Understanding of Self-assembly of Polypeptide-based Molecules and Bolaamphiphile from Atomistic and Mesoscale Simulations

Thesis Submitted to AcSIR For the Award of
the Degree of
DOCTOR OF PHILOSOPHY
In Chemical Sciences

By
Sujit Sarkar
Registration Number: 10CC11A26014

Under the guidance of

Dr. Kumar Vanka
(Research Supervisor)

&

Dr. Sudip Roy
(Research Co-supervisor)

CSIR-National Chemical Laboratory

March 2017

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

Certificate

This is to certify that the work incorporated in this Ph.D. thesis entitled **"Understanding of Self-assembly of Polypeptide-based Molecules and Bolaamphiphile from Atomistic and Mesoscale Simulations"** submitted by Mr. Sujit Sarkar to Academy of Scientific and Innovative Research (AcSIR) in fulfilment of the requirements for the award of the degree of **Doctor of Philosophy in Chemical Sciences**, embodies original research work under our guidance. We further certify that this work has not been submitted to any other University or Institutions in part or full for the award of any degree or diploma. Research material obtained from other sources has been duly acknowledged in the thesis. Any text, illustration, table, etc., used in the thesis from other sources, have been cited and acknowledged.

Signed:
Sujit Sarkar
Sujit Sarkar
(Research Fellow)
Physical and Material Chemistry,
CSIR- National Chemical Laboratory

Date : 23.03.2017

Signed:
Kumar Vanka
Dr. Kumar Vanka
(Research Supervisor)
Physical and Material Chemistry,
CSIR- National Chemical Laboratory

Date : 23.03.2017

Signed:
Sudip Roy
Dr. Sudip Roy
(Research Co-Supervisor)
Senior Researcher, Shell Technology, Bangalore

Date : 23.03.2017

Communications Channels
☎
NCL Level DID : 2590
NCL Board No. : +91-20-25902000
Four PRI Lines : +91-20-25902000
FAX
Director's Office : +91-20-25902601
COA's Office : +91-20-25902660
SPO's Office : +91 20 25902664
WEBSITE
www.ncl-india.org

DECLARATION

I hereby declare that this thesis entitled, "**Understanding of Self-assembly of Polypeptide-based Molecules and Bolaamphiphile from Atomistic and Mesoscale Simulations**" submitted for the degree of Doctor of Philosophy in Chemical Sciences to the Academy of Scientific & Innovative Research (AcSIR), has been carried out by me at the Physical and Materials Chemistry Division of CSIR-National Chemical Laboratory, Pune under the supervision of Dr. Kumar Vanka and Dr. Sudip Roy. The materials obtained from other sources have been duly acknowledged in the thesis. This work is original and has not been submitted in part or full by me for any other degree or diploma to this or to any other Institution or University.

Date: 23.03.2017

Place: CSIR- NCL, Pune

Sujit Sarkar

Sujit Sarkar

Ph. D. candidate

To

my Family

Acknowledgements

I am thankful to my research supervisor, Dr. Kumar Vanka for providing me an opportunity to complete my Ph.D. carrier and offering valuable suggestions and encouragement in stressful times. He has always given me valuable suggestions throughout my Ph. D. Tenure. I would like to express my sincere gratitude to my research co-supervisor Dr. Sudip Roy for being with me at most difficult times of my Ph. D tenure. He has taught me about the basics of subject of my Ph. D work. His constant encouragement and faith upon me helped me to evolve as a better researcher. All the valuable discussions with him regarding the work helped me to understand the area of my research in a better way. I would like to thank Dr. Suman Chakrabarty for learning a lot from his guidance during a collaborative research work with him. I would like to express my sincere gratitude to Dr. Rajnish Kumar for his continuous support in all aspects and also for providing me fellowship for a certain period of time.

I wish to express my sincere thanks to the Doctoral Advisory Committee members Dr. Nilanjana Sengupta, Dr. Nayana Vaval, and Dr. Rajnish Kumar for their suggestions about my research work and their valuable time during meetings.

I owe special thanks to all of my lab friends for providing a wonderful work environment throughout my Ph. D tenure. I find myself lucky for getting opportunity to work with Dr. Prithvi Raj Pandey. I have learned so many things from him about basics of my research and he always used to be my mentor. I would like to thank Santu Biswas and Dr. Souvik Chakraborty for having collaborative work with them, and learning a lot of things by useful discussions with them. I thank Dr. Chandan Kumar Choudhury for learning a lot scientific as well technical knowledge from him. I would like to thank Dr. Swagata Pahri and Dr. Anil Mhashal for always being helpful to me and giving suggestions to solve research problems. I want to thank Dr. Subhadip Das for having useful discussions regarding work. It has been wonderful experience to work with Mr. Nilesh Choudhury, Ms. Pragati Sharma, and Mr. Bappa Ghosh. Their encouragement and support always helped me to do better research.

