.4 Results

6.4.1 Dynamics of TNS in Solvents and Proteins

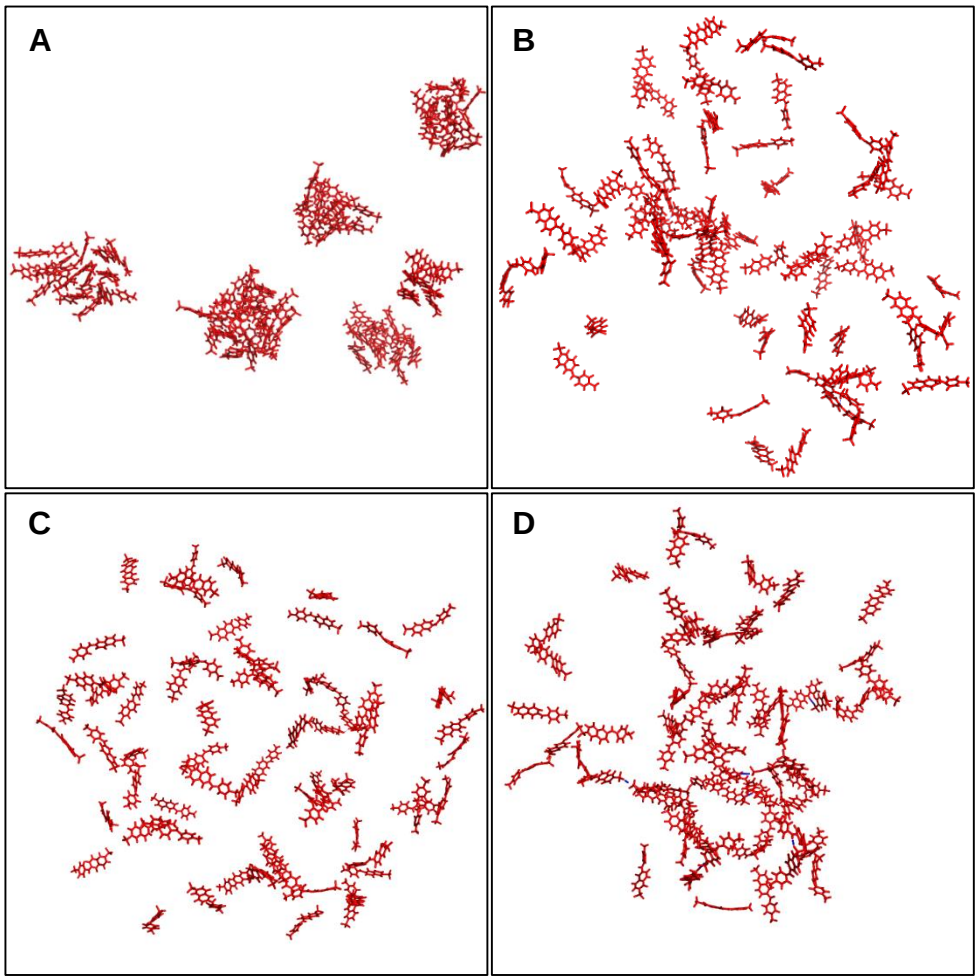


Figure 1. The snapshots of positions of 64 TNS molecules in (A) water, (B) ethanol, (C) DMF and (D) DMSO obtained at the end of 100 ns simulation.

In order to understand the conformational changes and aggregation behaviour of TNS in different solvent environments, 64 TNS molecules were simulated in water, ethanol, DMF and DMSO for 100 ns without any constrains. The last frames of all the simulations, representing the stabilized TNS-TNS interactions, are presented in Fig. 1. It clearly depicted that TNS formed larger aggregates in water whereas in non-aqueous solvents only monomers or smaller aggregates were found. In order to quantitatively measure the different aggregates formed during MD simulation, the fraction of aggregates with different sizes were calculated using the method described in Appendix B1. The analysis showed that the fraction of medium and higher order aggregates (more than four TNS molecules) are more in water whereas in non-aqueous solvents the fraction of higher order aggregates were minimal (Fig. 2A).

Similarly, to investigate the binding modes of TNS on different conformations of proteins, 50 molecules of TNS were simulated in the presence of native and two partially denatured conformations of α-LA and RNase A for 100 ns without any constrains. Since in the presence of BSA, TNS showed high quantum yield similar to its behaviour in DMF and DMSO, TNS was simulated in the presence of native conformation of BSA as well. The last frames of all the simulations are presented in Fig. 3-5. The simulation results suggested that all the three proteins could induce formation of TNS aggregates on their surface. However, there were significant differences in the fraction of distribution of aggregates. The quantitative analysis (Fig. 2B-C) showed that TNS preferably formed lower order aggregates in BSA where as in the presence of RNase A or α-LA the fraction of higher order aggregates were predominant. Upon denaturation of α-LA, the distribution of aggregates, in general, did not show any specific preference. However, as RNase A was denatured, the fraction of smaller aggregates increased over the higher order aggregates.

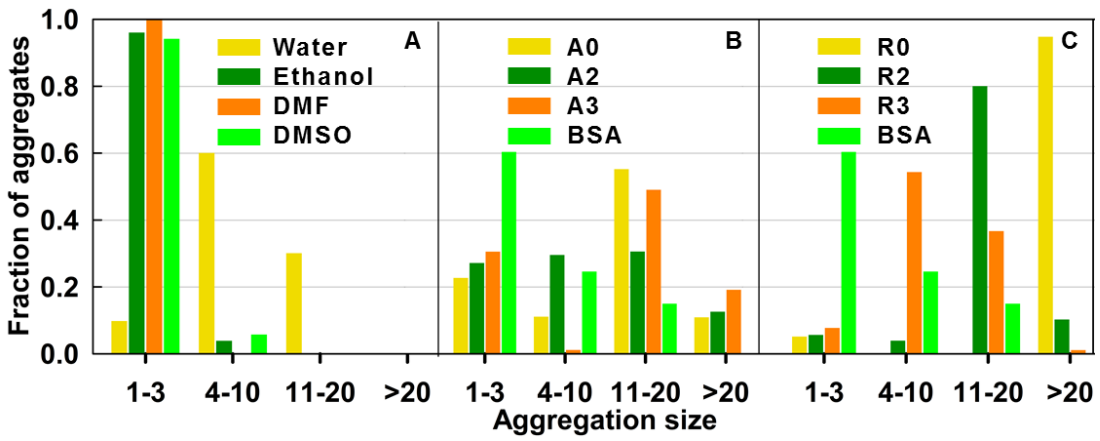


Figure 2. The aggregates of TNS formed in different conditions are categorized into four groups based on the number of TNS molecules in the aggregate (refer Appendix B1). The fraction of different sizes of aggregates of TNS (A) in solvents, (B) in α-LA in native (A0) and two denatured conformations (A2 & A3) and (C) in RNase A in native (R0) and two denatured conformations (R2 & R3). The fractions of TNS aggregates formed in the presence of native-state of BSA are plotted in panels (B) and (C) for comparison.

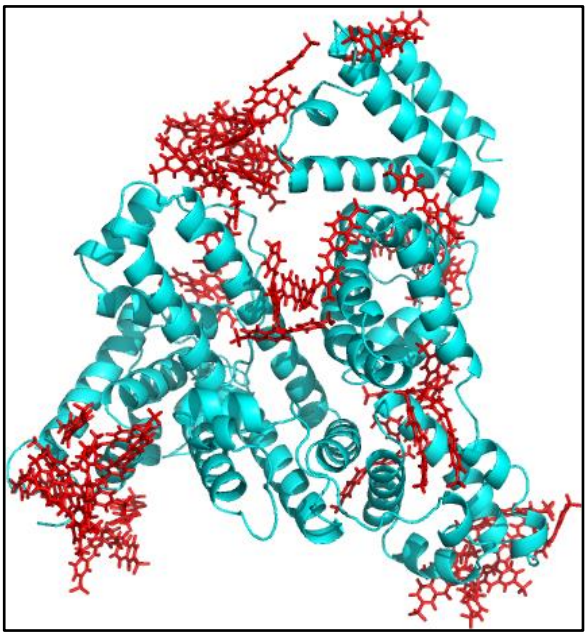
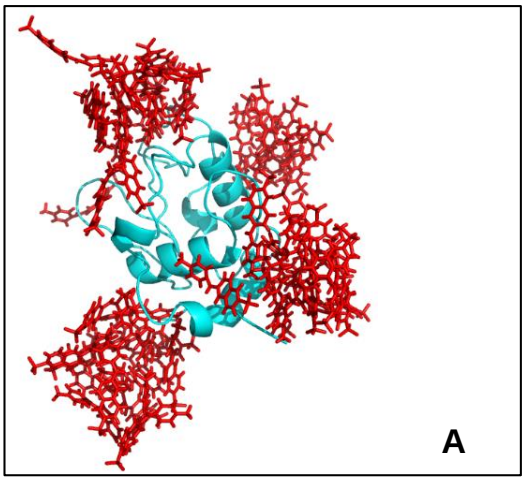
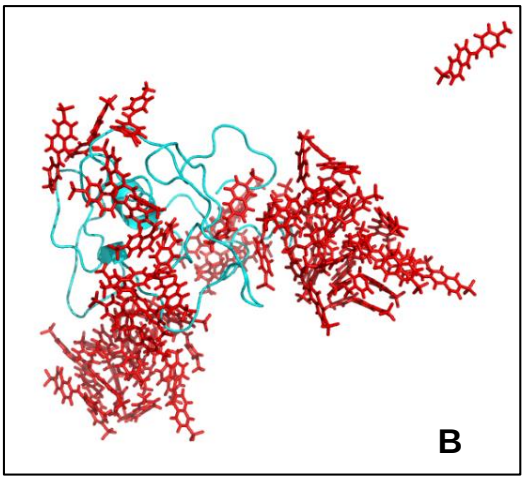


Figure 3 The snapshot of positions of 50 TNS molecules in the presence of BSA obtained at the end of 100 ns simulation. Cartoon diagram and stick models represent the protein and TNS, respectively.





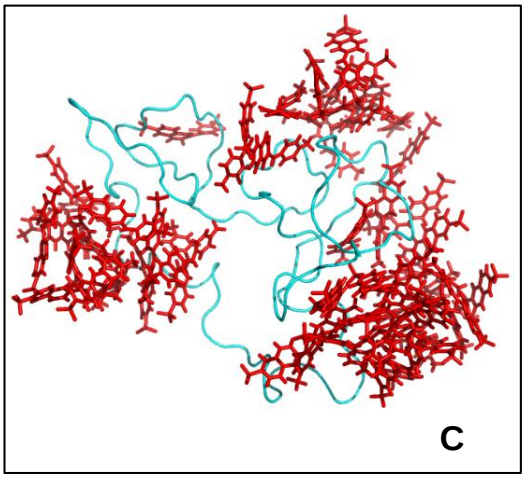


Figure 4 The snapshots of positions of 50 TNS molecules in the presence of (A) native and (B & C) two-denatured conformations of α-LA obtained at the end of 100 ns simulation. Cartoon diagrams and stick models represent the proteins and TNS, respectively.

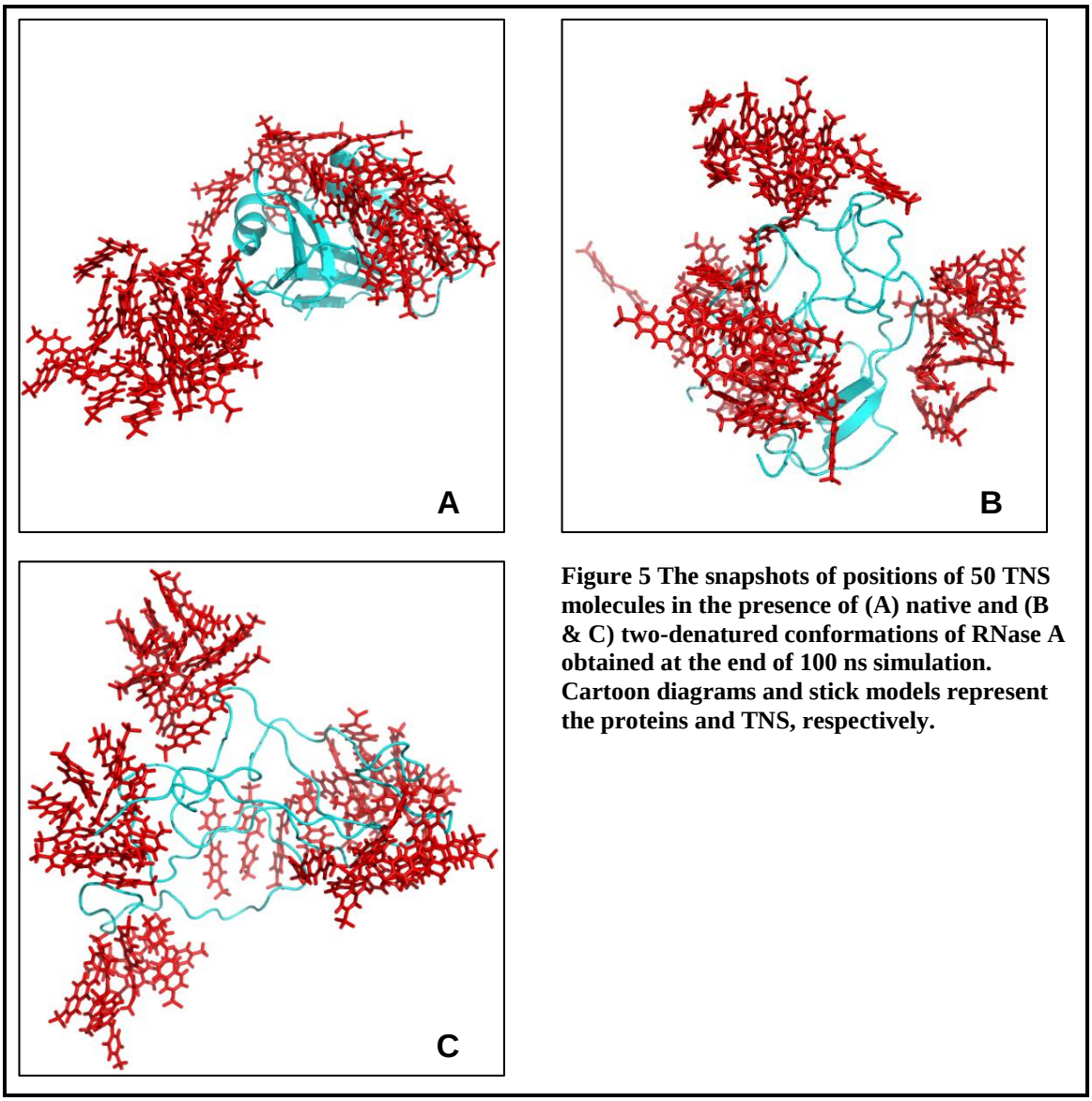


Figure 5 The snapshots of positions of 50 TNS molecules in the presence of (A) native and (B & C) two-denatured conformations of RNase A obtained at the end of 100 ns simulation. Cartoon diagrams and stick models represent the proteins and TNS, respectively.

6.4.2 Hydrogen Bonding Interactions of TNS

In order to analyse the interaction between TNS and the proteins, the number of H-bonding interactions of TNS with proteins and with solvents were calculated for last 10 ns of MD simulation using the method described in Appendix B2. Among all, the fraction of H-bonds formed by TNS was more with BSA followed by the denatured states of RNase A and α-LA (Fig. 6A). As expected, TNS formed higher number of H-bonding interactions with the solvent molecules (water) in all the cases irrespective of the proteins and their conformational states (Fig. 6B). For further insight, the fraction of H-bonds formed by TNS with different solvent molecules, water, ethanol, DMF and DMSO, were also calculated. The fraction was found to be above 0.95 in all the cases. The mean value of the fractions calculated from all the solvents is presented in Fig. 6B for reference. Comparison of Fig. 6A-B suggested that the fraction of H-bonds formed between TNS-solvent were reduced as the protein was added.

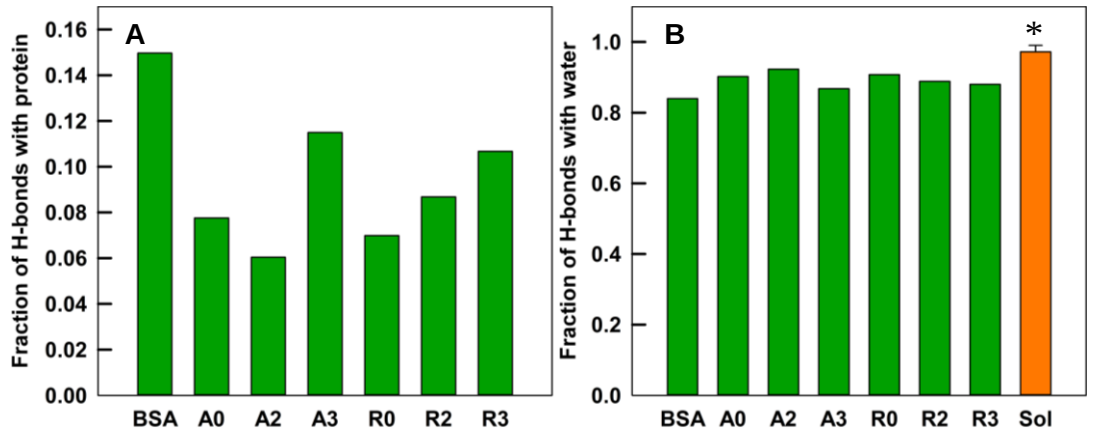


Figure 6. The fraction of H-bonds formed by TNS (A) with proteins and (B) with water in the presence of different proteins. The mean value of fractions of H-bonds formed by TNS in different solvents, i.e., water, DMF, and DMSO is presented as 'Sol' in panel B.

6.4.3 Order of Contacts

To probe the microenvironment around the TNS molecule, the numbers of other atoms around each TNS was calculated in different proteins and in different solvents as well. The contacts were classified into three types: TNS interacting with the protein atoms, TNS interacting with water and each TNS interacting with other TNS molecules. The numbers of contacts were counted and the fraction of each interaction over last 10 ns of MD simulation was evaluated by following the method mentioned in Appendix B3. The results are presented in Fig. 7A. The analysis clearly indicated that all the TNS molecules have contact with water in all the cases under study. Further, TNS-TNS interactions were found to be in significant fractions in all the conditions. Among all the proteins, BSA showed the highest fraction of TNS-protein contacts and the least fraction of TNS-TNS contacts. In α-LA, a slight increase in protein-TNS contacts was observed as the protein was denatured where as in RNase A, the changes were minimal.

In order to understand these interactions with respect to the self-aggregation property of TNS, the contacts were categorized into three types: zero order (no interaction), lower order and higher order contacts (more details in Appendix B3). All the three types of contacts for TNS with the proteins, solvent and other TNS molecules were calculated and are presented in Fig. 7B-D. In the case of DMF, DMSO and ethanol only zero order (no interaction with other TNS molecules) and lower order contacts were found. However, in water, almost 30% of the TNS molecules were participating in higher order contacts. Almost all the TNS molecules had contact with the solvent molecules. Also, the direct TNS-solvent contacts were less in water. These results suggest that TNS had more aggregation propensity in water whereas in non-aqueous solvents the numbers of aggregates were either reduced or not present. Further, the increased protein-TNS contacts in BSA could be attributed to the increase in lower order contacts suggesting that monomeric forms or smaller aggregates of TNS could be found on the protein (Fig. 3). This is complemented with the lowest number of higher order TNS-TNS contacts

observed in BSA. The slight increase in protein-TNS contacts in α-LA as the protein was denatured corresponded to increase in both lower order and higher order contacts around the protein which is complemented with slight decrease in higher order TNS-TNS contacts. This suggested that as the protein was denatured, the direct interaction between protein and TNS increased compared to TNS- protein bound TNS contacts (Fig. 4). In the case of RNase A, the higher order TNS-TNS contacts were reduced as the protein was denatured and the protein-TNS contacts and lower order TNS-TNS contacts were found to be increasing. This suggested that in the presence of denatured conformation of RNase A, the higher order aggregates of TNS were disturbed and the unfolded surface facilitated the formation of smaller TNS aggregates (Fig. 5).