I want to heartily thank my friends Dr. Achiyanta Dutta, and Dr. Susanta Das, and Dr. Himadri Pathak for being always with me in difficult times and for their valuable

suggestions. Life in CSIR-NCL have I would like to thank Dr. Arpan Manna, Dr. Sumantra Bhattacharya, and Dr. Kanak Roy for spending wonderful time with them.

It has been a pleasant journey with all my juniors, specially, Ujjwal Nandi, Sayantan Acharya, and Bipul Biswas. We had spent a lot of good times together. The discussion with them always helped me to learn new things. I would specially thank to Sudip Sasmal for always getting prompt help from him in all manner also during my thesis submission which made easier to accomplish different official formalities. I want to thank my roommate Mr. Soumyajyoti Chatterjee for having very good time throughout so many years.

I want to thank Santanu Pattanayak, Suman Chandra, Atreyee Banerjee, Manoj Nandi, Atanu Patra, Bittu Chandra, Chayanika Das, Manik C. Sil, Monalisa Gangopadhyay for having good time in NCL New Hostel with them. I want to heartily thank our hostel cooks for helping me to enjoy a hassle free life outside research lab.

At last I want to thank my mother, elder sister, and elder brother for always been extraordinary supportive during my Ph. D.

Contents

Declaration i

Acknowledgements iii

Abstract ix

List of Publications xi

1. Introduction ……. 1

2. Computational Methods …… 13

2.1. Classical all-atomistic molecular dynamics simulation ……. 14

2.1.1. Force Field ……. 15

2.1.2. Equation of motion integrator …… 17

2.1.3. Periodic boundary Condition …… 18

2.1.4. Energy minimization …….. 19

2.1.4.a. Steepest Descent …… 19

2.1.4.b. Conjugate Gradient Method ……. 20

2.1.5. Statistical Ensemble …… 20

2.1.6. Thermostat …… 21

2.1.7. Barostat …… 22

2.1.8. Free Energy ……. 23

2.2. Dissipative particle dynamics simulation ……. 24

3. Propensity of Self-assembled Leucine-Lysine Diblock Co-polymeric α-Helical Peptides to Remain in Parallel and Anti-parallel Alignments in Water …….. 29

3.1. Introduction …….. 30

3.2. Computational Methods …… 33

3.2.1. NPT Run of 200 ns …… 34

3.2.2. Calculation of end-to-end distance auto-correlation function … 35

3.2.3. Simulated Annealing Steps …… 37

3.2.4. NPT Run of 100 ns …….. 38

3.2.5. Calculation of helical order parameters ……. 39

3.2.6. Free Energy calculation ……. 41

3.3. Result & Discussions …… 43

3.3.1. Helicity of Peptide Chains ……. 43

3.3.2. Water Penetration in Lamellar Structure ……..... 44

3.3.3. Structural Arrangements …… 46

3.3.4. Origin of Stability of Lamellar Structures ……..... 50

3.3.5. Hydrogen Bonded Network ……. 53

3.3.6. Quantification of Stability: PMF Calculations …… 56

3.4. Conclusions ……. 59

4. Transferability of Different Classical Force Fields for Right and Left Handed α-Helices Constructed from Enantiomeric Amino Acids 64

4.1. Introduction …….. 65

4.2. Computational Details …… 69

4.2.1. QM optimization of helical peptides …… 69

4.2.2. Details of Classical MD Simulations …… 70

4.3. Result & Discussions ……... 73

4.3.1. Relative Energies of the Helical Conformers from QM Optimization …….. 73

4.3.2. Structural Properties of Helices - Comparison Between QM Optimized and Classical MD Structures …… …. 78

4.3.2.a. Comparison depending upon Φ and Ψ dihedral angles and i-i+4 H-Bonding ……... 79

4.3.2.b. Comparison depending upon helical order parameter …. 89

4.3.3. Quantification of Interaction between Side Chain and Main Chain …….. 97

4.4. Conclusions …….. 99

5. Structural integrity of bilayers formed by glycopolypeptide-polypropylene oxide di-block copolymers: Effect of chiral heterogeneity in glycopolypeptide …… 106

5.1. Introduction ……. 107

5.2. Computational Details …… 109

5.2.1. Simulation of single PPO_7 chain in water …… 110

5.2.2. Simulations of PPO_7 melt system …….. 111

5.2.3. Preparation of the systems
5.2.3.a. Glycopeptides block consisting of l-amino acids …….. 112
5.2.3.b. Glycopeptides block consisting of alternate d, l-amino acid …….. 113