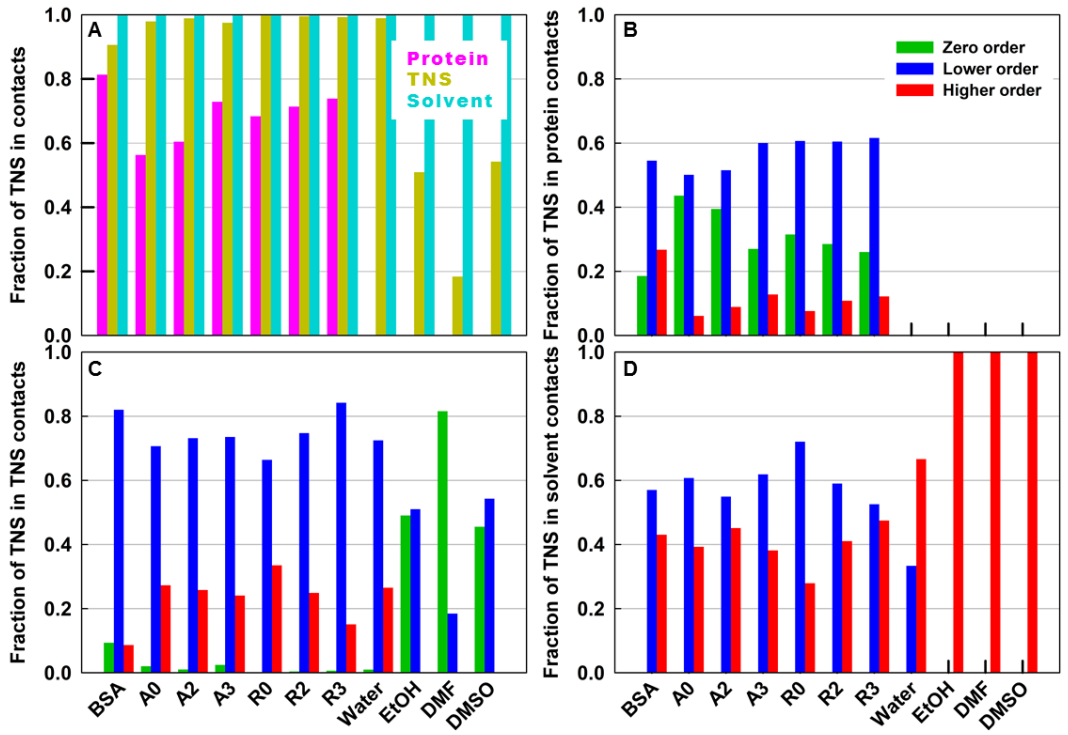


Figure 7. (A) The fraction of number of contacts of each TNS molecule with protein (pink), other TNS molecules (olive), and solvent (cyan) calculated in different solvents and in proteins. Further, each fraction of contacts, i.e., (B) TNS-protein, (C) TNS-other TNS molecules and (D) TNS-solvent is categorized into zero order (green), lower order (blue) and higher order (red) contacts based on number of neighbors (refer Appendix B3).

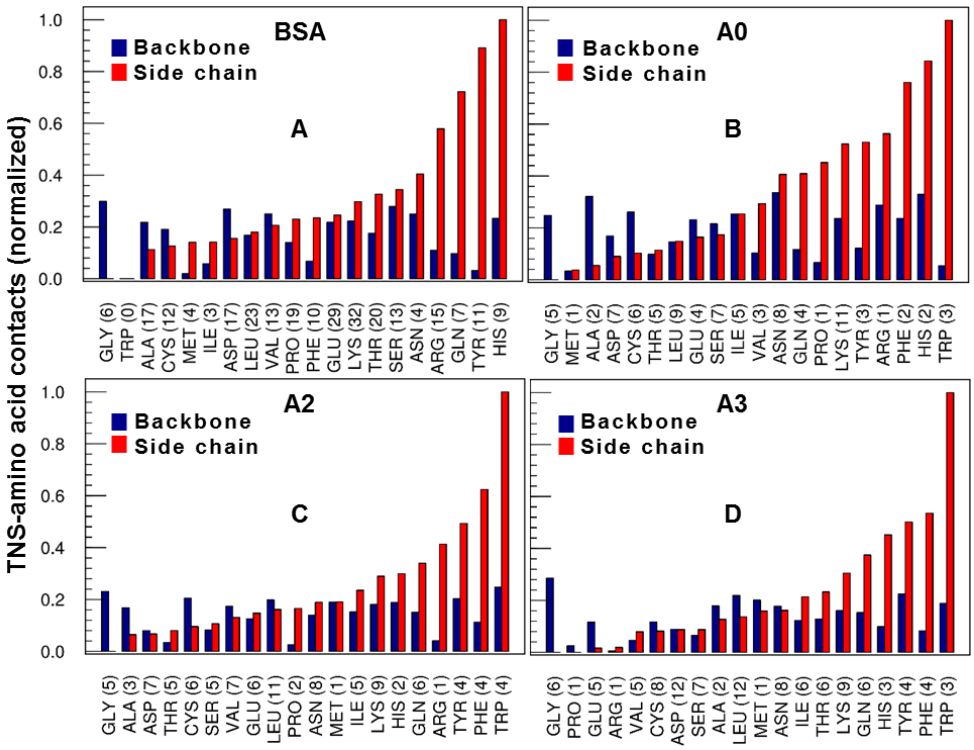


Figure 8. The number of contacts formed between TNS and backbone (blue) or side chain (red) atoms of each amino acid residue during the last 10 ns of MD simulation is counted and normalized with respect to the number of respective amino acids in the protein. The plots present the normalized TNS-amino acid contacts in (A) BSA, (B) native state of α-LA and (C & D) in denatured states of α-LA.

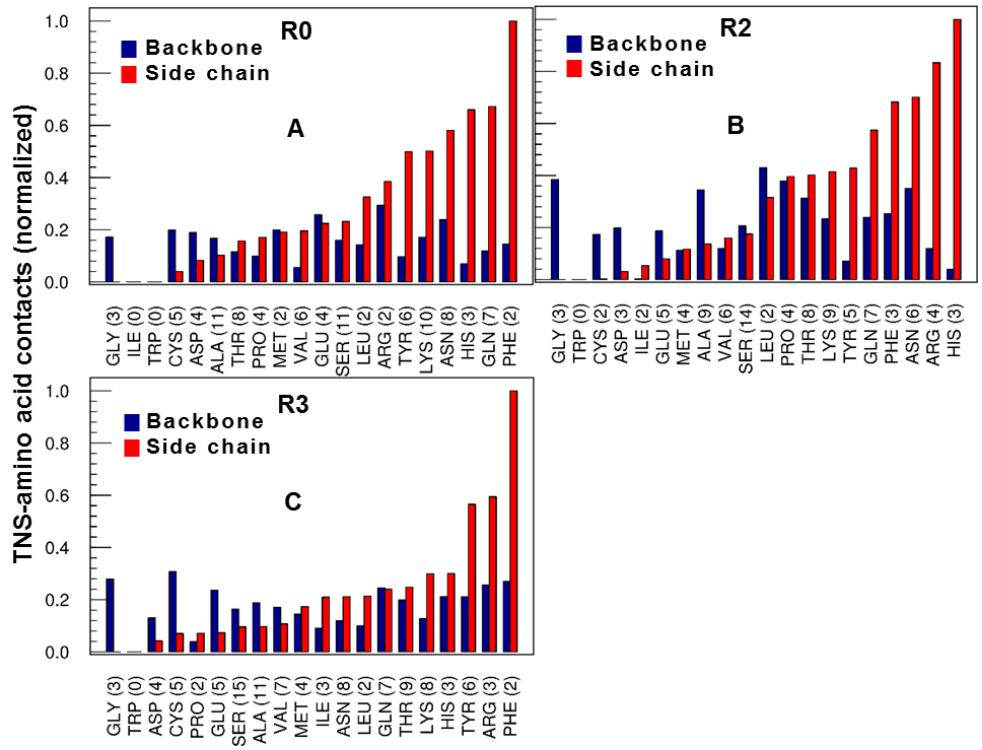


Figure 9. The normalized TNS-amino acid contacts in (A) native state of RNase A and (B & C) in denatured states of RNase A.

6.4.4 Amino Acid Specificity in TNS Binding

Though the binding of fluorescence probes such as ANS, bis-ANS, TNS, and PRODAN are assumed to be non-specific, there are studies suggesting the possible amino acid specificity in binding of these probes on proteins and selective increase in their fluorescence in such environments [301]. In order to obtain further insight, the amino acids interacting with TNS in all the proteins were analysed based on their contacts and normalized with respect to the number of each amino acid in the corresponding protein (Appendix B3) Interestingly, most of the interactions were found to be with basic residues (Arg, Lys and His), polar residues with amide side chains (Gln and Asn) and with solvent exposed aromatic amino acid residues (Trp, Tyr and Phe). In a few cases, significant interactions with proline were also found. In general, there were considerable amount of interactions with the peptide backbone of the protein chain as well which could be non-specific (Fig. 8-9).

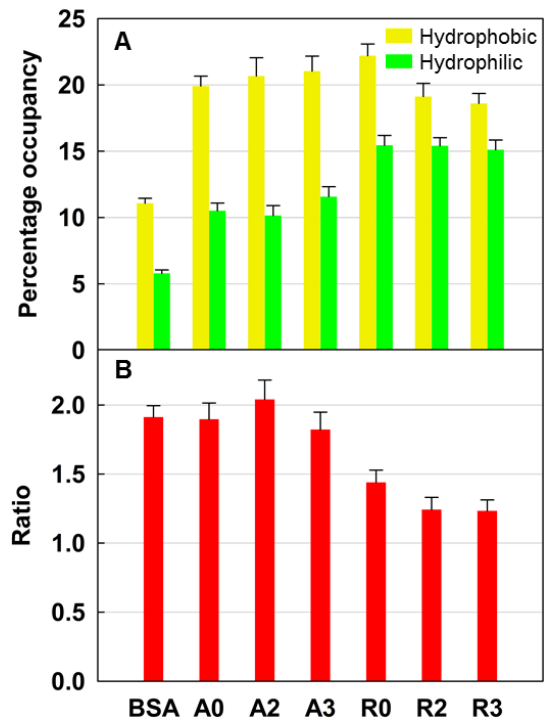


Figure 10. The solvent accessible surface area (SASA) of the proteins are classified in to hydrophobic and hydrophilic regions. (A) The percentage of hydrophobic (light green) and hydrophilic (dark green) surface areas occupied by TNS on different proteins. (B) The ratio of hydrophobic to hydrophilic surface area occupied by TNS in different proteins.

6.4.5 Solvent Accessible Surface Area (SASA)

To evaluate the surface area of proteins occupied by TNS molecules, the SASA values of the proteins was used as a measurement. The percentages of TNS occupied hydrophobic and hydrophilic SASA on the proteins were calculated (Fig. 10A) and the ratio of hydrophobic to hydrophilic SASA were also evaluated (Fig. 10B). The values suggested that TNS preferably occupied the hydrophobic surface patches on the proteins compared to the hydrophilic surfaces. In case of α-LA, upon denaturation of the protein, there was a slight increase in the occupancy of both hydrophobic and hydrophilic surfaces. In case of RNase A, the occupancy of TNS on the hydrophobic surface area was slightly reduced upon denaturation whereas the extent of occupied hydrophilic SASA remained the same.

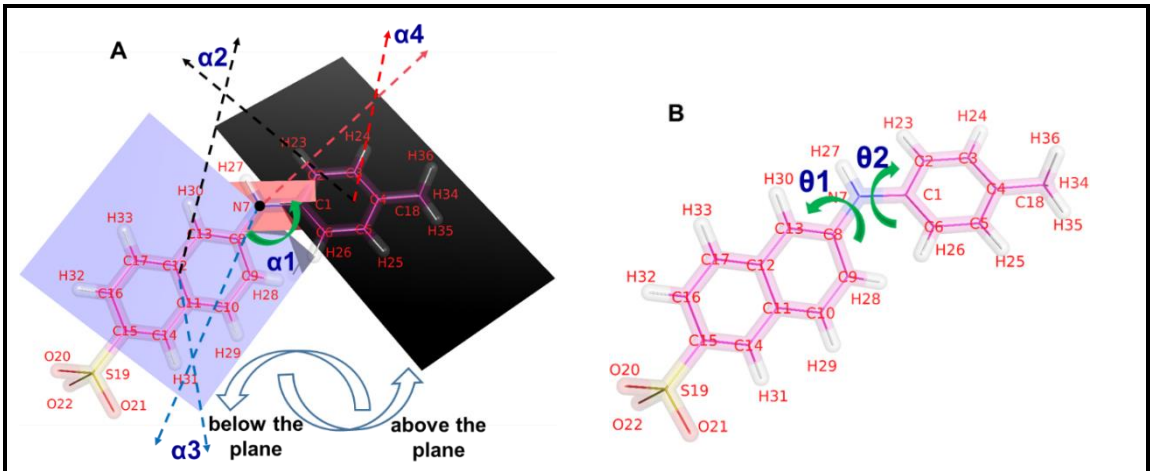


Figure 11. Four angles and two dihedrals defined for conformational analysis of TNS. (A) With reference to the center nitrogen atom N7 as the visual plane (pink), the naphthyl (blue) and tolyl (black) aromatic rings are represented below and above the plane, respectively. The angles defined are: α1 - angle between C8-N7-C1, α2 - angle between the planes of naphthyl (blue) and tolyl (black) rings, α3 - angle between the plane of naphthalene ring and the center nitrogen atom N7, and α4 - angle between the plane of tolyl ring and center nitrogen N7. (B) The dihedral angles defined are: θ1 - rotational angle of naphthyl ring with reference to the plane of central nitrogen N7 and θ2 - rotational angle of tolyl ring with reference to the plane of central nitrogen N7.

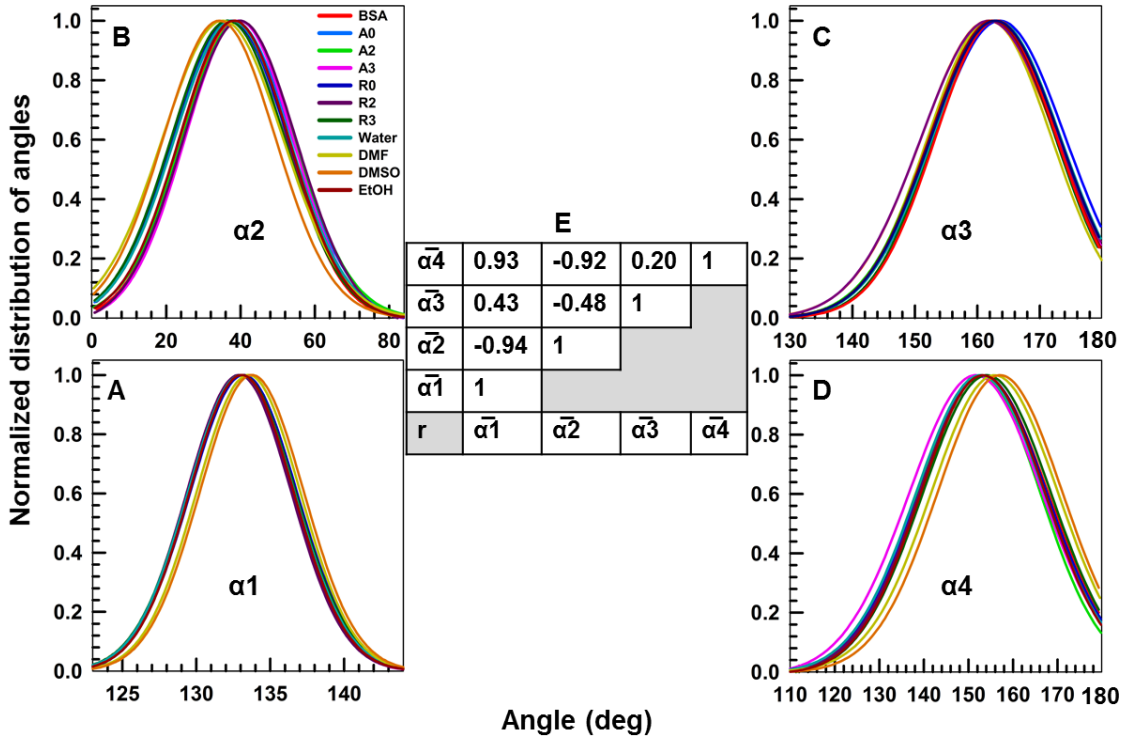


Figure 12. TNS conformational changes in different solution conditions is analyzed using four angles (refer Fig. 11). The distributions of TNS angles, clockwise from left bottom, (A) α1, (B) α2, (C) α3, and (D) α4 during MD simulation simulation in the presence of different solvents and the proteins and the color codes for different conditions are listed in panel B. The central panel (E) depicts the correlation-coefficients of the change in mean values between the angles (refer Table 2).

6.4.6 Angles and Dihedrals of TNS

To understand the conformational changes of TNS in solvents and upon binding to different states of the proteins, four angles and two dihedrals (Fig. 11) were investigated. The four angles are: α1 - the angle between C8-N7-C1; α2 - the angle between the planes of naphthyl and tolyl rings; α3 - the angle between the plane of naphthyl ring and nitrogen N7 at the centre; and α4 - the angle between the plane of tolyl ring and nitrogen N7 at the centre. The two dihedrals are: θ1 - rotational angle of naphthyl ring with reference to the plane of central nitrogen N7 and θ2 - rotational angle of tolyl ring with reference to the plane of central nitrogen N7. The calculated angles are shown in Fig. 12. The distributions of angles were fitted into Gaussian model and the calculated mean and standard deviation (SD) are presented in Table 2. In all the solvents and in the proteins, TNS showed similar distributions of α1 angle except in DMF and DMSO where the mean of the distributions of α1 were slightly increased. Notable changes were observed in the mean of α2 in different conditions, however, in DMF and DMSO the values were found to be exceptionally lower. The mean values of angle α3 varied between 161.8° and 163.6° with the lowest value in the unfolded conformation of α-LA (A2) to the highest value in RNase A (R0). There were evident differences in the mean values of α4 in different conditions with exceptionally higher mean values in the presence of DMF and DMSO. The correlation between the mean values of all the four angles of TNS in different environments were calculated (Fig. 12E). There is a negative correlation between the change in angles α1 and α2, i.e., if the bond angle between the two ring systems with respect to the central nitrogen was reduced, the planar angles between the rings would be increased. However, the change in α1 angle positively correlated with α4 suggesting that when the bond angle between the rings was reduced the planar angle of tolyl ring would be reduced with reference to the central nitrogen atom. Further, the angles α2 and α4 showed a negative correlation suggesting that increase in the angle between the two ring systems would decrease the planar angle of the tolyl ring with respect to the atom N7.

Similarly, the distribution of dihedral angles θ1 and θ2 were also calculated and fitted to bimodal distribution (Fig. 13 and Table 3). In all the environments, TNS showed two preferable dihedral angles with different fractions of distributions. The dihedral angle θ1 was centered around 164° and 195° with an exception for unfolded conformation of RNase A (R2) in which the dihedral was centered around 158.1° and 192.6°. The dihedral angle θ2 was centered around 169.7° and 190.6° with less variation in the distributions (less FWHM values) compared to dihedral θ1.

Table 2. Distribution of four angles of TNS (as defined in Fig. 11) during MD simulation in the presence of different solvents and the proteins

Protein/ Solvent	α1			α2			α3			α4		
	μ	σ	ƒ	μ	σ	ƒ	μ	σ	ƒ	μ	σ	ƒ
BSA	132.989	3.556	8.374	38.353	14.214	33.471	161.970	9.940	23.406	153.169	14.734	34.697
A0	133.130	3.555	8.372	38.037	15.942	37.540	162.726	10.379	24.440	152.905	14.264	33.588
A2	132.913	3.526	8.303	39.456	15.098	35.552	161.799	10.789	25.407	152.481	13.648	32.139
A3	132.937	3.513	8.273	39.573	14.539	34.235	162.676	10.199	24.016	151.649	14.936	35.171
R0	133.200	3.605	8.488	37.381	15.383	36.223	163.569	10.573	24.897	153.254	14.296	33.664
R2	132.895	3.520	8.288	39.775	15.251	35.913	161.872	10.879	25.619	153.074	14.060	33.109
R3	133.072	3.632	8.552	36.624	15.284	35.991	162.744	9.7890	23.051	154.143	14.411	33.935
Water	133.034	3.592	8.458	37.695	15.455	36.393	163.074	10.286	24.221	152.656	14.387	33.879
DMF	133.528	3.498	8.238	35.210	16.172	38.082	162.508	10.168	23.944	155.878	14.175	33.379
DMSO	133.756	3.515	8.277	34.305	15.087	35.528	163.014	9.726	22.903	157.087	14.015	33.003
Ethanol	133.069	3.513	8.272	38.237	14.722	34.668	162.884	10.34	24.349	153.255	13.964	32.884

Table 3. Distribution of two dihedral angles of TNS (as defined in Fig. 11) during MD simulation in the presence of different solvents and the proteins

Protein / Solvent	θ1						θ2					
	μ1	σ1	ƒ1	μ2	σ2	ƒ2	μ1	σ1	ƒ1	μ2	σ2	ƒ2
BSA	163.094	13.357	31.453	195.334	11.668	27.476	171.274	8.736	20.572	192.031	7.283	17.150
A0	163.798	12.044	28.361	195.133	13.725	32.319	168.626	7.491	17.640	189.640	8.323	19.599
A2	163.120	13.171	31.016	194.916	13.706	32.275	169.105	8.040	18.932	191.006	7.944	18.708
A3	164.753	12.341	29.060	195.874	13.642	32.124	169.538	8.459	19.918	191.154	7.806	18.381
R0	164.856	13.131	30.920	193.719	13.734	32.341	169.516	7.612	17.924	190.737	8.734	20.567
R2	158.113	11.356	26.741	192.642	15.213	35.825	171.239	8.796	20.713	192.253	7.063	16.631
R3	165.260	11.437	26.932	195.023	12.228	28.796	168.896	6.990	16.460	189.415	8.424	19.837
Water	166.604	13.662	32.171	195.479	13.110	30.87	169.662	8.373	19.716	191.392	8.181	19.265
DMF	165.279	13.083	30.807	196.435	12.827	30.204	169.502	7.908	18.621	189.338	8.723	20.542
DMSO	164.920	12.397	29.192	194.521	12.070	28.424	170.557	7.810	18.391	188.852	8.521	20.065
Ethanol	163.708	12.896	30.367	195.485	12.522	29.487	168.479	7.875	18.545	190.498	8.288	19.517

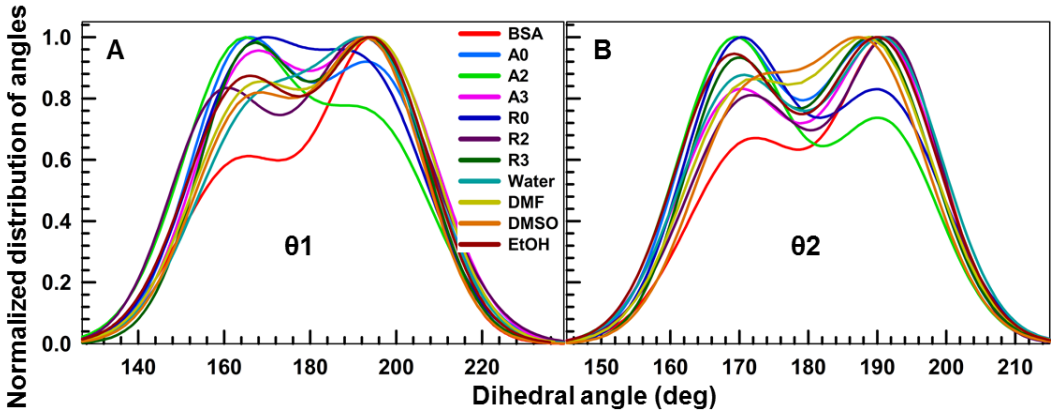


Figure 13. TNS conformational changes in different conditions analyzed using two dihedrals (refer Fig. 11). The distribution of dihedral angles (A) θ1 and (B) θ2 during MD simulation in different solvents and proteins as mentioned in the label.

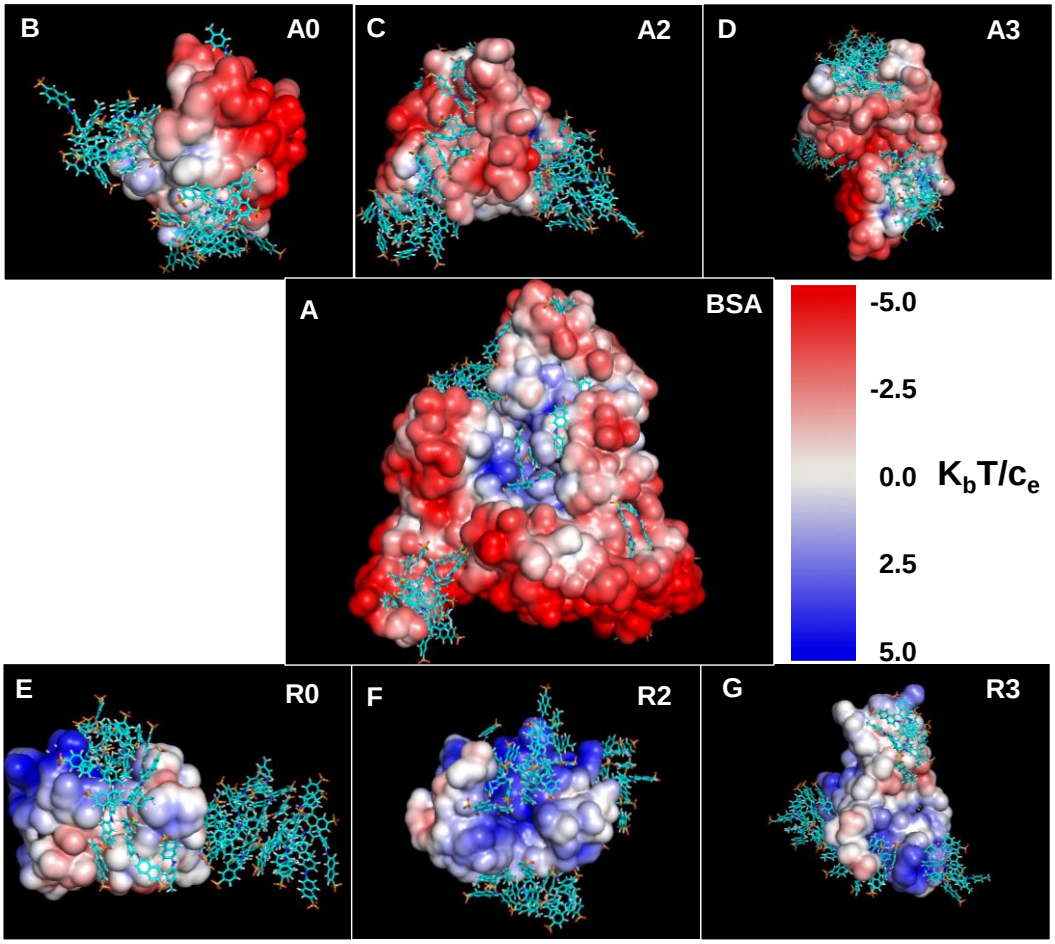


Figure 14. Electrostatic surface maps of (A) BSA, (B) native state of α-LA, (C & D) denatured states of α-LA, (E) native state of RNase A, and (F & G) denatured states of RNase A constructed using APBS tool are presented as surface models with bound TNS molecules as stick models.

6.4.7 Electrostatics Surface Analysis

In order to analyse the preferential binding regions of TNS on the protein surfaces, the representative electrostatics maps of the model proteins were created (Fig. 14). It was

consistently noticed that TNS preferably binds around the hydrophobic surface of the proteins and with the positively charges regions on the proteins. In the case of BSA, the aggregated forms of TNS molecule were found to be less whereas in the native forms of RNase A and α-LA a significant amount of TNS aggregates were observed. As RNase A and α-LA were denatured, more hydrophobic surfaces of the proteins were exposed which led to the redistribution of fraction of different aggregate forms of TNS molecules. Even in these cases, TNS preferably bound to the hydrophobic surfaces preferably adjacent to the positively charged regions of the proteins.

6.5 Discussion

6.5.1 TNS in Solvents

TNS fluorescence emission showed polarity dependent solvatochromic effect. As the polarity scale was increased, WL_{max} was shifted to longer wavelength. The highest quantum yield and exceptional shift in WL_{max} observed in the presence of DMF and DMSO (Chap. 5, Fig. 2) indicate that the solvent molecules with proton accepting group could enhance the quantum yield of the dye. MD simulation of TNS molecules in water suggests that the fraction of TNS molecules in free non-aggregated state is very low and the aggregated forms are more prevalent (Fig. 2A and 7C) which could quench the fluorescence. Nevertheless, in ethanol and DMSO, only lower order aggregates and monomeric forms could be found almost in equal proportions. In DMF, even the fraction of the lower order aggregates is significantly reduced (Fig. 2A and 7C). This is complemented by the decrease in solvent contacts of TNS in water compared to other non-aqueous solvents (Fig. 7D). Formation of such aggregates in water and destabilization of aggregates in non-aqueous solvents have been proposed in earlier studies as well [315]. In fact, self-aggregation of structurally similar dye ANS was also proposed to be weakened in the solvent with proton-accepting group such as dioxane (in the present study DMF and DMSO) and showing enhanced fluorescence emission [319]. These results support that the computational methods and the parameters used to probe TNS properties in different conditions by MD simulation are within the acceptable limits only. Comparing these results with experimentally measured fluorescence emission intensities and with earlier studies suggest that the aggregation of TNS molecules in water might quench the fluorescence emission probably due to partial exclusion of water as attributed in the case of ANS [365]. Further, similar environments for TNS could be found in the presence of proteins which could alter its fluorescence properties in the same manner (discussed below).

6.5.2 TNS in Different Conformations of Proteins

A simple count of interaction of each TNS molecule with proteins suggests that BSA has maximum number of TNS contacts and most of the contacts are in the form of lower order aggregates. However, in RNase A and α-LA, the monomeric form of TNS

bound on the surface of the proteins is reduced at the cost of increasing higher order aggregates (Fig. 2 - 5). This is complemented with the increase in TNS-TNS contacts in these proteins (Fig. 7C). Comparing these results with experimental results suggest that formation of aggregates on the surface of the protein would quench the fluorescence whereas binding of TNS in monomeric form or as lower order aggregates at the interior of the protein would enhance the fluorescence. Similar self-aggregation dependent fluorescence change upon binding to protein is proposed for ANS as well [319].

Further analysis on the specific interactions between TNS molecules and the proteins suggest that BSA forms more number of hydrogen bonding interactions with TNS even compared to the unfolded states of RNase A and α-LA (Fig. 6A). Thus, the higher order aggregates in water might be destabilized in the presence of BSA which is in fact observed as a reduction of H-bonds of TNS with solvents (Fig. 6B). Moreover, TNS molecules occupy higher fraction of hydrophobic surfaces on BSA than hydrophilic surfaces. Though the denaturation of α-LA, would increase the hydrophobic surface available for TNS occupancy, the percentage of occupancy is not varied (Fig. 10) and the distribution of different forms of aggregates also remain almost same (Fig. 7C). In case of RNase A, upon denaturation of the protein, the percentage occupancy of hydrophobic surfaces by TNS is reduced. This could be attributed to the increase in the number of lower order aggregates of TNS (Fig. 2B) which is presumably encouraged by interactions with the side chains of basic amino acids (Fig. 9) rather than hydrophobic pockets of the denatured protein. This could be due to the inherent basic nature of RNase A (pI 9.3).

The analysis of charges on the protein surfaces where TNS molecules are bound provides further insight on the charge-charge interactions. The electrostatic surface maps (Fig. 14) evidently show that monomeric or lower order aggregates of TNS preferably binds around hydrophobic surfaces or to the basic amino acid residues of BSA (Video file, V6). However, such monomeric (or lower order aggregates of) TNS molecules are not found on the surfaces of RNase A and α-LA even in their unfolded states. In fact, these two proteins encourage formation of higher order TNS aggregates on their surface which is lower in fraction on BSA. This predicts that the fraction of distribution of TNS in monomeric and in aggregated forms might vary from protein to protein based on their surface charge, availability of hydrophobic crevices, and the extent of unfolding. Thus, the amount of increase in fluorescence emission of TNS upon binding to denatured states of different proteins cannot be expected to be similar.

The further analysis on residue specificity of TNS interactions also suggest that in all the proteins TNS preferably binds to the side chains of basic or aromatic amino acid residues (Fig. 8-9). Earlier experimental studies with short peptides have also proposed that the binding of TNS with basic amino acids could be the initiating factor followed by hydrophobic interactions to position the aromatic ring systems in appropriate environment which could in fact enhance the fluorescence of the dye [278, 301]. Interestingly, for

ANS as well the ionic interactions between sulfonyl group and the side chains of basic amino acid residues is attributed to the stabilization of protein-dye complex and enhancement of the fluorescence [357]. Further, the most effective interactions of ANS are characterized to be with the arginine side chains followed by lysine and histidine [356]. Similar interactions have been identified for bis-ANS while interacting with *Bacillus subtilis* lipase using MD simulations [347, 357]. Moreover, the amino acids with polar side chains also contribute to the binding of TNS, however, it is clearly observed that the fraction of interactions with acidic side chains is very less. Persistent contacts of TNS with the backbone of the protein chains are also noticed in all the cases supporting the existence of weak non-specific interactions which could be probably found in all the proteins.

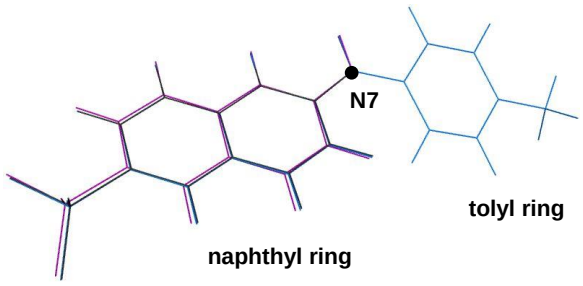


Figure 15. Representation of most accessible conformation of TNS in water (blue), DMF (red) and BSA (black). The central nitrogen atom N7 is labeled for reference.

6.5.3 Binding Modes of TNS

Since the fluorescence property of fluorophores could be altered by the conformational changes of the molecule [366], the conformational states accessed by TNS in different environments need to be examined. In order to accomplish this, four representative angles and two dihedrals of TNS were calculated in different solvents and in different proteins. There are notable changes observed in the mean values of all the angles (Table 2). In DMF and DMSO, the mean of the angles θ1, θ2 and θ4 are exceptionally increased/decreased suggesting that the monomeric form of TNS in these environments would be different from the aggregated forms (such as in water) and the monomers found on the surface of the protein (such as in BSA). Representative conformations of TNS in these three conditions are shown in Fig. 15. Further, the variations in the distribution of angle α1is found to be less (FWHM of < 8.5°) compared to the other angles. In general, the curve width followed the trend of $\alpha 2 \approx \alpha 4 > \alpha 3 >> \alpha 1$. The lower values of α1 suggest that the bending motion of the two-ring systems in TNS is more constrained during the binding whereas the extent of movements along the planes of the rings, as represented by other angles, could be relatively high. This is further supported by the direct analysis of two of the dihedrals. Each dihedrals θ1 and θ2 show two most-probable angles (Fig. 13) represented by bimodal distributions. Their variation (FWHM) is also comparable with the angles θ2, θ3 and θ4 (Table 2). This further confirms that the binding of TNS on proteins might be facilitated by planar movements than the restricted angular bending between the ring systems. Further examination on the correlations between the extent of changes among the four angles suggest that if the bond angle between the two aromatic rings (α1) in TNS is reduced, the angle between the planes of these rings (α2) would be increased and the planar angle of toluene ring with

respect to the central nitrogen atom (α4) would be decreased (Fig. 12E). These constrained-correlated changes could assist the favourable binding of TNS on the proteins.

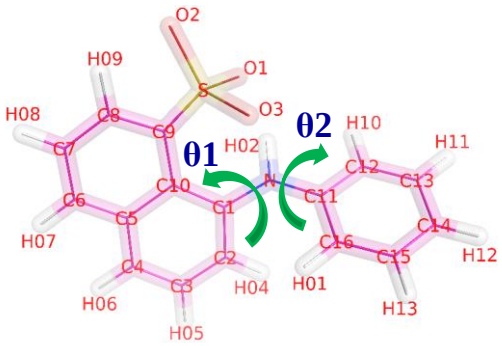


Figure 16. Stick model of ANS molecule with all the atoms labeled.

To the best of our knowledge, there is no crystal structure of TNS bound to a protein molecule is available. However, structures of proteins with ANS (Fig. 16) (similar dye as TNS) have been resolved. The conformations of protein bound ANS molecules were extracted from protein databank and aligned in PyMOL [155]. Also, the angle and dihedrals of ANS correspond to α1 and θ1 & θ2 of TNS, respectively, were calculated (Table 4). The results show that ANS does not undergo larger changes on the bending angle whereas limited dihedral movements are plausible upon binding. MD simulation of protein bound bis-ANS also supports the constrained changes in the conformations of the dye at different binding sites on lipase [347]. Hence, it may be assumed that these restricted dynamics could be a common nature of *N*-arylaminonaphthalene sulfonate dyes.

Table 4. The angle and dihedrals of protein bound ANS

PDB ids from which ANS conformations were obtained‡	Bond angle between C1-N-C11 (equivalent to α1)	Rotational angle of naphthyl ring (equivalent to θ1)	Rotational angle of phenyl ring (equivalent to θ2)
1txc-ans1	132.2	141.9	150.4
1txc-ans2	135.5	151.4	168.2
1txc-ans3	134.6	142.5	178.5
3pxq-ans1	129.2	142.2	143.7
3pxq-ans2	127.7	127.4	164.7
3pxq-ans3	127.7	169.1	124.4
4n3e-ans1	119.9	173.5	149.7
4n3e-ans2	125.2	143.3	142.8
4n3e-ans3	121.8	167.6	147.2
4n3e-ans4	121.1	157.6	145.3

‡ - ans1, ans 2, etc., represents that the same protein has multiple ANS molecules bound on it

6.5.4 Aggregation, Binding and Fluorescence

The overall analysis of the TNS configurations in different solvents and in different protein states suggest that TNS preferably aggregates in water whereas in non-aqueous solvents the number of higher order aggregates is largely reduced. In the presence of solvents with H-bonding acceptors, such as DMF and DMSO, TNS

preferably exists in its monomeric form. Experimental results suggest that in these conditions TNS fluorescence could be enhanced. Comparing these correlations with the TNS fluorescence changes in the presence of proteins predict that TNS prefers basic or aromatic amino acid residues for binding. The hydrophobic crevices along with positively charged surfaces on the protein could facilitate TNS molecules to bind in its monomeric form, thus enhance its fluorescence quantum yield. However, binding on the hydrophilic surfaces of the protein and exposed to water might encourage the aggregation of TNS molecules which would reduce the fluorescence emission. The protein-TNS structures obtained from the simulations suggest that there could be different microenvironments for TNS binding as proposed earlier [358]. An attempt to dissect at least three different TNS sub-structures as proposed from fluorescence life time measurements of different proteins [315, 358] provides the following: (a) TNS in monomeric form buried in the hydrophobic crevices of the proteins which could show high fluorescence quantum yield, (b) smaller-aggregates of TNS which are partially-exposed to the solvent, and (c) almost completely solvent-exposed higher-order aggregates which are similar to the TNS aggregates found in water in terms of solvent exposure, thus, may show similar lifetime and fluorescence quenching as suggested by the experiments.

6.6 Summary

Conformational state of TNS in different solvents and in proteins has been investigated using unconstrained MD simulations with multiple numbers of TNS molecules. The simulations provide hitherto unknown information on protein-TNS interactions. TNS preferably forms aggregates in aqueous environment which could be completely destabilized by non-aqueous solvents with proton-accepting groups. When protein surface could facilitate the binding of TNS in monomeric form through its hydrophobic pockets and with basic amino acid residues, the fluorescence emission of the dye might increase. However, in some proteins TNS forms aggregates of different proportions and their fluorescence emission is similar to the aqueous environment. If the unfolded conformations of such proteins could not create appropriate hydrophobic crevices to accommodate the aromatic ring systems and stabilize the monomeric state of TNS, the fluorescence emission would not be enhanced. Thus, fluorescence emission must be cautiously used as an evidence to interpret the unfolded conformations of proteins which might vary with different proteins based on their surface charge and ability to form hydrophobic pockets for the binding of TNS. Further, this study identifies the constrained-correlated motions of the aromatic rings in TNS upon binding to the proteins.

Bibliography

[1] Kettering, C. F., Shutts, L. W. & Andrews, D. H. (1930) A representation of the dynamic properties of molecules by mechanical models, *Phys. Rev.*, **36**, 531-543.