5.3. Result & Discussions …… 113

5.4. Conclusion …….. 121

6. Phase Diagram of Self-assembled Sophorolipid Morphologies from Mesoscale Simulations …… 127

6.1. Introduction …….. 128

6.2. Computational Methods ……. 131

6.2.1. Theoretical Background of DPD …… 131

6.2.2. Calculation of interaction parameters between DPD beads 132

6.2.3. Structure factor Calculation ……. 136

6.3. Result & Discussions …… 137

6.3.1. Morphologies of self-assembled structure …… 138

6.3.2. Arrangement of hydrophobic and hydrophilic groups of sophorolipids ……. 141

6.3.3. Orientation of chains and their flips ……. 145

6.4. Conclusions ……. 152

7. Conclusions ……. 159

Abstract

In this thesis, we have studied the reasons behind the self-assembly of polypeptide based molecules and bolaamphiphile from atomistic and mesoscale simulation, respectively. The self-assembly of monolayer consisting of α-helical diblocks co-polypeptide is mainly understood by performing all-atomistic molecular dynamics simulation. Bilayer consisting of self-assembled diblock chains of hydrophilic α-helical poly-glycopeptides and hydrophobic poly-propylene oxide is envisaged in details by molecular dynamics simulation. The self-assembly of bolaamphiphile has been studied by mesoscopic coarse-graining simulation.

Chapter 1: The aim of this thesis is to explore the reasons behind the self-assembly of polypeptide-based α-helical diblock chains and to obtain structural insight of these self-assembled morphologies. Molecular dynamics simulation is very useful to obtain those detailed information. The examples of different polypeptide based self-assemblies have been discussed in this chapter. Self-assembly of different diblock chains consisting of hydrophilic glycopeptide and hydrophobic segment and the utility of those self-assembled morphologies have been described in this chapter. Bolaamphiphiles containing hydrophilic head groups at the terminals of a long hydrophobic tail self assemble into different structures in water. The self-assembly of different such systems has been described and the necessity of performing dissipative dynamics simulation have been discussed.

Chapter 2: This chapter presents the basic theoretical background molecular dynamics simulation. The different energy minimization methods have also been discussed. The fundamentals of Umbrella Sampling technique have been discussed for free energy calculation. For coarse-grain mesoscale simulations, theoretical aspects of Dissipative Particle Dynamics simulation are given.

Chapter 3: Molecular dynamics simulation study of α-helical diblock co-polypeptides pre-assembled in parallel and anti-parallel alignments in water is presented. The structural basis of the stability which gives favourable interactions within these lamellar structures has been investigated. We obtained a structural insight of self-assembled di-block co-polypeptides along with interactions, which are favouring the stability of the structure. Further, we have

carried out hydrogen bond analysis and free energy of binding calculation to identify the reason behind the integrity stability of these self-assembled structures.

Chapter 4: In this chapter, the performance of various force fields to simulate naturally less occurring amino acid enantiomers is discussed. The quantum mechanically optimized conformers of four different types of amino acids residues have been used as benchmark to validate molecular dynamics simulation results using various force fields.

Chapter 5: In this chapter, the integrity of two bilayers has been compared, consisting of diblocks polyglycopeptide-polypropylene oxide chains. The built of hydrophilic α-helical glycopeptide blocks are different in these two bilayers, one consist of purely l amino acids, and another by alternate arrangement of d, l-amino acids. All atomistic molecular dynamics simulations of these two bilayers are performed to find the factors which keeps integrity of one bilayer between two types of bilayer.

Chapter 6: In this chapter, dissipative particle dynamics simulation is used as a tool to investigate the morphologies of self-assembled structures formed by bolaamphiphiles in water at various concentrations. To perform DPD simulation the different functional groups of bolaamphiphile have been mapped into beads and the interaction parameters are obtained from atomistic simulation. We have analyzed the arrangement of the bolaamphiphile chains in the self-assembled structure and their dynamics inside confined shell as well.

List of Publications

1. **Sujit Sarkar**, Prithvi Raj Pandey, and S. Roy*, Propensity of Self-assembled Leucine-Lysine Diblock Co-polymeric α-helical Peptides to Remain in Parallel and Anti-parallel Alignments in Water, *J. Phys. Chem. B*, 2015, **119**, pp 9520–9531.

2. Santu Biswas[†], **Sujit Sarkar**[†], Prithvi Raj Pandey, S. Roy*, Transferability of Different Classical Force Fields for Right and Left Handed α-helices Constructed from Enantiomeric Amino Acids *Phys. Chem. Chem. Phys.*, 2016, **18**, pp 5550-5563.

3. **Sujit Sarkar***, Souvik Chakraborty and Sudip Roy*, Phase Diagram of Self-assembled Sophorolipid Morphologies from Mesoscale Simulations, *communicated*

4. **Sujit Sarkar**, Sudip Roy, Suman Chakrabarty*, Structural integrity of bilayers formed by glycopolypeptide-polypropylene oxide di-block copolymers: Effect of chiral heterogeneity in glycopolypeptide, *manuscript under preparation.*

Chapter 1

Introduction

Self-assembly is a mechanism by which unorganised materials get organized through different local interactions to form a structure of particular shape. It is a very common and important phenomenon which is found in almost every living system. The cell itself is an example which formed through the complex process of self-assembly. There are wide varieties of materials like surfactant, co-polypeptide, lipid bilayers which can undergo self-assembly to form structures of various types. One of the common features among all these building block of self-assembly is that they comprise of two or more components which differ in hydrophobicity. The mismatch in their hydrophobicity causes segregation of these components from each other, and same component likely to be in vicinity. It results into the arrangement of these building blocks into a particular pattern.