[2] Hill, T. L. (1946) On steric effects, *J. Chem. Phys.*, **14**, 465-465.

[3] Westheimer, F. H. & Mayer, J. E. (1946) The theory of the racemization of optically active derivatives of diphenyl, *J. Chem. Phys.*, **14**, 733-738.

[4] Dostrovsky, I., Hughes, E. D. & Ingold, C. K. (1946) 50. Mechanism of substitution at a saturated carbon atom. Part XXXII. The role of steric hindrance. (Section G) magnitude of steric effects, range of occurrence of steric and polar effects, and place of the wagner rearrangement in nucleophilic substitution and elimination, *J. Chem. Soc.*, 173-194.

[5] Metropolis, N. & Ulam, S. (1949) The monte carlo method, *J. Amer. Statist. Assoc.*, **44**, 335-341.

[6] Kirkwood, J. G., Maun, E. K. & Alder, B. J. (1950) Radial Distribution Functions and the Equation of State of a Fluid Composed of Rigid Spherical Molecules, *J. Chem. Phys.*, **18**, 1040-1047.

[7] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., *et al.* (1953) Equation of state calculations by fast computing machines, *J. Chem. Phys.*, **21**, 1087-1092.

[8] Rosenbluth, M. N. & Rosenbluth, A. W. (1954) Further results on Monte Carlo equations of state, *J. Chem. Phys.*, **22**, 881-884.

[9] Fermi, E., Pasta, J. & Ulam, S. (1955) Studies of nonlinear problems, *Los Alamos Report LA-1940*, **978**.

[10] Leach, A. R. (2001), Molecular modelling: principles and applications, *Pearson education*.

[11] Alder, B. & Wainwright, T. (1957) Phase transition for a hard sphere system, *J. Chem. Phys.*, **27**, 1208.

[12] Alder, B. J. & Wainwright, T. (1959) Studies in molecular dynamics. I. General method, *J. Chem. Phys.*, **31**, 459-466.

[13] Kitaygorodsky, A. (1960) Calculation of conformations of organic molecules, *Tetrahedron*, **9**, 183-193.

[14] Kitaygorodsky, A. I. (1961) The interaction curve of non-bonded carbon and hydrogen atoms and its application, *Tetrahedron*, **14**, 230-236.

[15] Kendrew, J. C., Bodo, G., Dintzis, H. M., *et al.* (1958) A three-dimensional model of the myoglobin molecule obtained by X-ray analysis, *Nature*, **181**, 662-666.

[16] Hendrickson, J. B. (1961) Molecular geometry. I. machine computation of the common rings, *J. Am. Chem. Soc.*, **83**, 4537-4547.

[17] Rahman, A. (1964) Correlations in the motion of atoms in liquid argon, *Phys. Rev.*, **136**, A405.

[18] Rahman, A. & Stillinger, F. H. (1971) Molecular dynamics study of liquid water, *J. Chem. Phys.*, **55**, 3336-3359.

[19] Allinger, N. L., Tribble, M. T., Miller, M. A. & Wertz, D. H. (1971) Conformational analysis. LXIX. Improved force field for the calculation of the structures and energies of hydrocarbons, *J. Am. Chem. Soc.*, **93**, 1637-1648.

[20] Allinger, N. L. & Sprague, J. T. (1972) Conformational analysis. LXXXIV. Study of the structures and energies of some alkenes and cycloalkenes by the force field method, *J. Am. Chem. Soc.*, **94**, 5734-5747.

[21] Allinger, N., Tribble, M. & Miller, M. (1972) Conformational analysis—LXXIX: An improved force field for the calculation of the structures and energies of carbonyl compounds, *Tetrahedron*, **28**, 1173-1190.

[22] Allinger, N. L. & Sprague, J. T. (1973) Conformational analysis. XC. Calculation of the structures of hydrocarbons containing delocalized electronic systems by the molecular mechanics method, *J. Am. Chem. Soc.*, **95**, 3893-3907.

[23] Allinger, N. L. (1977) Conformational analysis. 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms, *J. Am. Chem. Soc.*, **99**, 8127-8134.

[24] Allinger, N. L., Yuh, Y. H. & Lii, J. H. (1989) Molecular mechanics. The MM3 force field for hydrocarbons. 1, *J. Am. Chem. Soc.*, **111**, 8551-8566.

[25] Allinger, N. L., Chen, K. & Lii, J. H. (1996) An improved force field (MM4) for saturated hydrocarbons, *J. Comput. Chem.*, **17**, 642-668.

[26] Boyd, D. B. (2013) Quantum chemistry program exchange, facilitator of theoretical and computational chemistry in pre-Internet history in *Pioneers of Quantum Chemistry* pp. 221-273, American Chemical Society.
[27] Engler, E. M., Andose, J. D. & Schleyer, P. V. R. (1973) Critical evaluation of molecular mechanics, *J. Am. Chem. Soc.*, **95**, 8005-8025.
[28] Brownlee, C., The protein papers, http://www.pnas.org/site/classics/classics1.xhtml.
[29] Pauling, L. & Corey, R. B. (1950) Two hydrogen-bonded spiral configurations of the polypeptide chain, *J. Am. Chem. Soc.*, **72**, 5349-5349.
[30] Pauling, L., Corey, R. B. & Branson, H. R. (1951) The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain, *Proc. Natl. Acad. Sci. USA*, **37**, 205-211.
[31] Kremer, J. M., Wilting, J. & Janssen, L. (1988) Drug binding to human alpha-1-acid glycoprotein in health and disease, *Pharmacol. Rev.*, **40**, 1-47.
[32] Frokjaer, S. & Otzen, D. E. (2005) Protein drug stability: a formulation challenge, *Nat. Rev. Drug Discov.*, **4**, 298-306.
[33] Li, J., Zhu, X. & Chen, J. Y. (2009) Building disease-specific drug-protein connectivity maps from molecular interaction networks and PubMed abstracts, *PLoS. Comput. Biol.*, **5**, e1000450.
[34] Dudley, J. T., Deshpande, T. & Butte, A. J. (2011) Exploiting drug–disease relationships for computational drug repositioning, *Brief. Bioinform.*, bbr013.
[35] Refolo, L. M., Pappolla, M. A., LaFrancois, J., *et al.* (2001) A cholesterol-lowering drug reduces β-amyloid pathology in a transgenic mouse model of Alzheimer's disease, *Neurobiol. Dis.*, **8**, 890-899.
[36] Levitt, M. & Warshel, A. (1975) Computer simulation of protein folding, *Nature*, **253**, 694-698.
[37] Momany, F. A., McGuire, R. F., Burgess, A. W. & Scheraga, H. A. (1975) Energy parameters in polypeptides. VII. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions, and intrinsic torsional potentials for the naturally occurring amino acids, *J. Phys. Chem.*, **79**, 2361-2381.
[38] McCammon, J. A., Gelin, B. R. & Karplus, M. (1977) Dynamics of folded proteins, *Nature*, **267**, 585-590.
[39] Artymiuk, P. J., Blake, C. C. F., Grace, D. E. P., *et al.* (1979) Crystallographic studies of the dynamic properties of lysozyme, *Nature*, **280**, 563-568.
[40] Frauenfelder, H., Petsko, G. A. & Tsernoglou, D. (1979) Temperature-dependent X-ray diffraction as a probe of protein structural dynamics, *Nature*, **280**, 558-563.
[41] Lifson, S., Hagler, A. T. & Dauber, P. (1979) Consistent force field studies of intermolecular forces in hydrogen-bonded crystals. 1. Carboxylic acids, amides, and the C:O.cntdot..cntdot..cntdot.H- hydrogen bonds, *J. Am. Chem. Soc.*, **101**, 5111-5121.
[42] Lifson, S. & Warshel, A. (1968) Consistent force field for calculations of conformations, vibrational spectra, and enthalpies of cycloalkane and n-alkane molecules, *J. Chem. Phys.*, **49**, 5116-5129.
[43] Warshel, A. & Lifson, S. (1970) Consistent force field calculations. II. crystal structures, sublimation energies, molecular and lattice vibrations, molecular conformations, and enthalpies of alkanes, *J. Chem. Phys.*, **53**, 582-594.
[44] Ermer, O. & Lifson, S. (1973) Consistent force field calculations. III. Vibrations, conformations, and heats of hydrogenation of nonconjugated olefins, *J. Am. Chem. Soc.*, **95**, 4121-4132.
[45] Hagler, A. T., Huler, E. & Lifson, S. (1974) Energy functions for peptides and proteins. I. Derivation of a consistent force field including the hydrogen bond from amide crystals, *J. Am. Chem. Soc.*, **96**, 5319-5327.
[46] Hagler, A. T. & Lifson, S. (1974) Energy functions for peptides and proteins. II. Amide hydrogen bond and calculation of amide crystal properties, *J. Am. Chem. Soc.*, **96**, 5327-5335.
[47] Weiner, P. K. & Kollman, P. A. (1981) AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions, *J. Comput. Chem.*, **2**, 287-303.
[48] Weiner, S. J., Kollman, P. A., Case, D. A., *et al.* (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins, *J. Am. Chem. Soc.*, **106**, 765-784.
[49] Van Gunsteren, W. & Berendsen, H. (1987) Groningen molecular simulation (GROMOS) library manual, *Biomos, Groningen*, **24**, 13.
[50] Scott, W. R., Hünenberger, P. H., Tironi, I. G., *et al.* (1999) The GROMOS biomolecular simulation program package, *J. Phys. Chem. A*, **103**, 3596-3607.
[51] Jorgensen, W. L. & Tirado-Rives, J. (1988) The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin, *J. Am. Chem. Soc.*, **110**, 1657-1666.

[52] Berendsen, H. J., van der Spoel, D. & van Drunen, R. (1995) GROMACS: a message-passing parallel molecular dynamics implementation, *Comput. Phys. Comm.*, **91**, 43-56.
[53] Warshel, A. & Levitt, M. (1976) Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme, *J. Mol. Biol.*, **103**, 227-249.
[54] The Nobel prize in chemistry 2013, https://www.nobelprize.org/nobel_prizes/chemistry/laureates/2013/.
[55] Rudd, R. E. & Broughton, J. Q. (1998) Coarse-grained molecular dynamics and the atomic limit of finite elements, *Phys. Rev. B*, **58**, R5893-R5896.
[56] Marrink, S. J., Risselada, H. J., Yefimov, S., *et al.* (2007) The MARTINI force field: coarse grained model for biomolecular simulations, *J. Phys. Chem. B*, **111**, 7812-7824.
[57] de Jong, D. H., Singh, G., Bennett, W. D., *et al.* (2012) Improved parameters for the martini coarse-grained protein force field, *J. Chem. Theory Comput.*, **9**, 687-697.
[58] Marrink, S. J. & Tieleman, D. P. (2013) Perspective on the Martini model, *Chem. Soc. Rev.*, **42**, 6801-6822.
[59] Sugita, Y. & Okamoto, Y. (1999) Replica-exchange molecular dynamics method for protein folding, *Chem. Phys. Lett.*, **314**, 141-151.
[60] Karplus, M. & Kuriyan, J. (2005) Molecular dynamics and protein function, *Proc. Natl. Acad. Sci. USA*, **102**, 6679-6685.
[61] Vaccaro, L., Koronakis, V. & Sansom, M. S. (2006) Flexibility in a drug transport accessory protein: molecular dynamics simulations of MexA, *Biophys. J.*, **91**, 558-564.
[62] Klepeis, J. L., Lindorff-Larsen, K., Dror, R. O. & Shaw, D. E. (2009) Long-timescale molecular dynamics simulations of protein structure and function, *Curr. Opin. Struct. Biol.*, **19**, 120-127.
[63] Ge, H., Wang, Y., Li, C., *et al.* (2013) Molecular dynamics-based virtual screening: accelerating the drug discovery process by high-performance computing, *J. Chem. Inf. Model.*, **53**, 2757-2764.
[64] MacKerell, A. D. & Nilsson, L. (2008) Molecular dynamics simulations of nucleic acid–protein complexes, *Curr. Opin. Struct. Biol.*, **18**, 194-199.
[65] Lebrun, A. & Lavery, R. (1996) Modelling Extreme Stretching of DNA, *Nucleic Acids Res.*, **24**, 2260-2267.
[66] Kosikov, K. M., Gorin, A. A., Zhurkin, V. B. & Olson, W. K. (1999) DNA stretching and compression: large-scale simulations of double helical structures1, *J. Mol. Biol.*, **289**, 1301-1326.
[67] Lebrun, A. & Lavery, R. (1998) Modeling the mechanics of a DNA oligomer, *J. Biomol. Struct. Dyn.*, **16**, 593-604.
[68] Strahan, G. D., Keniry, M. A. & Shafer, R. H. (1998) NMR Structure refinement and dynamics of the K+-[d(G3T4G3)]2 quadruplex via particle mesh Ewald molecular dynamics simulations, *Biophys. J.*, **75**, 968-981.
[69] Špačková, N. a., Berger, I. & Šponer, J. (1999) Nanosecond molecular dynamics simulations of parallel and antiparallel guanine quadruplex DNA molecules, *J. Am. Chem. Soc.*, **121**, 5519-5534.
[70] Fersht, A. R. (1995) Characterizing transition states in protein folding: an essential step in the puzzle, *Curr. Opin. Struct. Biol.*, **5**, 79-84.
[71] Du, R., Pande, V. S., Grosberg, A. Y., *et al.* (1998) On the transition coordinate for protein folding, *J. Chem. Phys.*, **108**, 334-350.
[72] Marchi, M. & Ballone, P. (1999) Adiabatic bias molecular dynamics: A method to navigate the conformational space of complex molecular systems, *J. Chem. Phys.*, **110**, 3697-3702.
[73] Geissler, P. L., Dellago, C. & Chandler, D. (1999) Kinetic pathways of ion pair dissociation in water, *J. Phys. Chem. B*, **103**, 3706-3710.
[74] Tiana, G. & Camilloni, C. (2012) Ratcheted molecular-dynamics simulations identify efficiently the transition state of protein folding, *J. Chem. Phys.*, **137**, 235101.
[75] Bolhuis, P. G., Chandler, D., Dellago, C. & Geissler, P. L. (2002) Transition path sampling: throwing ropes over rough mountain passes, in the dark, *Annu. Rev. Phys. Chem.*, **53**, 291-318.
[76] Faradjian, A. K. & Elber, R. (2004) Computing time scales from reaction coordinates by milestoning, *J. Chem. Phys.*, **120**, 10880-10889.
[77] Faccioli, P., Sega, M., Pederiva, F. & Orland, H. (2006) Dominant pathways in protein folding, *Phys. Rev. Lett.*, **97**, 108101.
[78] Hu, J., Ma, A. & Dinner, A. R. (2006) Bias annealing: A method for obtaining transition paths de novo, *J. Chem. Phys.*, **125**, 114101.
[79] Faccioli, P. (2008) Characterization of protein folding by dominant reaction pathways, *J. Phys. Chem. B*, **112**, 13756-13764.

[80] Sha, Z., Wu, R., Lu, Y., *et al.* (2009) Glass forming abilities of binary Cu100-xZrx (34, 35.5, and 38.2 at.%) metallic glasses: A LAMMPS study, *J. Appl. Phys.*, **105**, 3521.

[81] Hu, J., Ruan, X. & Chen, Y. P. (2009) Thermal conductivity and thermal rectification in graphene nanoribbons: A molecular dynamics study, *Nano Letters*, **9**, 2730-2735.

[82] Shih, C.-J., Lin, S., Sharma, R., *et al.* (2012) Understanding the pH-dependent behavior of graphene oxide aqueous solutions: A Comparative experimental and molecular dynamics simulation study, *Langmuir*, **28**, 235-241.

[83] Berber, S., Kwon, Y.-K. & Tománek, D. (2000) Unusually high thermal conductivity of carbon nanotubes, *Phys. Rev. Lett.*, **84**, 4613-4616.

[84] Smith, J. S., Bedrov, D. & Smith, G. D. (2003) A molecular dynamics simulation study of nanoparticle interactions in a model polymer-nanoparticle composite, *Composites Sci. Technol.*, **63**, 1599-1605.

[85] Yamakov, V., Wolf, D., Phillpot, S. R., *et al.* (2002) Dislocation processes in the deformation of nanocrystalline aluminium by molecular-dynamics simulation, *Nat. Mater*, **1**, 45-49.

[86] Yamakov, V., Wolf, D., Phillpot, S. R., *et al.* (2004) Deformation-mechanism map for nanocrystalline metals by molecular-dynamics simulation, *Nat. Mater*, **3**, 43-47.

[87] Zhang, Y., Liu, H. & Yang, W. (2000) Free energy calculation on enzyme reactions with an efficient iterative procedure to determine minimum energy paths on a combined ab initio QM/MM potential energy surface, *J. Chem. Phys.*, **112**, 3483-3492.

[88] Cui, Q., Elstner, M., Kaxiras, E., *et al.* (2001) A QM/MM implementation of the self-consistent charge density functional tight binding (SCC-DFTB) method, *J. Phys. Chem. B*, **105**, 569-585.

[89] French, S. A., Sokol, A. A., Bromley, S. T., *et al.* (2001) From CO2 to methanol by hybrid QM/MM embedding, *Angew. Chem. Int. Ed.*, **113**, 4569-4572.

[90] Acevedo, O. & Jorgensen, W. L. (2009) Advances in quantum and molecular mechanical (QM/MM) simulations for organic and enzymatic reactions, *Acc. Chem. Res.*, **43**, 142-151.

[91] Chandrasekhar, J., Shariffskul, S. & Jorgensen, W. L. (2002) QM/MM simulations for Diels-Alder reactions in water: Contribution of enhanced hydrogen bonding at the transition state to the solvent effect, *J. Phys. Chem. B*, **106**, 8078-8085.

[92] Reyes, C. M. & Kollman, P. A. (2000) Structure and thermodynamics of RNA-protein binding: using molecular dynamics and free energy analyses to calculate the free energies of binding and conformational change, *J. Mol. Biol.*, **297**, 1145-1158.

[93] Freddolino, P. L., Arkhipov, A. S., Larson, S. B., *et al.* (2006) Molecular dynamics simulations of the complete satellite tobacco mosaic virus, *Structure*, **14**, 437-449.

[94] Jayaram, B., McConnell, K., Dixit, S. B. & Beveridge, D. (1999) Free energy analysis of protein–DNA binding: the EcoRI endonuclease–DNA complex, *J. Comput. Phys.*, **151**, 333-357.

[95] Brooks, C., Brünger, A. & Karplus, M. (1985) Active site dynamics in protein molecules: A stochastic boundary molecular-dynamics approach, *Biopolymers*, **24**, 843-865.

[96] Abriata, L. A. & Dal Peraro, M. (2015) Assessing the potential of atomistic molecular dynamics simulations to probe reversible protein-protein recognition and binding, *Sci. Rep.*, **5**.

[97] Cuendet, M. A. & Michielin, O. (2008) Protein-protein interaction investigated by steered molecular dynamics: the TCR-pMHC complex, *Biophys. J.*, **95**, 3575-3590.

[98] Cau, Y., Fiorillo, A., Mori, M., *et al.* (2015) Molecular dynamics simulations and structural analysis of giardia duodenalis 14-3-3 protein–protein interactions, *J. Chem. Inf. Model.*, **55**, 2611-2622.

[99] Lindahl, E. & Sansom, M. S. (2008) Membrane proteins: molecular dynamics simulations, *Curr. Opin. Struct. Biol.*, **18**, 425-431.

[100] Jo, S., Kim, T. & Im, W. (2007) Automated builder and database of protein/membrane complexes for molecular dynamics simulations, *PloS ONE*, **2**, e880.

[101] Pickholz, M., Oliveira, O. N. & Skaf, M. S. (2006) Molecular dynamics simulations of neutral chlorpromazine in zwitterionic phospholipid monolayers, *J. Phys. Chem. B*, **110**, 8804-8814.

[102] Ulander, J. & Haymet, A. (2003) Permeation across hydrated DPPC lipid bilayers: simulation of the titrable amphiphilic drug valproic acid, *Biophys. J.*, **85**, 3475-3484.

[103] Wood, I. & Pickholz, M. (2013) Concentration effects of sumatriptan on the properties of model membranes by molecular dynamics simulations, *Eur. Biophys. J.*, **42**, 833-841.

[104] Tarek, M. (2005) Membrane electroporation: a molecular dynamics simulation, *Biophys. J.*, **88**, 4045-4053.