Water is most abundant in living organism. Therefore, the interactions between water and the other components of cells are very important in biological chemistry. Water being a polar molecule, can form hydrogen bond with each other or with other protic polar molecules. In contrast, nonpolar molecules do not interact with water, which makes it poorly soluble in water. As a result, non-polar molecules try to decrease their contact area with water, and they form aggregate with each other. Thus such interactions between polar and non-polar molecules and with water play vital roles in the formation of biological structures.

The concept of hydrophobicity has been a subject of much study in all fields of science, specially biology and chemistry. It was first introduced by Walter Kazumann for explaining protein denaturation.[1] There is twenty different amino acids which are abundant in nature. Their characteristics are controlled by their side chain which introduces a unique role in a protein structure. Amino acid residues are classified as hydrophobic (low propensity to be in contact with water), polar or charged (energetically favourable to be in contact with water) depending upon the tendencies of the side chain to interact with polar solvent like water. Hydrophobic effect is represented by the relative hydrophobicity of amino acid side chains and their interaction guides the three dimensional structure and stability of protein. Hydrophobicity is a phenomenon which depends upon many factors, like size of molecules, its polarity, etc.[2] The segregation of hydrophobic and hydrophilic phases or components creates in a molecular interface that extends over lengths that are longer compared to the distances over which molecules interact with one another in a homogeneous liquid. These interfaces distinguish among mesoscopic structures, such as micelles, bilayers or micro-emulsions. Hydrophobic molecules interact through week van der Waals weak attractive forces with water molecules.

The hydrophobicity is measured from the transfer of energy of any solute from water to any non-polar solvent like n-octanol. Cohn and Edsall[3] have shown that the free energies of transfer increase with the number of –CH group.[4] They have measured the free energies of transfer of the amino acids from water to ethanol, and the data used to provide the numerical values in one of the first scales of hydrophobicity. Free energy of transfer has been plotted as a function of surface area accessed by solvent, and two straight lines found from fitting these points. Hydrophobic amino acids and hydrophilic amino acids are fitted in different lines. The free energy of transfer from polar to non-polar solvent is found to be linearly proportional with the surface area. The accessible surface increases with the length of side chains of amino acids. Thus, free energy of transfer is found to be more for amino acids having longer chain length.

Proteins are natural polymers consisting of α-amino carboxylic acids. There are 20 naturally occurring amino acids. In a protein, the amino-acid residues are linked together by amide linkage (–OC–NH–) which are also called peptide linkages. Rotation about the C–N bond is restricted because of double bond nature of C-N bond, therefore the amide group is planar. The flexibility of protein chains originates from the rotation about the asymmetric carbon atom around the amino-acid residue. Biopolymers such as proteins have a hierarchical structure. The primary structure is simply the sequence of amino-acid residues in the polymer chain. The secondary structure of biopolymers results from hydrogen bonding between –CO and –NH groups. This leads to the so-called α-helix and β-sheet structures of proteins. The amino acids differ in structure by the side chains. These side chains attribute different chemical, physical and structural properties to the final peptide or protein. Depending on the side-chain substituent, an amino acid can be classified as being acidic, basic or neutral.

Peptides have distinct characteristic local structural conformations of its strand depending upon the secondary structures which formed by different kinds hydrogen bonding. Two types of secondary structures mainly exist, the α-helix and the ß-sheet. The *α-helix* is a coiled strand where every 3.6 residues complete one turn. The side-chains of the amino acid groups in α-helices project to the outside. There is formation of hydrogen bonds between the oxygen of the C=O group of one protein residue and the hydrogen of the N-H group of the fourth amino acids in the α-helix. The inter residue hydrogen bonds gives stability to the α-helix conformation. Polypeptides are usually α-helical when it uncharged, and have low solubility. After ionization, it adopts the random coil conformation and it dissolves well in water. In presence of strong acid like tri-fluoro acetic acid, the α-helical structure gets perturbed, inter-

molecular hydrogen bonds ruptured, and conformation of polypeptide becomes random coil.[5-7] In the β-sheet structures, the polypeptide backbone is almost fully stretched. This allows the peptide NH and C=O to point out at right angles to the line of the backbone. NH and C=O groups alternate along each edge. The hydrogen bond is formed in between two strands (inter-strand) rather than within strand (intra-strand).

In last few years, a lot of attention has been given to polypeptide based self-assemblies and different kinds of aggregates with complex structures have been discovered.[8, 9] In these self-assembled structures, polypeptide chains can serve as either hydrophilic or hydrophobic blocks. The excellent biocompatibility and biodegradability properties of polypeptide copolymers and their self-assemblies exert important applications in bio-related fields, such as tissue engineering, and drug delivery.[10] Peptide building blocks have been used for the self-assembly to organic nanotube by Ghadiri and co-workers.[11, 12] They have designed a new class of material using cyclic polypeptides by alternatively arranging even number of D- and L- amino acids, which interact through hydrogen bonding and form an array of self-assembled nanotube. The internal diameter of the nanotube is found to be approximately 7-8 Å, and it depends upon the number of the amino acids in the cyclic peptide. These tubular structures can be applied for various applications. They preferentially interact with bacterial membrane, enhance membrane permeability, and break transmembrane ion-potential which causes cell death.[13] Savin et al [14] observed a formation of micelles from poly (butadiene)-block-poly (L-lysine) (PB-b-PLL) block copolymers in aqueous solution. PLL remain on the shell and PB found to be at the core of micelle. Lecommandoux and coworkers[15] have synthesized and studied self-assembly of poly-(butadiene)-block-poly(L-glutamic acid) (PB-b- PLGA) block copolymers. PLGA forms shell of the self-assembled aggregates in aqueous solution. Spherical micelles and vesicles were obtained by varying composition of sample. Oleic acid and linolenic acid sophorolipids self-assemble to form different structures.[16] They have two hydrophilic head groups connected through hydrophobic tail, called bolaamphiphile. Glycolipid based bola contains larger sophorose head group and smaller -COOH head groups, separated by $–C_{15}$ alky tail.[17] The morphology of their self-assembled structure changes with degree of ionisation of –COOH group.

Molecular dynamics simulation is a tool by which structure, conformation, different physical and dynamic properties of a bulk system are represented. It gives microscopic molecular level informations with reasonable accuracy. MD simulation have wide applications in calculating

physical properties of different materials like solid, liquid, colloid, polymers, biological system, etc. MD simulation was established[18, 19] in late 70' s and it has been developed from the simulation of system of 100 atoms to system of biological importance with millions of atoms with advancement of computational power. The initial conformation of the system is taken from either experimental structures or optimized structure obtained from quantum mechanical calculation. The simulation can be performed at different length and time scale. Atomistic simulation reproduces the systems in most detailed manner. However, coarse-grain simulation is found to be very effective in producing the correct morphology. The atomistic simulation is not feasible for the very large system where long time simulation is required.

In the Chapter 3, we have used MD simulation to investigate the interaction in monolayers consisting of α-helical b-poly(leucine)-b-poly(lysine) diblocks chains. The interactions among these diblocks chains stabilising the bilayer have been analysed. We also investigated the effect when a hydrophilic block of co-polypeptide comes near a hydrophobic block during self assembly. Whether there is any possibility of hydrophilic moieties to fall within the aggregated hydrophobic moieties or *vice-versa* in the self assembled structure is investigated. Therefore, two types of monolayer have been constructed. In one, hydrophobic as well hydrophilic segments are closely arranged to themselves, and in another hydrophobic blocks are nearer to hydrophilic segment. MD simulation enables us to acquire detailed insight of these systems.

For successful performance of MD simulation, it is important to evaluate potential of the system accurately. The different forms of potential functions are used to calculate bonded and non bonded potential acting between atoms. The different functional forms and the parameters used to calculate potentials are called force field. The functional forms of force field used to calculate bonded potentials is harmonic in nature, and calculation becomes simpler and faster even for large systems. The non-bonded potential consists of Lennard-Jones 12-6 potential and coulombic interaction. Although, force fields potential functional forms are simple in nature, but its parameterization is very crucial. The accurate description of structure, conformation of the molecule which is obtained from MD simulation, depends upon the proper parameterization technique. There are varieties of available force fields like OPLS[20], AMBER[21], CHARMM[22], GROMOS[23]. These force field parameters have been developed to best produce certain kind of target physical properties, conformation etc. The success of these force fields is judged by their ability to reproduce experimentally available properties.

Generally, L-amino acids are abundant in nature. Thus, all the available force fields have been developed with respect to their performance on L-amino acids. Although L-amino acid is abundant in nature, D-amino acids also exist.[24, 25] Peptidoglycan is an important source of D-amino acid residues.[26] Interestingly, naturally occurring L-amino acids form right handed helices more preferentially, whereas D-amino acids preferentially form left-handed helices.[27] Therefore, it is very important to study the robustness of these recent force fields to explain other possible conformations. In the chapter 4, we have compared different force fields on the basis of reproducing structures of right and left-handed α-helical homo-polypeptide consisting of L and D enatiomers of four amino acids belonging to different classes.