[105] Venable, R. M., Zhang, Y., Hardy, B. J. & Pastor, R. W. (1993) Molecular dynamics simulations of a lipid bilayer and of hexadecane: an investigation of membrane fluidity, *Science*, **262**, 223-223.

[106] Van der Ploeg, P. & Berendsen, H. (1982) Molecular dynamics simulation of a bilayer membrane, *J. Chem. Phys.*, **76**, 3271-3276.
[107] Silva, J. R., Pan, H., Wu, D., *et al.* (2009) A multiscale model linking ion-channel molecular dynamics and electrostatics to the cardiac action potential, *Proc. Natl. Acad. Sci. USA*, **106**, 11102-11106.
[108] Jensen, M. Ø., Jogini, V., Borhani, D. W., *et al.* (2012) Mechanism of voltage gating in potassium channels, *Science*, **336**, 229-233.
[109] Ma, B. & Nussinov, R. (2002) Molecular dynamics simulations of alanine rich β-sheet oligomers: Insight into amyloid formation, *Protein Sci.*, **11**, 2335-2350.
[110] Gsponer, J., Haberthür, U. & Caflisch, A. (2003) The role of side-chain interactions in the early steps of aggregation: Molecular dynamics simulations of an amyloid-forming peptide from the yeast prion Sup35, *Proc. Natl. Acad. Sci. USA*, **100**, 5154-5159.
[111] Suplatov, D. A., Besenmatter, W., Svedas, V. K. & Svendsen, A. (2012) Bioinformatic analysis of alpha/beta-hydrolase fold enzymes reveals subfamily-specific positions responsible for discrimination of amidase and lipase activities, *Protein Eng Des Sel*, **25**, 689-97.
[112] Skjot, M., De Maria, L., Chatterjee, R., *et al.* (2009) Understanding the plasticity of the alpha/beta hydrolase fold: lid swapping on the Candida antarctica lipase B results in chimeras with interesting biocatalytic properties, *Chembiochem*, **10**, 520-7.
[113] van Pouderoyen, G., Eggert, T., Jaeger, K. E. & Dijkstra, B. W. (2001) The crystal structure of Bacillus subtilis lipase: a minimal alpha/beta hydrolase fold enzyme, *J Mol Biol*, **309**, 215-26.
[114] Lewkowitsch, J. & Warburton, G. H. (1895), Chemical technology and analysis of oils, fats and waxes, *Macmillan and co., limited*.
[115] Sarda, L. & Desnuelle, P. (1958) Actions of pancreatic lipase on esters in emulsions, *Biochim Biophys Acta*, **30**, 513-21.
[116] Entressangles, B. & Desnuelle, P. (1968) Action of pancreatic lipase on aggregated glyceride molecules in an isotropic system, *Biochim Biophys Acta*, **159**, 285-95.
[117] Chapus, C., Semeriva, M., Bovier-Lapierre, C. & Desnuelle, P. (1976) Mechanism of pancreatic lipase action. 1. Interfacial activation of pancreatic lipase, *Biochemistry*, **15**, 4980-7.
[118] Malakhova, E. A., Kurganov, B. I., Levashov, A. V., *et al.* (1983) New approach to the study of enzyme reactions with the participation of water-insoluble substrates. Pancreatic lipase incorporated into the inverted micelles of a surface-active substance in an organic solvent, *Dokl Akad Nauk SSSR*, **270**, 474-7.
[119] Inada, Y., Nishimura, H., Takahashi, K., *et al.* (1984) Ester synthesis catalyzed by polyethylene glycol-modified lipase in benzene, *Biochem Biophys Res Commun*, **122**, 845-50.
[120] Kim, K. H., Kwon, D. Y. & Rhee, J. S. (1984) Effects of organic solvents on lipase for fat splitting, *Lipids*, **19**, 975-7.
[121] Zaks, A. & Klibanov, A. M. (1984) Enzymatic Catalysis in Organic Media at 100° C, *Science*, **224**, 1249-1251.
[122] Zaks, A. & Klibanov, A. M. (1985) Enzyme-catalyzed processes in organic solvents, *Proc. Natl. Acad. Sci. USA*, **82**, 3192-3196.
[123] Zaks, A. & Klibanov, A. M. (1988) Enzymatic catalysis in nonaqueous solvents, *J. Biol. Chem.*, **263**, 3194-201.
[124] Houde, A., Kademi, A. & Leblanc, D. (2004) Lipases and their industrial applications: an overview, *Appl. Biochem. Biotechnol.*, **118**, 155-70.
[125] Shah, S., Sharma, S. & Gupta, M. N. (2004) Biodiesel Preparation by Lipase-Catalyzed Transesterification of Jatropha Oil, *Energy & Fuels*, **18**, 154-159.
[126] Semeriva, M. & Desnuelle, P. (1976) Pancreatic lipase and colipase: an example of heterogeneous biocatalysis, *Horiz Biochem Biophys*, **2**, 32-59.
[127] Rosenheim, O. & Shaw-Mackenzie, J. A. (1910) PROCEEDINGS OF THE PHYSIOLOGICAL SOCIETY: February 19, 1910, *J. Physiol*, **40**, viii-xxvi.
[128] Maylie, M. F., Charles, M., Gache, C. & Desnuelle, P. (1971) Isolation and partial identification of a pancreatic colipase, *Biochim Biophys Acta*, **229**, 286-9.
[129] Morgan, R. G. & Hoffman, N. E. (1971) The interaction of lipase, lipase cofactor and bile salts in triglyceride hydrolysis, *Biochim. Biophys. Acta.*, **248**, 143-8.
[130] Erlanson, C. & Borgström, B. (1972) Purification and further characterization of co-lipase from porcine pancreas, *Biochim. Biophys. Acta*, **271**, 400-412.
[131] Rathelot, J., Julien, R., Bosc-Bierne, I., *et al.* (1981) Horse pancreatic lipase. Interaction with colipase from various species, *Biochimie*, **63**, 227-34.

[132] Hermoso, J., Pignol, D., Kerfelec, B., *et al.* (1996) Lipase activation by nonionic detergents. The crystal structure of the porcine lipase-colipase-tetraethylene glycol monooctyl ether complex, *J. Biol. Chem.*, **271**, 18007-16.

[133] Hjorth, A., Carriere, F., Cudrey, C., *et al.* (1993) A structural domain (the lid) found in pancreatic lipases is absent in the guinea pig (phospho)lipase, *Biochemistry*, **32**, 4702-7.

[134] Holmquist, M., Clausen, I. G., Patkar, S., *et al.* (1995) Probing a functional role of Glu87 and Trp89 in the lid of Humicola lanuginosa lipase through transesterification reactions in organic solvent, *J Protein Chem*, **14**, 217-24.

[135] Chang, R. C., Chen, J. C. & Shaw, J. F. (1996) Facile purification of highly active recombinant Staphylococcus hyicus lipase fragment and characterization of a putative lid region, *Biochem Biophys Res Commun*, **228**, 774-9.

[136] Peters, G. H., Olsen, O. H., Svendsen, A. & Wade, R. C. (1996) Theoretical investigation of the dynamics of the active site lid in Rhizomucor miehei lipase, *Biophysical Journal*, **71**, 119-29.

[137] Lowe, M. E. (1997) Colipase stabilizes the lid domain of pancreatic triglyceride lipase, *J Biol Chem*, **272**, 9-12.

[138] Belle, V., Fournel, A., Woudstra, M., *et al.* (2007) Probing the opening of the pancreatic lipase lid using site-directed spin labeling and EPR spectroscopy, *Biochemistry*, **46**, 2205-14.

[139] Ranaldi, S., Belle, V., Woudstra, M., *et al.* (2010) Amplitude of pancreatic lipase lid opening in solution and identification of spin label conformational subensembles by combining continuous wave and pulsed EPR spectroscopy and molecular dynamics, *Biochemistry*, **49**, 2140-9.

[140] Selvan, A., Seniya, C., Chandrasekaran, S. N., *et al.* (2010) Molecular dynamics simulations of human and dog gastric lipases: Insights into domain movements, *FEBS Letters*, **584**, 4599-4605.

[141] Papaleo, E. & Invernizzi, G. (2011) Conformational plasticity of the calcium-binding pocket in the Burkholderia glumae lipase: remodeling induced by mutation of calcium coordinating residues, *Biopolymers*, **95**, 117-26.

[142] Burney, P. R. & Pfaendtner, J. (2013) Structural and dynamic features of Candida rugosa lipase 1 in water, octane, toluene, and ionic liquids BMIM-PF6 and BMIM-NO3, *J Phys Chem B*, **117**, 2662-70.

[143] Ottosson, J., Fransson, L. & Hult, K. (2002) Substrate entropy in enzyme enantioselectivity: an experimental and molecular modeling study of a lipase, *Protein Sci*, **11**, 1462-71.

[144] Gruber, C. C. & Pleiss, J. (2012) Molecular modeling of lipase binding to a substrate-water interface, *Methods Mol Biol*, **861**, 313-27.

[145] Sali, A. & Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints, *J Mol Biol*, **234**, 779-815.

[146] Roussel, A., Yang, Y., Ferrato, F., *et al.* (1998) Structure and activity of rat pancreatic lipase-related protein 2, *J Biol Chem*, **273**, 32121-8.

[147] Eisenberg, D., Luthy, R. & Bowie, J. U. (1997) VERIFY3D: assessment of protein models with three-dimensional profiles, *Methods Enzymol*, **277**, 396-404.

[148] Laskowski, R. A., Macarthur, M. W., Moss, D. S. & Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Cryst.*, **26**, 283-291.

[149] Lindahl, E., Hess, B. & van der Spoel, D. (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis, *J. Mol. Model.*, **7**, 306-317.

[150] Van Der Spoel, D., Lindahl, E., Hess, B., *et al.* (2005) GROMACS: Fast, flexible, and free, *J. Comput. Chem.*, **26**, 1701-1718.

[151] Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. (2008) GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation, *J. Chem. Theory Comput.*, **4**, 435-447.

[152] Essmann, U., Perera, L., Berkowitz, M. L., *et al.* (1995) A smooth particle mesh Ewald method, *J. Chem. Phys.*, **103**, 8577-8593.

[153] Parrinello, M. & Rahman, A. (1981) Polymorphic transitions in single crystals: A new molecular dynamics method, *J. Appl. Phys.*, **52**, 7182-7190.

[154] Hess, B. (2008) P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation, *J. Chem. Theory Comput.*, **4**, 116-122.

[155] Schrodinger, LLC (2010) The PyMOL Molecular Graphics System, Version 1.3r1 in

[156] Pettersen, E. F., Goddard, T. D., Huang, C. C., *et al.* (2004) UCSF Chimera-a visualization system for exploratory research and analysis, *J Comput Chem*, **25**, 1605-12.

[157] Laskowski, R. A. (2001) PDBsum: summaries and analyses of PDB structures, *Nucleic Acids Research*, **29**, 221-222.

[158] Dolinsky, T. J., Czodrowski, P., Li, H., *et al.* (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations, *Nucleic Acids Research*, **35**, W522-W525.
[159] Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. (2004) PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations, *Nucleic. Acids. Res*, **32**, W665-W667.
[160] Baker, N. A., Sept, D., Joseph, S., *et al.* (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome, *Proc. Natl. Acad. Sci. USA*, **98**, 10037-10041.
[161] van Tilbeurgh, H., Egloff, M. P., Martinez, C., *et al.* (1993) Interfacial activation of the lipase-procolipase complex by mixed micelles revealed by X-ray crystallography, *Nature*, **362**, 814-20.
[162] Bourne, Y., Martinez, C., Kerfelec, B., *et al.* (1994) Horse pancreatic lipase. The crystal structure refined at 2.3 A resolution, *J. Mol. Biol*, **238**, 709-32.
[163] Ganoth, A., Friedman, R., Nachliel, E. & Gutman, M. (2006) A Molecular Dynamics Study and Free Energy Analysis of Complexes between the Mlc1p Protein and Two IQ Motif Peptides, *Biophysical Journal*, **91**, 2436-2450.
[164] Sharp, K. A. & Honig, B. (1990) Electrostatic Interactions in Macromolecules: Theory and Applications, *Annu. Rev. Biophys. Biophys. Chem*, **19**, 301-332.
[165] Chong, L. T., Duan, Y., Wang, L., *et al.* (1999) Molecular dynamics and free-energy calculations applied to affinity maturation in antibody 48G7, *Proc. Natl. Acad. Sci. USA*, **96**, 14330-14335.
[166] Egloff, M.-P., Marguet, F., Buono, G., *et al.* (1995) The 2.46 Å Resolution Structure of the Pancreatic Lipase-Colipase Complex Inhibited by a C11 Alkyl Phosphonate, *Biochemistry*, **34**, 2751-2762.
[167] Eydoux, C., Spinelli, S., Davis, T. L., *et al.* (2008) Structure of human pancreatic lipase-related protein 2 with the lid in an open conformation, *Biochemistry*, **47**, 9553-64.
[168] Ranaldi, S., Belle, V., Woudstra, M., *et al.* (2009) Lid opening and unfolding in human pancreatic lipase at low pH revealed by site-directed spin labeling EPR and FTIR spectroscopy, *Biochemistry*, **48**, 630-8.
[169] Bezzine, S., Ferrato, F., Ivanova, M. G., *et al.* (1999) Human pancreatic lipase: colipase dependence and interfacial binding of lid domain mutants, *Biochemistry*, **38**, 5499-510.
[170] Withers-Martinez, C., Carriere, F., Verger, R., *et al.* (1996) A pancreatic lipase with a phospholipase A1 activity: crystal structure of a chimeric pancreatic lipase-related protein 2 from guinea pig, *Structure*, **4**, 1363-74.
[171] Jayne, S., Kerfelec, B., Foglizzo, E., *et al.* (2002) Activation of horse PLRP2 by bile salts does not require colipase, *Biochemistry*, **41**, 8422-8.
[172] Johnson, Q. R., Nellas, R. B. & Shen, T. (2012) Solvent-dependent gating motions of an extremophilic lipase from Pseudomonas aeruginosa, *Biochemistry*, **51**, 6238-45.
[173] Cherukuvada, S. L., Seshasayee, A. S., Raghunathan, K., *et al.* (2005) Evidence of a double-lid movement in Pseudomonas aeruginosa lipase: insights from molecular dynamics simulations, *PLoS. Comput. Biol*, **1**, e28.
[174] Barbe, S., Cortes, J., Simeon, T., *et al.* (2011) A mixed molecular modeling-robotics approach to investigate lipase large molecular motions, *Proteins*, **79**, 2517-29.
[175] Trodler, P., Schmid, R. D. & Pleiss, J. (2009) Modeling of solvent-dependent conformational transitions in Burkholderia cepacia lipase, *BMC Struct. Biol.*, **9**, 38.
[176] Rehm, S., Trodler, P. & Pleiss, J. (2010) Solvent-induced lid opening in lipases: a molecular dynamics study, *Protein Sci.*, **19**, 2122-30.
[177] Green, D. E., Murer, E., Hultin, H. O., *et al.* (1965) Association of integrated metabolic pathways with membranes. I. Glycolytic enzymes of the red blood corpuscle and yeast, *Arch. Biochem. Biophys.*, **112**, 635-47.
[178] Welch, G. R. (1978) On the role of organized multienzyme systems in cellular metabolism: A general synthesis, *Prog. Biophys. Mol. Biol.*, **32**, 103-191.
[179] Fletcher, P. D. I., Freedman, R. B., Oldfield, C., *et al.* (1984) Reactivity of alpha-chymotrypsin in water-in-oil microemulsions, *Colloids Surf.*, **10**, 193-203.
[180] Martinek, K., Levashov, A. V., Klyachko, N. L., *et al.* (1981) The principles of enzyme stabilization. VI. Catalysis by water-soluble enzymes entrapped into reversed micelles of surfactants in organic solvents, *Biochim. Biophys. Acta*, **657**, 277-94.
[181] Misiorowski, R. L. & Wells, M. A. (1974) The activity of phospholipase A2 in reversed micelles of phosphatidylcholine in diethyl ether: effect of water and cations, *Biochemistry*, **13**, 4921-7.
[182] Darszon, A., Strasser, R. J. & Montal, M. (1979) Rhodopsin-phospholipid complexes in apolar environments: photochemical characterization, *Biochemistry*, **18**, 5205-5213.

[183] Douzou, P., Debey, P. & Franks, F. (1978) Supercooled water as medium for enzyme reactions at subzero temperatures, *Biochim. Biophys. Acta.*, **523**, 1-8.
[184] Menger, F. M. & Yamada, K. (1979) Enzyme catalysis in water pools, *J. Am. Chem. Soc.*, **101**, 6731-6734.
[185] Samama, J. P., Lee, K. M. & Biellmann, J. F. (1987) Enzymes and microemulsions. Activity and kinetic properties of liver alcohol dehydrogenase in ionic water-in-oil microemulsions, *Eur. J. Biochem.*, **163**, 609-17.
[186] Wolf, R. & Luisi, P. L. (1979) Micellar solubilization of enzymes in hydrocarbon solvents. Enzymatic activity and spectroscopic properties of ribonuclease in N-octane, *Biochem. Biophys. Res. Commun.*, **89**, 209-17.
[187] Schoenfeld, M., Montal, M. & Feher, G. (1980) Reaction center-phospholipid complex in organic solvents: formation and properties, *Biochemistry*, **19**, 1535-1542.
[188] Luisi, P. L. (1985) Enzymes Hosted in Reverse Micelles in Hydrocarbon Solution, *Angew. Chem. Int. Ed. Engl.*, **24**, 439-450.
[189] Martinek, K., Levashov, A. V., Klyachko, N., *et al.* (1986) Micellar enzymology, *Eur. J. Biochem.*, **155**, 453-68.
[190] Waks, M. (1986) Proteins and peptides in water-restricted environments, *Proteins: Struct. Funct. Bioinform.*, **1**, 4-15.
[191] Kazandjian, R. Z., Dordick, J. S. & Klibanov, A. M. (1986) Enzymatic analyses in organic solvents, *Biotechnol. Bioeng.*, **28**, 417-21.
[192] Kauzmann, W. (1956) Structural factors in protein denaturation, *J. Cell Physiol. Suppl.*, **47**, 113-31.
[193] Kauzmann, W. (1959) Some factors in the interpretation of protein denaturation, *Adv. Protein Chem.*, **14**, 1-63.
[194] Nelson, D. L. & Cox, M. M. (2011), Lehninger Principles of Biochemistry, *Macmillan Higher Education*.
[195] van Erp, S. H., Kamenskaya, E. O. & Khmelnitsky, Y. L. (1991) The effect of water content and nature of organic solvent on enzyme activity in low-water media. A quantitative description, *Eur. J. Biochem.*, **202**, 379-84.
[196] Garza-Ramos, G., Fernandez-Velasco, D. A., Ramirez, L., *et al.* (1992) Enzyme activation by denaturants in organic solvent systems with a low water content, *Eur. J. Biochem.*, **205**, 509-17.
[197] Ogino, H., Miyamoto, K. & Ishikawa, H. (1994) Organic-solvent-tolerant bacterium which secretes organic-solvent-stable lipolytic enzyme, *Appl. Environ. Microbiol.*, **60**, 3884-6.
[198] Yang, H., Cao, S. G., Han, S. P., *et al.* (1996) Enhancing the stereoselectivity and activity of Candida species lipase in organic solvent by noncovalent enzyme modification, *Ann. NY. Acad. Sci.*, **799**, 358-63.
[199] Jonsson, Wehtje, E., Adlercreutz, P. & Mattiasson, B. (1999) Thermodynamic and kinetic aspects on water vs. organic solvent as reaction media in the enzyme-catalysed reduction of ketones, *Biochim. Biophys. Acta*, **1430**, 313-22.
[200] Wescott, C. R. & Klibanov, A. M. (1994) The solvent dependence of enzyme specificity, *Biochim. Biophys. Acta*, **1206**, 1-9.
[201] Clark, D. S. (2004) Characteristics of nearly dry enzymes in organic solvents: implications for biocatalysis in the absence of water, *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, **359**, 1299-307.
[202] Norin, M., Haeffner, F., Hult, K. & Edholm, O. (1994) Molecular dynamics simulations of an enzyme surrounded by vacuum, water, or a hydrophobic solvent, *Biophys. J.*, **67**, 548-59.
[203] Trodler, P. & Pleiss, J. (2008) Modeling structure and flexibility of Candida antarctica lipase B in organic solvents, *BMC Struct. Biol.*, **8**, 9.
[204] Li, C., Tan, T., Zhang, H. & Feng, W. (2010) Analysis of the conformational stability and activity of Candida antarctica lipase B in organic solvents: insight from molecular dynamics and quantum mechanics/simulations, *J. Biol. Chem.*, **285**, 28434-41.
[205] Wedberg, R., Abildskov, J. & Peters, G. H. (2012) Protein dynamics in organic media at varying water activity studied by molecular dynamics simulation, *J. Phys. Chem. B*, **116**, 2575-85.
[206] Park, H. J., Joo, J. C., Park, K., *et al.* (2013) Prediction of the solvent affecting site and the computational design of stable Candida antarctica lipase B in a hydrophilic organic solvent, *J. Biotechnol.*, **163**, 346-52.
[207] Skjold-Jørgensen, J., Bhatia, V. K., Vind, J., *et al.* (2015) The Enzymatic Activity of Lipases Correlates with Polarity-Induced Conformational Changes: A Trp-Induced Quenching Fluorescence Study, *Biochemistry*, **54**, 4186-4196.