The information on suitable force field to simulate the peptide consisting of naturally less occurring D-amino acid is essential for studying self-assembly of system containing such type of amino acid enantiomer. Carbohydrates is an important molecule which involves in various biological processes like signal transmission, cell−cell recognition, and protein folding.[28] Glycopeptide is a peptide based molecule where carbohydrate is attached to side chain of amino acid residue. The multiple –OH groups which are attached to the side chain of glucose make the glycopeptide residue hydrophilic. When this hydrophilic glycopeptides residue is attached to hydrophobic segment in a di-block chain it becomes amphiphilic. Therefore, they can self-assemble into different structure like micelles, vesicle, etc.[29] The micelles compose of hydrophobic core and hydrophilic exterior. The hydrophobic core can encapsulate hydrophobic drugs where the hydrophilic shell on the exterior makes it water-soluble. Polymersome is spherical in shape and creates a hollow inside containing water, enclosed by bilayer membrane. It can be utilized for carrying drug as they are stable in throughout the blood circulation time.[30] Polymersome of different thickness and permeability have been synthesized by changing their molecular weight and compositions.[31] Numerous studies have been done on polymersome to use it as a drug carrier for cancer therapy. It can also be modified to use for targeted drug delivery by attaching a biological reorganisation unit on the surface. Polymersome consisting of hydroxyproline rich glycoprotein block co-polymers form scaffolds by aggregation of hydrophobic peptide and hydrophilic glycopeptides.[32] Holowka et al have found that the helical conformation of a hydrophobic peptide segment like poly (γ-benzyl-L-glutamate) (PBLG) gives rise to a stable vesicle.[33]

The integrity of a bilayer depends upon the packing of its constituent's chains. The di-block chains having one of the components as α-helical poly glycopeptides have the effect on the integrity of bilayer. The chirality of the constituent amino acid plays a role in stability of bilayer through the efficient packing of the constituent chains. Das et al[34] observed that presence of an ordered helical glycopolypeptide segment in b-glycopeptides-b-polypropylene oxide chains is required for self-assembly to form polymersome in aqueous solution. Polypropylene is a FDA- approved hydrophobic material, and it makes polymer chain biocompatible when it combined with hydrophilic glycopeptides block. This encouraged us to do detailed study of bilayer formed by glycopeptides-b-poly-(propylene oxide) through MD simulation in chapter 5. The effect of chirality is envisaged by constructing glycopeptide block by purely l-amino acids in one bilayer and by alternate d, l-amino acid in another bilyer.

So far we have discussed about a particular phase of self-assembled morphology which is lamellar. The lamellar phases are formed by the amphilphilic di-block chains in water, and the ratios of di-block chain to water are maintained approximately 50% (w/w). All-atomistic molecular dynamics simulation has been very successful to probe the interaction among the di-block chains. In the last working chapter, we have advanced our study slightly further from diblocks polypeptide chains. The self-assembly of the polymeric chains having more than one components of different hydrophobicity is interesting. In the chapter 6, the self-assembly of bola-amphiphile chains is envisaged by mesoscale simulation method. Linolenic acid bolaamphiphile which is a glycolipid having sophorose disaccharide and propanoic acid head group attached to a long unsaturated alkyl tail is considered for study.

The bolaamphiphiles can spontaneously assemble to form many microstructures like vesicles, micelles or bilayers, etc depending on their molecular composition, interactions among different moieties, and most importantly on their concentration.[35] These microstructures have application in drug delivery[36], gene delivery[37], etc. Baccile et al[17] experimentally observed different self-assembled morphologies formed by sophorolipid chains in water. The spherical micelle is found at low concentration of sophorolipid, and it becomes cylindrical at higher concentration. The phase diagram of a tri-component system constituting of water, oil and surfactant is derived by Larson et al[35] applying Monte Carlo simulation. The change from spherical to cylindrical morphology and finally the lamellar structure is obtained by increasing concentration of the surfactant in a surfactant-water mixture. Although a significant amount of study has been done but still many questions remain unanswered like

how the mechanism of formation of different microstructures in phase diagram is influenced by various parameters like relative strengths of hydrophobic-hydrophilic interactions, concentration, relative size of hydrophilic and hydrophobic moieties in the molecule etc. Therefore we have studied self-assembly of bolaamphiphile with varying its concentration in water by using dissipative particle dynamics simulation. The advantage of DPD simulation is that it can simulate system over higher length and time scale. The interaction parameters for performing DPD simulation are mapped from experimental data and all-atomistic MD simulation. This study has appeared to be very useful to probe the self-assembly of bolaamphiphile chains and to find the arrangement of chains in different aggregated morphologies.

The layout of the thesis is as follows.

In chapter 2, computational methods are discussed.

In chapter 3, reason behind the self-assembly of monolayer consisting of α-helical diblock co-polypeptide is envisaged.

In chapter 4, comparative study on suitable force fields for simulating α-helical polypeptide consisting of naturally less occurring amino acids.

In chapter 5, the integrities of two bilayers is compared where the built of glycopeptide segment differs in the constituent glycopeptide-polyepropylene oxide diblocks chains.

In chapter 6, self-assembly of linolenic acid sophorolipid in water have been studied by mesoscopic simulation.

In chapter 7, conclusions from the above studies are presented.

References

[1] Kauzmann, W., in *Advances in Protein Chemistry*, edited by C.B. Anfinsen, M. L. A. K. B., and John, T. E. (Academic Press, 1959), pp. 1-63.

[2] Charton, M., and Charton, B. I., "The Structural Dependence of Amino Acid Hydrophobicity Parameters," Journal of Theoretical Biology 99, 629-644 (1982).

[3] Cohn, E. J. E., J. T., *Proteins, Amino Acids, and Peptides as Ions and Dipolar Ions* (Reinhold, New York, 1943).

[4] Richards, F., "Areas, Volumes, Packing and Protein Structure," Annu Rev Biophys Bioeng 6, 151-176 (1977).

[5] Lin, J., Liu, N., Chen, J., and Zhou, D., "Conformational Changes Coupled with the Isotropic–Anisotropic Transition Part 1. Experimental Phenomena and Theoretical Considerations," Polymer 41, 6189-6194 (2000).

[6] Lin, J., Zhou, H., and Zhou, D., "Conformational Changes Coupled with the Isotropic–Anisotropic Transition Part 2. Effect of the External Orientational Field," Polymer 42, 549-554 (2001).

[7] Lin, J., and Abe, A., "Liquid Crystal Formation Coupled with the Coil−Helix Transition in the Ternary System Poly(Γ-Benzyl L-Glutamate)/Dichloroacetic Acid/Dichloroethane," Polymer 29, 2584-2589 (1996).

[8] Schlaad, H., "Solution Properties of Polypeptide-Based Copolymers," Peptide Hybrid Polymers 53-73 (2006).

[9] Carlsen, A., and Lecommandoux, S., "Self-Assembly of Polypeptide-Based Block Copolymer Amphiphiles," Current Opinion in Colloid & Interface Science 14, 329-339 (2009).

[10] Gazit, E., "Self-Assembled Peptide Nanostructures: The Design of Molecular Building Blocks and Their Technological Utilization," Chemical Society Reviews 36, 1263-1269 (2007).

[11] Ghadiri, M. R., Granja, J. R., Milligan, R. A., McRee, D. E., and Khazanovich, N., "Self-Assembling Organic Nanotubes Based on a Cyclic Peptide Architecture," Nature 366, 324-327 (1993).

[12] Hartgerink, J. D., Granja, J. R., Milligan, R. A., and Ghadiri, M. R., "Self-Assembling Peptide Nanotubes," Journal of American Chemical Society 118, 43-50 (1996).

[13] Fernandez-Lopez, S., Kim, H.-S., Choi, E. C., Delgado, M., Granja, J. R., Khasanov, A., Kraehenbuehl, K., Long, G., Weinberger, D. A., Wilcoxen, K. M., and Ghadiri, M. R.,

"Antibacterial Agents Based on the Cyclic D,L-[Alpha]-Peptide Architecture," Nature 412, 452-455 (2001).

[14] Gebhardt, K. E., Ahn, S., Venkatachalam, G., and Savin, D. A., "Rod-Sphere Transition in Polybutadiene−Poly(L-Lysine) Block Copolymer Assemblies," Journal of American Chemical Society 23, 2851-2856 (2007).

[15] Chécot, F., Lecommandoux, S., Klok, H.-A., and Gnanou, Y., "From Supramolecular Polymersomes to Stimuli-Responsive Nano-Capsules Based on Poly(Diene-B-Peptide) Diblock Copolymers," The European Physical Journal E 10, 25-35 (2003).

[16] Dhasaiyan, P., Pandey, P. R., Visaveliya, N., Roy, S., and Prasad, B. L. V., "Vesicle Structures from Bolaamphiphilic Biosurfactants: Experimental and Molecular Dynamics Simulation Studies on the Effect of Unsaturation on Sophorolipid Self-Assemblies," Chemistry – A European Journal 20, 6246-6250 (2014).

[17] Baccile, N., Nassif, N., Malfatti, L., Van Bogaert, I. N. A., Soetaert, W., Pehau-Arnaudet, G., and Babonneau, F., "Sophorolipids: A Yeast-Derived Glycolipid as Greener Structure Directing Agents for Self-Assembled Nanomaterials," Green Chemistry 12, 1564-1567 (2010).

[18] McCammon, J. A., Gelin, B. R., and Karplus, M., "Dynamics of Folded Proteins," Nature 267, 585-590 (1977).

[19] Warshel, A., and Levitt, M., "Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme," Journal of Molecular Biology 103, 227-249 (1976).