[208] Bordes, F., Barbe, S., Escalier, P., *et al.* (2010) Exploring the conformational states and rearrangements of Yarrowia lipolytica Lipase, *Biophys. J.*, **99**, 2225-34.
[209] James, J. J., Lakshmi, B. S., Seshasayee, A. S. & Gautam, P. (2007) Activation of Candida rugosa lipase at alkane-aqueous interfaces: a molecular dynamics study, *FEBS Lett.*, **581**, 4377-83.
[210] Peters, G. H., van Aalten, D. M., Svendsen, A. & Bywater, R. (1997) Essential dynamics of lipase binding sites: the effect of inhibitors of different chain length, *Protein Eng.*, **10**, 149-58.
[211] Peters, G. H., van Aalten, D. M., Edholm, O., *et al.* (1996) Dynamics of proteins in different solvent systems: analysis of essential motion in lipases, *Biophys. J.*, **71**, 2245-55.
[212] van der Spoel, D., van Maaren, P. J. & Caleman, C. (2012) GROMACS molecule & liquid database, *Bioinformatics*, **28**, 752-753.
[213] Caleman, C., van Maaren, P. J., Hong, M., *et al.* (2012) Force Field Benchmark of Organic Liquids: Density, Enthalpy of Vaporization, Heat Capacities, Surface Tension, Isothermal Compressibility, Volumetric Expansion Coefficient, and Dielectric Constant, *J. Chem. Theory Comput.*, **8**, 61-74.
[214] Rother, K., Hildebrand, P. W., Goede, A., *et al.* (2009) Voronoia: analyzing packing in protein structures, *Nucleic Acids Res.*, **37**, D393-5.
[215] Allen, M. P. & Tildesley, D. J. (1987), Computer simulation of liquids, *Clarendon Press*.
[216] Amadei, A., Linssen, A. B. & Berendsen, H. J. (1993) Essential dynamics of proteins, *Proteins*, **17**, 412-25.
[217] Amadei, A., Ceruso, M. A. & Di Nola, A. (1999) On the convergence of the conformational coordinates basis set obtained by the essential dynamics analysis of proteins' molecular dynamics simulations, *Proteins: Struct. Funct. Bioinform.*, **36**, 419-424.
[218] Adcock, S. A. & McCammon, J. A. (2006) Molecular dynamics: survey of methods for simulating the activity of proteins, *Chem. Rev.*, **106**, 1589-615.
[219] Bruccoleri, R. E. & Karplus, M. (1990) Conformational sampling using high-temperature molecular dynamics, *Biopolymers*, **29**, 1847-1862.
[220] Dunlop, E. A., Dodd, K. M., Seymour, L. A. & Tee, A. R. (2009) Mammalian target of rapamycin complex 1-mediated phosphorylation of eukaryotic initiation factor 4E-binding protein 1 requires multiple protein–protein interactions for substrate recognition, *Cell. Signal.*, **21**, 1073-1084.
[221] Yang-Yen, H.-F., Chambard, J.-C., Sun, Y.-L., *et al.* (1990) Transcriptional interference between c-Jun and the glucocorticoid receptor: mutual inhibition of DNA binding due to direct protein-protein interaction, *Cell*, **62**, 1205-1215.
[222] Suhane, T., Laskar, S., Advani, S., *et al.* (2014) Both charged linker region and ATPase domain of Hsp90 are essential for Rad51 dependent DNA repair, *Eukaryot. cell*, EC. 00159-14.
[223] Bader, G. D. & Hogue, C. W. (2002) Analyzing yeast protein–protein interaction data obtained from different sources, *Nat. biotechnol.*, **20**, 991-997.
[224] Uetz, P., Giot, L., Cagney, G., *et al.* (2000) A comprehensive analysis of protein–protein interactions in Saccharomyces cerevisiae, *Nature*, **403**, 623-627.
[225] Blatch, G. L. & Lässle, M. (1999) The tetratricopeptide repeat: a structural motif mediating protein-protein interactions, *Bioessays*, **21**, 932-939.
[226] Blaszczyk, M., Harmer, N. J., Chirgadze, D. Y., *et al.* (2015) Achieving high signal-to-noise in cell regulatory systems: Spatial organization of multiprotein transmembrane assemblies of FGFR and MET receptors, *Prog. Biophys. Mol. Biol.*, **118**, 103-111.
[227] Kučera, O. & Cifra, M. (2013) Cell-to-cell signaling through light: just a ghost of chance?, *Cell Commun. Signal*, **11**, 87-87.
[228] Neumann, S., Løvdok, L., Bentele, K., *et al.* (2014) Exponential Signaling Gain at the Receptor Level Enhances Signal-to-Noise Ratio in Bacterial Chemotaxis, *PLoS ONE*, **9**, e87815.
[229] Robinson, P. N. & Godfrey, M. (2000) The molecular genetics of Marfan syndrome and related microfibrillopathies, *J. Med. Genet.*, **37**, 9-25.
[230] Pena, V., Liu, S., Bujnicki, J. M., *et al.* (2007) Structure of a multipartite protein-protein interaction domain in splicing factor prp8 and its link to retinitis pigmentosa, *Mol. Cell*, **25**, 615-624.
[231] Mollet, G., Salomon, R., Gribouval, O., *et al.* (2002) The gene mutated in juvenile nephronophthisis type 4 encodes a novel protein that interacts with nephrocystin, *Nat. genet.*, **32**, 300-305.
[232] Guney, E. & Oliva, B. (2012) Exploiting Protein-Protein Interaction Networks for Genome-Wide Disease-Gene Prioritization, *PLoS ONE*, **7**, e43557.
[233] Ruf, R. G., Xu, P.-X., Silvius, D., *et al.* (2004) SIX1 mutations cause branchio-oto-renal syndrome by disruption of EYA1–SIX1–DNA complexes, *Proc. Natl. Acad. Sci. USA*, **101**, 8090-8095.

[234] Brook-Carter, P. T., Peral, B., Ward, C. J., *et al.* (1994) Deletion of the TSC2 and PKD1 genes associated with severe infantile polycystic kidney disease—a contiguous gene syndrome, *Nat. Genet.*, **8**, 328-332.
[235] Fry, D. C. & Vassilev, L. T. (2005) Targeting protein–protein interactions for cancer therapy, *Int. J. Mol. Med.*, **83**, 955-963.
[236] Porta-Pardo, E., Garcia-Alonso, L., Hrabe, T., *et al.* (2015) A pan-cancer catalogue of cancer driver protein interaction interfaces, *PLoS Comput. Biol.*, **11**, e1004518.
[237] Morris, J. A., Kandpal, G., Ma, L. & Austin, C. P. (2003) DISC1 (Disrupted-In-Schizophrenia 1) is a centrosome-associated protein that interacts with MAP1A, MIPT3, ATF4/5 and NUDEL: regulation and loss of interaction with mutation, *Hum. Mol. Genet.*, **12**, 1591-1608.
[238] O'Connor, V., Heuss, C., De Bello, W. M., *et al.* (1997) Disruption of syntaxin-mediated protein interactions blocks neurotransmitter secretion, *Proc. Natl. Acad. Sci. USA*, **94**, 12186-12191.
[239] Arkin, M. R. & Wells, J. A. (2004) Small-molecule inhibitors of protein–protein interactions: progressing towards the dream, *Nat. Rev. Drug Discov.*, **3**, 301-317.
[240] Shangary, S. & Wang, S. (2009) Small-molecule inhibitors of the MDM2-p53 protein-protein interaction to reactivate p53 function: a novel approach for cancer therapy, *Annu. Rev. Pharmacol. Toxicol.*, **49**, 223.
[241] Wells, J. A. & McClendon, C. L. (2007) Reaching for high-hanging fruit in drug discovery at protein–protein interfaces, *Nature*, **450**, 1001-1009.
[242] Pawson, T. & Nash, P. (2003) Assembly of cell regulatory systems through protein interaction domains, *Science*, **300**, 445-452.
[243] Virkamäki, A., Ueki, K. & Kahn, C. R. (1999) Protein–protein interaction in insulin signaling and the molecular mechanisms of insulin resistance, *J. Clin. Invest.*, **103**, 931-943.
[244] Xiao, S.-H. & Manley, J. L. (1997) Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing, *Genes Dev.*, **11**, 334-344.
[245] Choudhary, C., Kumar, C., Gnad, F., *et al.* (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions, *Science*, **325**, 834-840.
[246] Rabut, G. & Peter, M. (2008) Function and regulation of protein neddylation, *EMBO Rep.*, **9**, 969-976.
[247] Geiss-Friedlander, R. & Melchior, F. (2007) Concepts in sumoylation: a decade on, *Nat. Rev. Mol. Cell Biol.*, **8**, 947-956.
[248] Strott, C. A. (2002) Sulfonation and molecular action, *Endocr. Rev.*, **23**, 703-732.
[249] Apweiler, R., Hermjakob, H. & Sharon, N. (1999) On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database, *Biochim. Biophys. Acta*, **1473**, 4-8.
[250] Jaeger, K.-E., Ransac, S., Dijkstra, B. W., *et al.* (1994) Bacterial lipases, *FEMS Microbiol. Rev.*, **15**, 29-63.
[251] Jaeger, K.-E. & Reetz, M. T. (1998) Microbial lipases form versatile tools for biotechnology, *Trends Biotechnol.*, **16**, 396-403.
[252] Gupta, R., Gupta, N. & Rathi, P. (2004) Bacterial lipases: an overview of production, purification and biochemical properties, *Appl. Microbiol. Biotechnol.*, **64**, 763-781.
[253] Haque, N. & Prabhu, N. P. (2016) Lid closure dynamics of porcine pancreatic lipase in aqueous solution, *Biochim. Biophys. Acta*, **1860**, 2313-2325.
[254] Borgstrom, B. & Erlanson, C. (1973) Pancreatic lipase and co-lipase. Interactions and effects of bile salts and other detergents, *Eur. J. Biochem.*, **37**, 60-8.
[255] Patton, J. S., Albertsson, P. A., Erlanson, C. & Borgstrom, B. (1978) Binding of porcine pancreatic lipase and colipase in the absence of substrate studies by two-phase partition and affinity chromatography, *J. Biol. Chem.*, **253**, 4195-202.
[256] Charles, M., Astier, M., Sauve, P. & Desnuell, P. (1975) Interactions of colipase with bile salt micelles. 1. Ultracentrifugation studies, *Eur. J. Biochem.*, **58**, 555-559.
[257] Sari, H., Entressangles, B. & Desnuelle, P. (1975) Interactions of colipase with bile salt micelles. 2. Study by dialysis and spectrophotometry, *Eur. J. Biochem.*, **58**, 561-565.
[258] Donner, J., Spink, C. H., Borgstom, B. & Sjoholm, I. (1976) Interactions between pancreatic lipase, co-lipase, and taurodeoxycholate in the absence of triglyceride substrate, *Biochemistry*, **15**, 5413-7.
[259] Semeriva, M. & Desnuelle, P. (1979) Pancreatic lipase and colipase. An example of heterogeneous biocatalysis, *Adv. Enzymol. Relat. Areas Mol. Biol.*, **48**, 319-70.

[260] Rathelot, J., Julien, R., Canioni, P., *et al.* (1975) Studies on the effect of bile salt and colipase on enzymatic lipolysis. Improved method for the determination of pancreatic lipase and colipase, *Biochimie,* **57**, 1117-22.
[261] Maylie, M. F., Charles, M., Astier, M. & Desnuelle, P. (1973) On porcine pancreatic colipase: large scale purification and some properties, *Biochem. Biophys. Res. Commun. ,* **52**, 291-7.
[262] Pignol, D., Ayvazian, L., Kerfelec, B., *et al.* (2000) Critical role of micelles in pancreatic lipase activation revealed by small angle neutron scattering, *J. Biol. Chem.,* **275**, 4220-4.
[263] Hermoso, J., Pignol, D., Penel, S., *et al.* (1997) Neutron crystallographic evidence of lipase-colipase complex activation by a micelle, *EMBO J.*, **16**, 5531-6.
[264] Charles, M., Semeriva, M. & Chabre, M. (1980) Small-angle neutron scattering study of the association between porcine pancreatic colipase and taurodeoxycholate micelles, *J. Mol. Biol.,* **139**, 297-317.
[265] Breg, J. N., Sarda, L., Cozzone, P. J., *et al.* (1995) Solution structure of porcine pancreatic procolipase as determined from 1H homonuclear two-dimensional and three-dimensional NMR, *Eur. J. Biochem.,* **227**, 663-72.
[266] Koziara, K. B., Stroet, M., Malde, A. K. & Mark, A. E. (2014) Testing and validation of the Automated Topology Builder (ATB) version 2.0: prediction of hydration free enthalpies, *Journal of computer-aided molecular design,* **28**, 221-233.
[267] van Tilbeurgh, H., Sarda, L., Verger, R. & Cambillau, C. (1992) Structure of the pancreatic lipase-procolipase complex, *Nature,* **359**, 159-62.
[268] Haque, N. & Prabhu, N. P. (2016) Lid dynamics of porcine pancreatic lipase in non-aqueous solvents, *Biochim. Biophys. Acta,* **1860**, 2326-2334.
[269] Small, D. M. (1968) Size and Structure of Bile Salt Micelles in *Molecular Association in Biological and Related Systems* pp. 31-52, American Chemical Society.
[270] Mazer, N. A., Carey, M. C., Kwasnick, R. F. & Benedek, G. B. (1979) Quasielastic light scattering studies of aqueous biliary lipid systems. Size, shape, and thermodynamics of bile salt micelles, *Biochemistry,* **18**, 3064-3075.
[271] Brockerhoff, H. (1973) A model of pancreatic lipase and the orientation of enzymes at interfaces, *Chem. Phys. Lipids,* **10**, 215-22.
[272] Chapus, C., Sari, H., Semeriva, M. & Desnuelle, P. (1975) Role of colipase in the interfacial adsorption of pancreatic lipase at hydrophilic interfaces, *FEBS Lett.,* **58**, 155-8.
[273] Borgstrom, B. (1977) Physico-chemical characteristics of the lipase-colipase-bile salt system, *Expos. Annu. Biochim. Med. ,* **33**, 173-82.
[274] Winkler, F. K., D'Arcy, A. & Hunziker, W. (1990) Structure of human pancreatic lipase, *Nature,* **343**, 771-774.
[275] Carmona, P., Navarro, R. & Hernanz, A. (1997), Spectroscopy of Biological Molecules: Modern Trends, *Kluwer Academic Publishers*.
[276] Gauglitz, G. & Vo-Dinh, T. (2006), Bioanalysis in Handbook of Spectroscopy, *Wiley*.
[277] McClure, W. O. & Edelman, G. M. (1966) Fluorescent probes for conformational states of proteins. I. Mechanism of fluorescence of 2-p-toluidinylnaphthalene-6-sulfonate, a hydrophobic probe, *Biochemistry,* **5**, 1908-1919.
[278] Beyer, C. F., Craig, L. C. & Gibbons, W. A. (1972) Interaction of the fluorescent probe 2-p-toluidinylnaphthalene-6-sulfonate with peptides. Structural requirements for binding and fluorescence enhancement, *Biochemistry,* **11**, 4920-4926.
[279] Santra, M. K., Dasgupta, D. & Panda, D. (2006) Pyrene excimer fluorescence of yeast alcohol dehydrogenase: a sensitive probe to investigate ligand binding and unfolding pathway of the enzyme, *Photochem. Photobiol.,* **82**, 480-6.
[280] Bernal-Perez, L. F., Prokai, L. & Ryu, Y. (2012) Selective N-terminal fluorescent labeling of proteins using 4-chloro-7-nitrobenzofurazan: a method to distinguish protein N-terminal acetylation, *Anal. Biochem.,* **428**, 13-5.
[281] Hawe, A., Sutter, M. & Jiskoot, W. (2008) Extrinsic fluorescent dyes as tools for protein characterization, *Pharm. Res.,* **25**, 1487-1499.
[282] Turro, N. J., Lei, X.-G., Ananthapadmanabhan, K. P. & Aronson, M. (1995) Spectroscopic Probe Analysis of Protein-Surfactant Interactions: The BSA/SDS System, *Langmuir,* **11**, 2525-2533.
[283] McClure, W. O. & Edelman, G. M. (1967) Fluorescent Probes for Conformational States of Proteins. II. The Binding of 2-p-Toluidinylnaphthalene-6-sulfonate to α-Chymotrypsin, *Biochemistry,* **6**, 559-566.
[284] Wang, J. L. & Edelman, G. M. (1971) Fluorescent probes for conformational states of proteins. IV. The pepsinogen-pepsin conversion, *J. Biol. Chem.,* **246**, 1185-91.

[285] Tejaswi Naidu, K. & Prakash Prabhu, N. (2011) Protein–Surfactant Interaction: Sodium Dodecyl Sulfate-Induced Unfolding of Ribonuclease A, *J. Phys. Chem. B,* **115**, 14760-14767.
[286] Stoitsova, S. R., Boteva, R. N. & Doyle, R. J. (2003) Binding of hydrophobic ligands by Pseudomonas aeruginosa PA-I lectin, *Biochim. Biophys. Acta.*, **20**, 213-9.
[287] Coons, A. H. & Kaplan, M. H. (1950) Localization of antigen in tissue cells; improvements in a method for the detection of antigen by means of fluorescent antibody, *J. Exp. Med.*, **91**, 1-13.
[288] Bompiani, A. & Rossi, A. (1951) Relation between fluorescence and blood proteins in the terminal stage of pregnancy, in labor, in puerperium and in placental blood, *Arch. Ostet. Ginecol.*, **56**, 883-91.
[289] Biegeleisen, J. Z., Jr., Scott, L. V. & Lewis, V., Jr. (1959) Rapid diagnosis of herpes simplex virus infections with fluorescent antibody, *Science*, **129**, 640-1.
[290] Weber, G. (1960) Fluorescence-polarization spectrum and electronic-energy transfer in proteins, *Biochem. J.*, **75**, 345-52.
[291] Chen, L., Yang, R., Hu, X. L., *et al.* (2013) Expression analysis of chlorophyllid alpha binding protein, a secretory, red fluorescence protein in the midgut of silkworm, Bombyx mori, *Insect. Sci.*, **21**, 1744-7917.
[292] Wang, J. C., Wang, P. Y., Chen, H. I., *et al.* (2013) Lie Group Analysis of the Photo-Induced Fluorescence of Drosophila Oogenesis with the Asymmetrically Localized Gurken Protein, *PLoS One,* **8**, e65143.
[293] Albani, J. R. (2014) Origin of Tryptophan Fluorescence Lifetimes Part 1. Fluorescence Lifetimes Origin of Tryptophan Free in Solution, *J. Fluoresc.*, **24**, 93-104.
[294] Albani, J. R. (2014) Origin of Tryptophan Fluorescence Lifetimes. Part 2: Fluorescence Lifetimes Origin of Tryptophan in Proteins, *J. Fluoresc.*, **24**, 105-117.
[295] Scott, J. N. & Callis, P. R. (2013) Insensitivity of Tryptophan Fluorescence to Local Charge Mutations, *J. Phys. Chem. B,* **117**, 9598-9605.
[296] McMillan, A. W., Kier, B. L., Shu, I., *et al.* (2013) Fluorescence of Tryptophan in Designed Hairpin and Trp-Cage Miniproteins: Measurements of Fluorescence Yields and Calculations by Quantum Mechanical Molecular Dynamics Simulations, *J. Phys. Chem. B,* **117**, 1790-1809.
[297] Tasaki, I., Watanabe, A. & Hallett, M. (1971) Properties of squid axon membrane as revealed by a hydrophobic probe, 2-p-toluidinylnaphthalene-6-sulfonate, *Proc. Natl. Acad. Sci. USA,* **68**, 938-41.
[298] Follenius, A. & Gerard, D. (1984) Fluorescence investigations of calmodulin hydrophobic sites, *Biochem. Biophys. Res. Commun.*, **119**, 1154-60.
[299] Wang, Z. J., Ren, L. X., Zhao, Y. Q., *et al.* (2007) Metal ions-induced conformational change of P23 by using TNS as fluorescence probe, *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, **66**, 1323-6.
[300] Niu, S., Gopidas, K. R., Turro, N. J. & Gabor, G. (1992) Formation of premicellar clusters of 2-p-toluidino naphthalene-6-sulfonate with cationic detergents, *Langmuir,* **8**, 1271-1277.
[301] Beyer, C. F., Craig, L. C. & Gibbons, W. A. (1973) Structural requirements for binding and fluorescence enhancement of the fluorescent probe TNS with peptides, *Nat. New. Biol.*, **241**, 78-80.
[302] Fioroni, M., Burger, K., Mark, A. E. & Roccatano, D. (2001) Model of 1,1,1,3,3,3-Hexafluoro-propan-2-ol for Molecular Dynamics Simulations, *J. Phys. Chem. B,* **105**, 10967-10975.
[303] Warren, D. B., Chalmers, D. K., Hutchison, K., *et al.* (2006) Molecular dynamics simulations of spontaneous bile salt aggregation, *Colloids Surf. Physicochem. Eng. Aspects,* **280**, 182-193.
[304] Cui, Y., Xing, C. & Ran, Y. (2010) Molecular dynamics simulations of hydrotropic solubilization and self-aggregation of nicotinamide, *J. Pharm. Sci.*, **99**, 3048-59.
[305] Brandman, R., Brandman, Y. & Pande, V. S. (2012) A-site residues move independently from P-site residues in all-atom molecular dynamics simulations of the 70S bacterial ribosome, *PLoS One,* **7**, e29377.
[306] Tieleman, D. P., van der Spoel, D. & Berendsen, H. J. C. (2000) Molecular Dynamics Simulations of Dodecylphosphocholine Micelles at Three Different Aggregate Sizes: Micellar Structure and Chain Relaxation, *J. Phys. Chem. B,* **104**, 6380-6388.
[307] Raghavan, K., Reddy, M. R. & Berkowitz, M. L. (1992) A molecular dynamics study of the structure and dynamics of water between dilauroylphosphatidylethanolamine bilayers, *Langmuir,* **8**, 233-240.
[308] de Vries, A. H., Mark, A. E. & Marrink, S. J. (2004) Molecular dynamics simulation of the spontaneous formation of a small DPPC vesicle in water in atomistic detail, *J. Am. Chem. Soc.*, **126**, 4488-9.

[309] Malde, A. K., Zuo, L., Breeze, M., *et al.* (2011) An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0, *J. Chem. Theory Comput.*, **7**, 4026-4037.
[310] Stumpe, M. C. & Grubmuller, H. (2009) Urea impedes the hydrophobic collapse of partially unfolded proteins, *Biophys. J.*, **96**, 3744-52.
[311] Bussi, G., Donadio, D. & Parrinello, M. (2007) Canonical sampling through velocity rescaling, *J. Chem. Phys.*, **126**, 014101.
[312] Trott, O. & Olson, A. J. (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, *J. Comput. Chem.*, **31**, 455-461.
[313] Jones, G., Willett, P. & Glen, R. C. (1995) Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation, *J. Mol. Biol.*, **245**, 43-53.
[314] Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions, *Protein Eng.*, **8**, 127-34.
[315] Albani, J. R. (2009) Fluorescence origin of 6,P-toluidinyl-naphthalene-2-sulfonate (TNS) bound to proteins, *J. Fluoresc.*, **19**, 399-408.
[316] Reichardt, C. (1979) Empirical Parameters of Solvent Polarity as Linear Free-Energy Relationships, *Angew. Chem. Int. Ed. Engl.*, **18**, 98-110.
[317] Heyne, B. (2016) Self-assembly of organic dyes in supramolecular aggregates, *Photochem. Photobiol. Sci.*, **15**, 1103-1114.
[318] Lakowicz, J. R. (2007), Principles of Fluorescence Spectroscopy, *Springer US*.
[319] Uversky, V. N., Winter, S. & Löber, G. (1998) Self-association of 8-anilino-1-naphthalene-sulfonate molecules: Spectroscopic characterization and application to the investigation of protein folding, *Biochim. Biophys. Acta*, **1388**, 133-142.
[320] Basak, S. & Chattopadhyay, K. (2014) Studies of protein folding and dynamics using single molecule fluorescence spectroscopy, *Phys. Chem. Chem. Phys.*, **16**, 11139-11149.
[321] Royer, C. A. & Scarlata, S. F. (2008) Chapter 5 Fluorescence approaches to quantifying biomolecular interactions in *Methods in Enzymology* pp. 79-106, Academic Press.
[322] Krishnamoorthy, G. (2003) Fluorescence spectroscopy in molecular description of biological processes, *Indian J. Biochem. Biophys.*, **40**, 147-59.
[323] Brucale, M., Schuler, B. & Samorì, B. (2014) Single-Molecule Studies of Intrinsically Disordered Proteins, *Chem. Rev.*, **114**, 3281-3317.
[324] Royer, C. a. (2006) Probing protein folding and conformational transitions with fluorescence probing protein folding and conformational transitions with fluorescence, *Chem. Rev.*, **106**, 1769-1784.
[325] Hall, M. D., Yasgar, A., Peryea, T., *et al.* (2016) Fluorescence polarization assays in high-throughput screening and drug discovery: a review, *Methods. Appl. Fluoresc.*, **4**, 022001-022001.
[326] Bachovchin, D. a., Brown, S. J., Rosen, H. & Cravatt, B. F. (2009) Identification of selective inhibitors of uncharacterized enzymes by high-throughput screening with fluorescent activity-based probes, *Nat. Biotechnol.*, **27**, 387-94.
[327] Munishkina, L. A. & Fink, A. L. (2007) Fluorescence as a method to reveal structures and membrane-interactions of amyloidogenic proteins, *Biochim. Biophys. Acta*, **1768**, 1862-1885.
[328] Edelman, G. M. & McClure, W. O. (1968) Fluorescent probes and the conformation of proteins, *Acc. Chem. Res.*, **1**, 65-70.
[329] Haynes, D. H. & Staerk, H. (1974) 1-Anilino-8-naphthalenesulfonate: a fluorescent probe of membrane surface structure, composition and mobility, *J. Membrane Biol.* , **17**, 313-340.
[330] Brand, L. & Gohlke, J. R. (1972) Fluorescence probes for structure, *Annu. Rev. Biochem.*, **41**, 843-868
[331] Azzi, A. (1975) The application of fluorescent probes in membrane studies, *Q. Rev. Biophys.*, **8**, 237-316.
[332] Rocha, S., Keersmaecker, H. D., Uji-i, H., *et al.* (2014), Fluorescence Spectroscopy and Microscopy: Methods and Protocols pp. 793-812, *Humana Press.*
[333] Kumar, E. K. & Prabhu, N. P. (2014) Differential effects of ionic and non-ionic surfactants on lysozyme fibrillation, *Phys. Chem. Chem. Phys.*, **16**, 24076-88.
[334] Honda, C., Kamizono, H., Matsumoto, K. I. & Endo, K. (2004) Studies on bovine serum albumin-sodium dodecyl sulfate complexes using pyrene fluorescence probe and 5-doxylstearic acid spin probe, *J. Colloid Interface Sci.*, **278**, 310-317.
[335] Mishra, R., Sjolander, D. & Hammarstrom, P. (2011) Spectroscopic characterization of diverse amyloid fibrils in vitro by the fluorescent dye Nile red, *Mol. Biosyst.*, **7**, 1232-1240.

[336] Palais, C., Machaidze, G., Gurny, R. & Arvinte, T. (2009) New methods allowing the detection of protein aggregates, *MAbs*, **1**, 142–150.
[337] Fernandez-Flores, a. (2011) A review of amyloid staining: methods and artifacts, *Biotech. Histochem.*, **86**, 293–301.
[338] Wagner, B. D. (2009) The use of coumarins as environmentally-sensitive fluorescent probes of heterogeneous inclusion systems, *Molecules*, **14**, 210-237.
[339] Haidekker, M. A. & Theodorakis, E. A. (2007) Molecular rotors-fluorescent biosensors for viscosity and flow, *Org. Biomol. Chem.*, **5**, 1669-78.
[340] Sainlos, M. & Imperiali, B. (2007) Tools for investigating peptide–protein interactions: peptide incorporation of environment-sensitive fluorophores via on-resin derivatization, *Nat. Protoc.*, **2**, 3201-3209.
[341] Raghuraman, H., Pradhan, S. K. & Chattopadhyay, A. (2004) Effect of urea on the organization and dynamics of Triton X-100 micelles: a fluorescence approach, *J. Phys. Chem. B*, **108**, 2489-2496.
[342] Katilius, E., Katiliene, Z. & Woodbury, N. W. (2005) Methods to create fluorescent biosensors using aptamers with fluorescent base analogs in, Google Patents,
[343] Gorbenko, G. P. (2011) Fluorescence spectroscopy of protein oligomerization in membranes, *J. Fluoresc.*, **21**, 945-951.
[344] Noronha, M., Santos, R., Paci, E., *et al.* (2009) Fluorescence Lifetimes of Tyrosine Residues in Cytochrome c″ as Local Probes to Study Protein Unfolding, *J. Phys. Chem. B*, **113**, 4466-4474.
[345] Miller, J. N. (2005) Fluorescence energy transfer methods in bioanalysis, *The Analyst*, **130**, 265-270.
[346] Bothra, A., Bhattacharyya, A., Mukhopadhyay, C., *et al.* (1998) A fluorescence spectroscopic and molecular dynamics study of bis-ANS/protein interaction, *J. Biomol. Struct. Dyn.*, **15**, 959-66.
[347] Kamal, M. Z., Ali, J. & Rao, N. M. (2013) Binding of bis-ANS to Bacillus subtilis lipase: A combined computational and experimental investigation, *Biochim. Biophys. Acta*, **1834**, 1501-1509.
[348] Sharma, K. K., Kumar, G. S., Murphy, A. S. & Kester, K. (1998) Identification of 1,1'-bi(4-anilino)naphthalene-5,5'-disulfonic acid binding sequences in alpha-crystallin, *J. Bio. Chem.*, **273**, 15474-15478.
[349] Flanagan, M. T. & Ainsworth, S. (1968) The binding of aromatic sulphonic acids to bovine serum albumin, *Biochim. Biophys. Acta*, **168**, 16-26.
[350] Datta, A., Mandal, D., Kumar Pal, S., *et al.* (1998) Interaction of Triton X-100 with cyclodextrins. A fluorescence study, *J. Chem. Soc. Faraday Trans.*, **94**, 3471-3475.
[351] Frank, J., Holzwarth, J. F., van Hoek, A., *et al.* (1997) Kinetics and equilibrium binding of the dyes TNS and RH421 to ribulose 1,5-bisphosphate carboxylase/oxygenase (RUBISCO), *J. Chem. Soc. Faraday Trans.*, **93**, 2379-2385.
[352] Lin, W. Y., Eads, C. D. & Villafranca, J. J. (1991) Fluorescent probes for measuring the binding constants and distances between the metal Ions bound to, *Biochemistry*, **30**, 3421-3426.
[353] Seale, J. W., Gorovits, B. M., Ybarra, J. & Horowitz, P. M. (1996) Reversible oligomerization and denaturation of the chaperonin GroES, *Biochemistry*, **35**, 4079-4083.
[354] Wang, Z. J., Ren, L. X., Zhao, Y. Q., *et al.* (2008) Investigation on the binding of TNS to centrin, an EF-hand protein, *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, **70**, 892-897.
[355] D'Alfonso, L., Collini, M. & Baldini, G. (1999) Evidence of heterogeneous 1-anilinonaphthalene-8-sulfonate binding to β-lactoglobulin from fluorescence spectroscopy, *Biochim. Biophys. Acta*, **1432**, 194-202.
[356] Matulis, D. & Lovrien, R. (1998) 1-Anilino-8-naphthalene sulfonate anion-protein binding depends primarily on ion pair formation, *Biophys. j.*, **74**, 422-429.
[357] Gasymov, O. K. & Glasgow, B. J. (2007) ANS fluorescence: Potential to augment the identification of the external binding sites of proteins, *Biochim. Biophys. Acta*, **1774**, 403-411.
[358] Wang, R. & Bright, F. V. (1993) Detailed Investigation of 2-p-toluidinylnaphthalene-6-sulfonate (TNS) Binding to Bovine Serum Albumin (BSA) by Steady-State and Time-Resolved Fluorescence Spectroscopy, *Appl. Spectrosc.*, **47**, 792-799.
[359] Cardamone, M. & Puri, N. K. (1992) Spectrofluorimetric assessment of the surface hydrophobicity of proteins, *Biochem. J.*, **282**, 589-593.
[360] Greene, F. C. (1984) Interactions of anionic and cationic fluorescent probes with proteins: The effect of charge, *J. Protein Chem.*, **3**, 167-180.
[361] Haque, N. & Prabhu, N. P. (2014) Insights into protein–TNS (2-p-toluidinylnaphthalene-6-sulfonate) interaction using molecular dynamics simulation, *J. Mol. Struc.*, **1068**, 261-269.

[362] Das, K., Sarkar, N. & Bhattacharyya, K. (1993) Interaction of urea with fluorophores bound to protein surfaces, *J. Chem. Soc. Faraday Trans.*, **89**, 1959-1959.
[363] Rappé, A. K., Casewit, C. J., Colwell, K., *et al.* (1992) UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations, *J. Am. Chem. Soc.*, **114**, 10024-10035.
[364] Van Gunsteren, W. & Berendsen, H. (1988) A leap-frog algorithm for stochastic dynamics, *Mol. Simul.*, **1**, 173-185.
[365] Penzer, G. R. (1972) 1-Anilinonaphtalene-8-sulphonate: The depedence of emission spectra on molecular conformation studied by fluorescence and proton-magnetic resonance, *Eur. J. Biochem.*, **25**, 218-228.
[366] Wurthner, F., Kaiser, T. E. & Saha-Moller, C. R. (2011) J-aggregates: From serendipitous discovery to supramolecular engineering of functional dye materials, *Angew. Chem. Int. Ed. Engl.*, **50**, 3376-3410.

Appendix A1

	Donor	Acceptor		Donor	Acceptor		Donor	Acceptor
1	ARG266N	CYS262O	31	THR256OG1	GLY251O	61	GLN245N	GLN245NE2
2	LEU265N	CYS262O	32	THR256OG1	GLU254OE1	62	SER244OG	ILE242O
3	ASN263ND2	GLN239O	33	THR256OG1	ARG257N	63	SER244OG	LEU243O
4	ALA261N	PHE259O	34	THR256N	GLY251O	64	SER244OG	SER244O
5	VAL260N	ASN241OD1	35	GLY255N	GLY251O	65	SER244OG	ILE246O
6	VAL260N	ASN241ND2	36	GLY255N	ILE252O	66	SER244N	ILE242O
7	PHE259N	ASP258OD1	37	GLU254N	ASP250O	67	ILE242N	ASN241OD1
8	PHE259N	ASP258OD2	38	GLU254N	GLY251O	68	ASN241ND2	ILE242O
9	ASP258N	ARG257NE	39	GLU254N	ILE252O	69	ASN241ND2	SER244OG
10	ASP258N	ARG257NH1	40	TRP253N	ILE249O	70	ASN241ND2	ASP258OD2
11	ARG257NH2	ILE242O	41	TRP253N	ASP250O	71	ASN241ND2	ASP258O
12	ARG257NH2	SER244OG	42	ILE252N	ASP248O	72	ASN241ND2	VAL260O
13	ARG257NH2	ILE246O	43	ILE252N	ILE249O	73	LYS240N	GLN239OE1
14	ARG257NH2	ASP248OD1	44	GLY251N	ASP248OD1	74	GLN239NE2	GLN239N
15	ARG257NH2	ASP248OD2	45	GLY251N	ASP248OD2	75	GLN239NE2	LYS240O
16	ARG257NH2	ASP250OD1	46	GLY251N	ASP248O	76	GLN239N	GLN239OE1
17	ARG257NH2	ASP250OD2	47	ASP250N	ASP248OD1	77	GLN234NE2	CYS238O
18	ARG257NH1	SER244OG	48	ASP250N	ASP248OD2	78	GLN234NE2	GLN239O
19	ARG257NH1	ILE246O	49	ILE249N	ASP248OD1	79	GLN234NE2	LYS240N
20	ARG257NH1	ASP248OD1	50	ILE249N	ASP248OD2	80	TYR115N	ASP250OD1
21	ARG257NH1	GLU254OE2	51	ASP248N	ARG257NE	81	TYR115N	ASP250OD2
22	ARG257NH1	THR256OG1	52	ASP248N	ARG257NH1	82	ARG112NH2	GLU254OE1
23	ARG257NH1	THR256O	53	VAL247N	GLN245O	83	ARG112NH2	GLU254OE2
24	ARG257NE	ASP248OD1	54	ILE246N	SER244OG	84	ARG112NH2	GLU254O
25	ARG257NE	ASP248OD2	55	ILE246N	SER244O	85	ARG112NH1	TRP253O
26	ARG257NE	ASP248O	56	GLN245NE2	LEU243O	86	ARG112NH1	GLU254OE1
27	ARG257NE	THR256O	57	GLN245NE2	SER244O	87	ARG112NH1	GLU254OE2
28	ARG257NE	ASP258O	58	GLN245NE2	GLN245O	88	LYS81NZ	GLU254OE2
29	ARG257N	GLY251O	59	GLN245N	SER244OG	89	LYS81NZ	GLU254O
30	ARG257N	GLY255O	60	GLN245N	GLN245OE1			

Appendix A2

	Donor	Acceptor		Donor	Acceptor		Donor	Acceptor
1	ARG23NH1	ASP248OD1	25	TYR115N	GLY251O	49	GLN234NE2	CYS238O
2	ARG23NH1	ASP248OD2	26	TYR115N	TRP253NE1	50	GLN234NE2	GLN239O
3	ARG23NH2	ASP248OD1	27	TYR115N	LEU254O	51	MET235N	ASN263OD1
4	ARG23NH2	ASP248OD2	28	TYR115N	GLY255O	52	MET235N	ASN263ND2
5	LYS81N	LEU254O	29	TYR115N	ASP258OD2	53	GLY237N	GLN239OE1
6	LYS81N	ASP258OD1	30	TYR115N	ASP258O	54	GLY237N	GLN239NE2
7	LYS81NZ	GLY251O	31	TYR115OH	TRP253NE1	55	CYS238N	GLN239OE1
8	LYS81NZ	GLY255O	32	TYR115OH	TRP253O	56	CYS238N	ASN263OD1
9	LYS81NZ	THR256O	33	TYR115OH	LEU254O	57	GLN239N	GLN239OE1
10	LYS81NZ	ASP258OD1	34	TYR115OH	ASP258O	58	GLN239N	GLN239NE2
11	LYS81NZ	ASP258OD2	35	HIS152NE2	HIS264ND1	59	GLN239N	ASN241OD1
12	LYS81NZ	ASP258O	36	HIS152NE2	HIS264O	60	GLN239NE2	CYS238O
13	ARG112NE	ASP258OD1	37	SER153OG	HIS264ND1	61	GLN239NE2	GLN239O
14	ARG112NE	ASP258OD2	38	THR205OG1	ASN263OD1	62	GLN239NE2	ASN241OD1
15	ARG112NE	ASP258O	39	THR205OG1	ASN263ND2	63	LYS240N	CYS238O
16	ARG112NH1	ASP258OD1	40	ASN213ND2	CYS238O	64	LYS240N	ASN241OD1
17	ARG112NH1	ASP258OD2	41	ASN213ND2	GLN239O	65	LYS240NZ	ILE242O
18	ARG112NH2	ASP258OD1	42	ASN213ND2	ASN241OD1	66	ASN241N	CYS238O
19	ARG112NH2	ASP258OD2	43	ASN213ND2	ILE242O	67	ASN241N	ASN241OD1
20	THR113N	ASP258OD2	44	ASN213ND2	VAL260O	68	ASN241ND2	CYS238O
21	GLY114N	GLY255O	45	LEU214N	ILE242O	69	ASN241ND2	GLN239O
22	GLY114N	ASP258OD2	46	LEU214N	SER244O	70	ASN241ND2	ASN241O
23	TYR115N	ILE249O	47	LYS233N	GLN239OE1	71	ASN241ND2	ILE242O
24	TYR115N	VAL250O	48	GLN234N	GLN239OE1	72	ASN241ND2	ASP258O

Appendix A2

	Donor	Acceptor		Donor	Acceptor		Donor	Acceptor
73	ASN241ND2	VAL260O	132	TRP253N	VAL250O	190	ARG257NH1	THR256O
74	ILE242N	ASN241OD1	133	TRP253N	GLY251O	191	ARG257NH1	ASP258OD1
75	ILE242N	ASN241ND2	134	TRP253N	LEU254O	192	ARG257NH1	PHE259O
76	LEU243N	LYS240O	135	TRP253NE1	GLN245OE1	193	ARG257NH2	ASN241OD1
77	LEU243N	ASN241O	136	TRP253NE1	GLN245O	194	ARG257NH2	ILE246O
78	LEU243N	SER244OG	137	TRP253NE1	ILE249O	195	ARG257NH2	VAL247O
79	LEU243N	THR256OG1	138	TRP253NE1	VAL250N	196	ARG257NH2	ASP248OD1
80	SER244N	LYS240O	139	TRP253NE1	VAL250O	197	ARG257NH2	ASP248OD2
81	SER244N	ILE242O	140	LEU254N	VAL250O	198	ARG257NH2	THR256OG1
82	SER244OG	ILE242O	141	LEU254N	GLY251O	199	ARG257NH2	ASP258OD1
83	SER244OG	LEU243O	142	GLY255N	ASP248OD1	200	ARG257NH2	ASP258OD2
84	SER244OG	SER244O	143	GLY255N	GLY251O	201	ARG257NH2	PHE259O
85	SER244OG	GLN245OE1	144	GLY255N	ILE252O	202	ASP258N	GLY255O
86	SER244OG	GLN245O	145	GLY255N	TRP253O	203	ASP258N	THR256OG1
87	SER244OG	ILE246O	146	GLY255N	ARG257O	204	ASP258N	THR256O
88	SER244OG	TRP253O	147	THR256N	LEU243O	205	ASP258N	ARG257NH1
89	SER244OG	LEU254O	148	THR256N	SER244OG	206	ASP258N	ASP258OD1
90	SER244OG	THR256OG1	149	THR256N	GLN245O	207	ASP258N	ASP258OD2
91	SER244OG	ASP258O	150	THR256N	ASP248OD1	208	ASP258N	ASP258O
92	GLN245N	LEU243O	151	THR256N	ASP248OD2	209	ASP258N	PHE259O
93	GLN245N	SER244OG	152	THR256N	VAL250O	210	PHE259N	ARG257NH1
94	GLN245N	GLN245OE1	153	THR256N	GLY251O	211	PHE259N	ARG257O
95	GLN245N	GLN245NE2	154	THR256N	TRP253O	212	PHE259N	ASP258OD1
96	GLN245N	THR256OG1	155	THR256N	ASP258O	213	PHE259N	ASP258OD2
97	GLN245NE2	ILE242O	156	THR256OG1	ASN241O	214	VAL260N	ASN241OD1
98	GLN245NE2	LEU243O	157	THR256OG1	LEU243O	215	VAL260N	LEU254O
99	GLN245NE2	SER244O	158	THR256OG1	GLN245O	216	ALA261N	PHE259O
100	GLN245NE2	GLN245O	159	THR256OG1	VAL247O	217	CYS262N	ASN241OD1
101	GLN245NE2	ILE246O	160	THR256OG1	ASP248OD1	218	CYS262N	VAL260O
102	GLN245NE2	TRP253NE1	161	THR256OG1	ASP248OD2	219	ASN263N	ASN263ND2
103	ILE246N	SER244OG	162	THR256OG1	VAL250O	220	ASN263ND2	CYS238O
104	ILE246N	SER244O	163	THR256OG1	GLY251O	221	HIS264N	ALA261O
105	ILE246N	GLN245OE1	164	THR256OG1	ARG257N	222	HIS264N	HIS264ND1
106	ILE246N	GLN245NE2	165	THR256OG1	ASP258O	223	LEU265N	ALA261O
107	VAL247N	SER244O	166	ARG257N	SER244OG	224	LEU265N	CYS262O
108	VAL247N	GLN245O	167	ARG257N	ASP248OD1	225	LEU265N	HIS264ND1
109	VAL247N	ILE252O	168	ARG257N	GLY251O	226	ARG266N	CYS262O
110	ASP248N	ILE246O	169	ARG257N	TRP253O	227	ARG266N	ASN263O
111	ASP248N	GLY251O	170	ARG257N	GLY255O	228	ARG266N	HIS264O
112	ASP248N	ILE252O	171	ARG257N	ASP258O	229	ARG266NE	ASN263OD1
113	ILE249N	VAL247O	172	ARG257NE	ASN241OD1	230	ARG266NH1	CYS262O
114	ILE249N	ASP248OD1	173	ARG257NE	SER244OG	231	ARG266NH1	ASN263OD1
115	ILE249N	ASP248OD2	174	ARG257NE	VAL247O	232	ARG266NH2	CYS262O
116	ILE249N	THR256OG1	175	ARG257NE	ASP248OD1	233	SER267N	ASN263O
117	VAL250N	ASP248OD1	176	ARG257NE	ASP248OD2	234	SER267N	HIS264O
118	VAL250N	ASP248OD2	177	ARG257NE	THR256OG1	235	SER267OG	ASN263O
119	VAL250N	ASP248O	178	ARG257NE	THR256O	236	SER267OG	HIS264O
120	VAL250N	THR256OG1	179	ARG257NE	ASP258OD1	237	TYR268N	HIS264O
121	GLY251N	ASP248OD1	180	ARG257NE	ASP258O			
122	GLY251N	ASP248OD2	181	ARG257NE	PHE259O			
123	GLY251N	ASP248O	182	ARG257NH1	ASN241OD1			
124	GLY251N	ILE249O	183	ARG257NH1	LEU243O			
125	ILE252N	VAL247O	184	ARG257NH1	SER244OG			
126	ILE252N	ASP248OD1	185	ARG257NH1	ILE246O			
127	ILE252N	ASP248OD2	186	ARG257NH1	VAL247O			
128	ILE252N	ASP248O	187	ARG257NH1	ASP248OD1			
129	ILE252N	ILE249O	188	ARG257NH1	TRP253O			
130	ILE252N	LEU254O	189	ARG257NH1	THR256OG1			
131	TRP253N	ILE249O						

Appendix A3

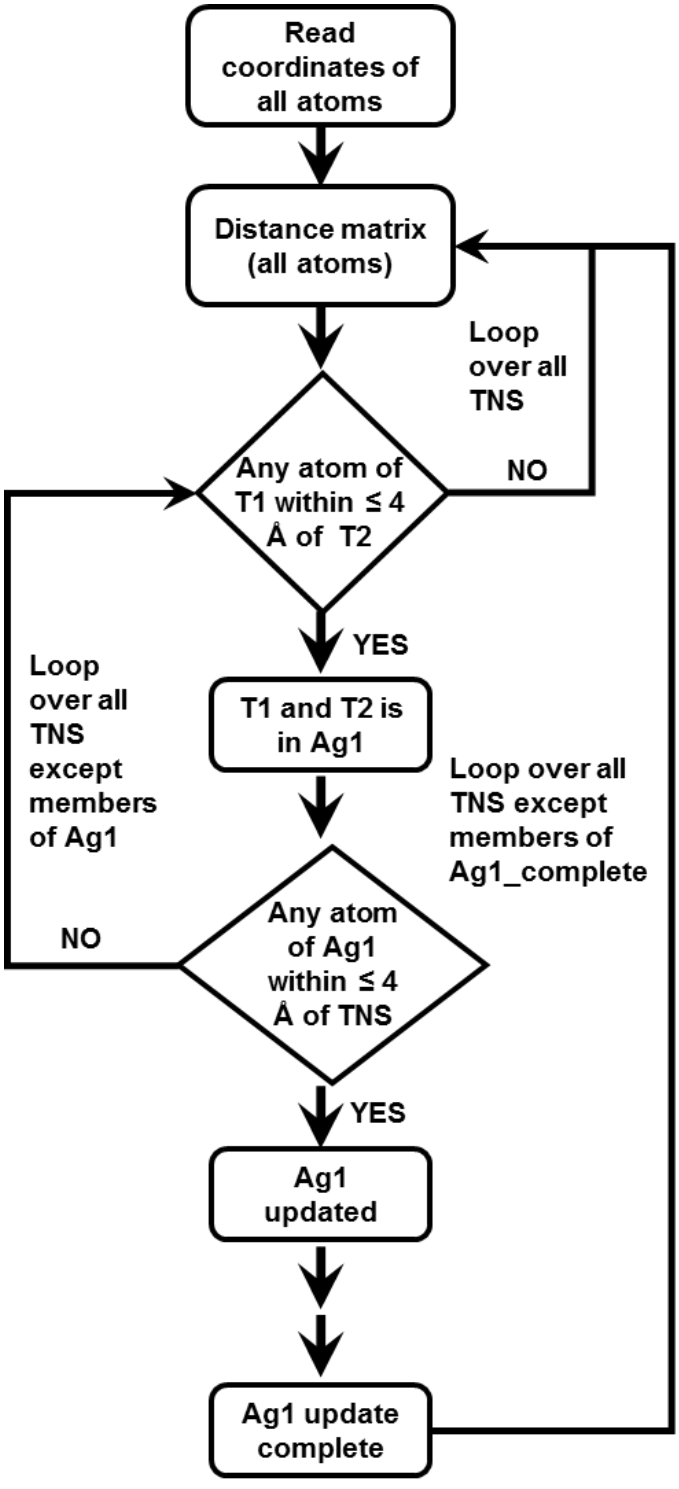
	Donor	Acceptor		Donor	Acceptor		Donor	Acceptor
1	LYS81NZ	ASP258OD1	61	GLN245N	GLN245OE1	122	ARG257N	LEU254O
2	LYS81NZ	ASP258OD2	62	GLN245NE2	LEU243O	123	ARG257N	GLY255N
3	LYS81NZ	PHE259O	63	GLN245NE2	SER244O	124	ARG257N	GLY255O
4	TYR115N	TRP253NE1	64	GLN245NE2	GLN245N	125	ARG257NE	GLN239OE1
5	TYR115N	TRP253O	65	GLN245NE2	GLN245O	126	ARG257NE	ILE246O
6	TYR115OH	ILE249O	66	GLN245NE2	ILE246O	127	ARG257NE	ASP258N
7	TYR115OH	TRP253NE1	67	ILE246N	SER244OG	128	ARG257NE	ASP258OD1
8	TYR115OH	ARG257O	68	ILE246N	SER244O	129	ARG257NE	ASP258OD2
9	HIS152NE2	HIS264ND1	69	ILE246N	GLN245OE1	130	ARG257NE	ASP258O
10	HIS152NE2	HIS264O	70	ILE246N	GLN245NE2	131	ARG257NH1	GLN239OE1
11	SER153OG	HIS264ND1	71	VAL247N	SER244O	132	ARG257NH1	SER244OG
12	SER153OG	HIS264NE2	72	VAL247N	GLN245O	133	ARG257NH1	ILE246O
13	SER153OG	HIS264O	73	VAL247N	ASP248O	134	ARG257NH1	ASP258OD1
14	THR205OG1	ASN263O	74	ASP248N	ILE246O	135	ARG257NH1	ASP258OD2
15	THR205OG1	HIS264O	75	ASP248N	VAL250O	136	ARG257NH2	GLN239OE1
16	ASN213ND2	ASN241OD1	76	ILE249N	VAL247O	137	ARG257NH2	GLN239NE2
17	ASN213ND2	ASN241ND2	77	ILE249N	ASP248OD1	138	ARG257NH2	SER244OG
18	ASN213ND2	ASN241O	78	ILE249N	ASP248OD2	139	ARG257NH2	ILE246O
19	ASN213ND2	ILE242N	79	VAL250N	ASP248OD1	140	ARG257NH2	ASP258OD1
20	ASN213ND2	ILE242O	80	VAL250N	ASP248OD2	141	ARG257NH2	ASP258OD2
21	ASN213ND2	SER244N	81	VAL250N	ASP248O	142	ARG257NH2	ASP258O
22	ASN213ND2	SER244OG	82	GLY251N	ASP248OD1	143	ASP258N	ARG257NH1
23	LEU214N	SER244OG	83	GLY251N	ASP248OD2	144	ASP258N	ARG257NH2
24	LEU214N	SER244O	84	GLY251N	ASP248O	145	ASP258N	ASP258OD1
25	LYS233NZ	ASN241O	85	GLY251N	ILE249O	146	ASP258N	ASP258OD2
26	LYS233NZ	ILE242O	86	GLY251N	TRP253O	147	PHE259N	ILE252O
27	GLN234NE2	ASN241O	87	ILE252N	ASP248OD1	148	PHE259N	TRP253NE1
28	GLN234NE2	ILE242O	88	ILE252N	ASP248OD2	149	PHE259N	ASP258OD1
29	MET235N	ASN263OD1	89	ILE252N	ASP248O	150	VAL260N	GLN239OE1
30	GLY237N	ASN263OD1	90	ILE252N	ILE249O	151	VAL260N	ASN241OD1
31	CYS238N	ASN263OD1	91	ILE252N	TRP253O	152	VAL260N	ASP258OD1
32	GLN239N	GLN239OE1	92	TRP253N	ASP248OD1	153	VAL260N	ASP258O
33	GLN239N	ASN263OD1	93	TRP253N	ASP248OD2	154	ALA261N	ASN241OD1
34	GLN239NE2	CYS238O	94	TRP253N	ILE249O	155	ALA261N	PHE259O
35	GLN239NE2	LYS240O	95	TRP253N	VAL250O	156	CYS262N	VAL260O
36	GLN239NE2	ASP258OD1	96	TRP253N	GLY251N	157	ASN263N	VAL260O
37	GLN239NE2	ASP258O	97	TRP253N	GLY251O	158	ASN263N	ALA261O
38	GLN239NE2	VAL260O	98	TRP253N	THR256OG1	159	ASN263N	ASN263OD1
39	LYS240N	GLN239OE1	99	TRP253NE1	GLY255O	160	ASN263ND2	CYS238O
40	LYS240N	GLN239NE2	100	TRP253NE1	ARG257O	161	ASN263ND2	GLN239O
41	LYS240NZ	CYS238O	101	LEU254N	VAL250O	162	HIS264N	ALA261O
42	ASN241N	ASN263OD1	102	LEU254N	GLY251O	163	HIS264N	ASN263OD1
43	ASN241N	ASN263ND2	103	LEU254N	ILE252O	164	LEU265N	ALA261O
44	ASN241ND2	VAL260O	104	LEU254N	TRP253NE1	165	LEU265N	CYS262O
45	ASN241ND2	ASN263ND2	105	LEU254N	ARG257O	166	ARG266N	CYS262O
46	ILE242N	LYS240O	106	GLY255N	GLY251O	167	ARG266N	ASN263O
47	ILE242N	ASN241OD1	107	GLY255N	ILE252O	168	ARG266N	HIS264O
48	LEU243N	ASN241O	108	GLY255N	TRP253NE1	169	ARG266NH1	CYS262O
49	SER244N	ASN241OD1	109	GLY255N	TRP253O	170	SER267N	HIS264O
50	SER244N	ILE242O	110	GLY255N	ARG257O	171	SER267OG	HIS264O
51	SER244N	GLN245OE1	111	THR256N	GLY251O	172	TYR270OH	ASN263OD1
52	SER244OG	ASN241OD1	112	THR256N	TRP253O	173	TYR270OH	ASN263ND2
53	SER244OG	ILE242O	113	THR256N	LEU254O	174	TYR270OH	ASN263O
54	SER244OG	LEU243O	114	THR256OG1	ASP248OD1			
55	SER244OG	SER244O	115	THR256OG1	ASP248OD2			
56	SER244OG	GLN245OE1	116	THR256OG1	VAL250O			
57	SER244OG	GLN245O	117	THR256OG1	GLY251O			
58	SER244OG	ILE246O	118	THR256OG1	TRP253O			
59	GLN245N	LEU243O	119	THR256OG1	THR256O			
60	GLN245N	SER244OG	120	THR256OG1	ARG257N			
			121	ARG257N	GLY251O			

Appendix B1

Characterization of aggregates

Aggregates were calculated using the distance coordinates between every atom in a particular time frame during simulation.

- A distance matrix consisting of all the atoms was created.
- All the atoms of each TNS (T1) was compared against all the atoms of another TNS (T2) molecule. If the distance was ≤ 4 Å, the molecules were considered to be forming an aggregate (Ag1).
- Then, Ag1 was checked against all the atoms of another TNS, excluding T1 and T2. If the distance condition (≤ 4 Å) was met, the Ag1 was updated.
- If Ag1 cannot be updated anymore, it means that the aggregate is complete.
- The steps 2 and 3 were repeatedly performed on all TNS molecules, excluding the members of Ag1. The new aggregates were numbered as Ag2, Ag3 and so on.



Appendix B2

Number of Hydrogen bonds:

- Last 10 ns frame was used for the hydrogen bond analysis.
- The hydrogen bonds were calculated with the help of the g_hbond tool of the GROMACS.
- All the hydrogen bonds formed by TNS molecule with protein, TNS and solvents were considered.
- R-code was written to analyze the g_hbond output file.
- Proper sorting and counting was implemented.
- Care was taken not to over-count the hydrogen bond number, e.g all the three oxygen of TNS showing hydrogen bond with an amino acid atom was counted as one.

Appendix B3

Order of contacts and amino acid specificity

- Order of contacts was calculated for last 10 ns at an interval of every100 ps (i.e., for 101 frames).
- For each TNS, list of atoms in contact with other components of the system (protein, solvent and other TNS molecules) was obtained using Gromacs tool for every chosen time frame. The file consisted of all the possible interactions of all the TNS molecules.
- In-house R-code was written to dissect each type of interactions, as given below.
- For 101 frames, with 50 TNS molecules (or 64 TNS in the cases of solvent-only simulations) in each frame, the interactions were classified into TNS-protein, TNS-TNS and TNS-solvent, and the number of interactions were counted.
- The number of each interactions were divided by 101 to calculate the average of interactions/frame.
- These values were further divided by 50 (or 64 in the cases of solvent-only simulations) to evaluate the fraction of TNS molecules interacting with protein or with solvent or with other TNS molecules.
- The number of TNS-protein contacts was in the range of 1 to 50 and TNS-TNS & TNS-solvent contacts were in the range of 1 to 90
- For TNS-Protein contacts, three categories were assigned such that zero, lower and higher order for 0, 1-20 and above 21 contacts, respectively.
- For TNS-TNS and TNS-solvents contacts, three categories were assigned such that zero, lower and higher order for 0, 1-50 and above 51 contacts, respectively.
- From TNS-Protein contact information, the individual amino acids interacting with TNS were counted. The amino acid-TNS contacts were classified into TNS-side chain and TNS-main-chain contacts based on atom labels.
- The counted contacts were normalized with respect to the number of each amino acid residues available in the protein.

List of Publications:

[1] N. Haque, N. Prabhu, Insights into protein–TNS (2-p-toluidinylnaphthalene-6-sulfonate) interaction using molecular dynamics simulation, Journal of Molecular Structure, 1068 (2014) 261-269.

[2] N. Haque, N.P. Prabhu, Lid dynamics of porcine pancreatic lipase in non-aqueous solvents, Biochimica et Biophysica Acta (BBA) - General Subjects, 1860 (2016) 2326-2334.

[3] N. Haque, N.P. Prabhu, Lid closure dynamics of porcine pancreatic lipase in aqueous solution, Biochimica et Biophysica Acta (BBA) - General Subjects, 1860 (2016) 2313-2325.

[4] E.K. Kumar, N. Haque, N.P. Prabhu, Kinetics of Protein Fibril Formation: Methods and Mechanisms, International Journal of Biological Macromolecules, (2016).

[5] N. Haque, N.P. Prabhu, Role of Bile Salt in the Formation and Stabilization of its Ternary Complex with Pancreatic Lipase and Colipase (under preparation)

[6] N. Haque, K. K. Baratam, N.P. Prabhu, Binding Modes of Fluorescence Probe TNS with Different Conformational States of Proteins: A Computational Study (under preparation)