[20] Jorgensen, W. L., and Tirado-Rives, J., "The Opls [Optimized Potentials for Liquid Simulations] Potential Functions for Proteins, Energy Minimizations for Crystals of Cyclic Peptides and Crambin," Journal of the American Chemical Society 110, 1657-1666 (1988).

[21] Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A., "A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules," Journal of the American Chemical Society 117, 5179-5197 (1995).

[22] Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M., "Charmm: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations," Journal of Computational Chemistry 4, 187-217 (1983).

[23] Schmid, N., Eichenberger, A. P., Choutko, A., Riniker, S., Winger, M., Mark, A. E., and van Gunsteren, W. F., "Definition and Testing of the Gromos Force-Field Versions 54a7 and 54b7," European Biophysics Journal 40, 843 (2011).

[24] Wade, D., Boman, A., Wåhlin, B., Drain, C. M., Andreu, D., Boman, H. G., and Merrifield, R. B., "All-D Amino Acid-Containing Channel-Forming Antibiotic Peptides," Proceedings of the National Academy of Sciences 87, 4761-4765 (1990).

[25] Oren, Z., Hong, J., and Shai, Y., "A Repertoire of Novel Antibacterial Diastereomeric Peptides with Selective Cytolytic Activity," Journal of Biological Chemistry 272, 14643-14649 (1997).

[26] Cava, F., Lam, H., de Pedro, M. A., and Waldor, M. K., "Emerging Knowledge of Regulatory Roles of D-Amino Acids in Bacteria," Cellular and Molecular Life Sciences 68, 817-831 (2011).

[27] Zawadzke, L. E., and Berg, J. M., "The Structure of a Centrosymmetric Protein Crystal," Proteins: Structure, Function, and Bioinformatics 16, 301-305 (1993).

[28] Bertozzi, C. R., and Kiessling, L. L., "Chemical Glycobiology," Science 291, 2357-2364 (2001).

[29] Pati, D., Kalva, N., Das, S., Kumaraswamy, G., Sen Gupta, S., and Ambade, A. V., "Multiple Topologies from Glycopolypeptide–Dendron Conjugate Self-Assembly: Nanorods, Micelles, and Organogels," Journal of the American Chemical Society 134, 7796-7802 (2012).

[30] Discher, B. M., Won, Y.-Y., Ege, D. S., Lee, J. C.-M., Bates, F. S., Discher, D. E., and Hammer, D. A., "Polymersomes: Tough Vesicles Made from Diblock Copolymers," Science 284, 1143-1146 (1999).

[31] Bermudez, H., Brannan, A. K., Hammer, D. A., Bates, F. S., and Discher, D. E., "Molecular Weight Dependence of Polymersome Membrane Structure, Elasticity, and Stability," Macromolecules 35, 8203-8208 (2002).

[32] Wegenhart, B., Tan, L., Held, M., Kieliszewski, M., and Chen, L., "Aggregate Structure of Hydroxyproline-Rich Glycoprotein (Hrgp) and Hrgp Assisted Dispersion of Carbon Nanotubes," Nanoscale Research Letters 1, 154-159 (2006).

[33] Holowka, E. P., Pochan, D. J., and Deming, T. J., "Charged Polypeptide Vesicles with Controllable Diameter," Journal of American Chemical Society 127, 12423-12428 (2005).

[34] Das, S., Sharma, D. K., Chakrabarty, S., Chowdhury, A., and Sen, S., "Bioactive Polymersomes Self-Assembled from Amphiphilic Ppo-Glycopolypeptides: Synthesis, Characterization, and Dual-Dye Encapsulation," Langmuir 31, 3402-3412 (2015).

[35] Larson, R. G., "Self- Assembly of Surfactant Liquid Crystalline Phases by Monte Carlo Simulation," The Journal of Chemical Physics 91, 2479-2488 (1989).

[36] Grinberg, S., Kolot, V., Linder, C., Shaubi, E., Kas'yanov, V., Deckelbaum, R. J., and Heldman, E., "Synthesis of Novel Cationic Bolaamphiphiles from Vernonia Oil and Their Aggregated Structures," Chemistry and Physics of Lipids 153, 85-97 (2008).

[37] Jain, N., Arntz, Y., Goldschmidt, V., Duportail, G., Mély, Y., and Klymchenko, A. S., "New Unsymmetrical Bolaamphiphiles: Synthesis, Assembly with DNA, and Application for Gene Delivery," Bioconjugate Chemistry 21, 2110-2118 (2010).

Chapter 2

Computational Methods

We have used computer simulation as a tool to understand the self-assembly behaviour of polypeptide based molecules and bolaamphiphile. The self-assembly of di-block co-polypeptide and glycopeptides have been understood using all-atomistic classical molecular dynamics simulation. Dissipative particle dynamics simulation which is mesoscale simulation method has been utilized to investigate the self-assembly behaviour of bolaamphiphile.

2.1 Classical all-atomistic molecular dynamics simulation

We collect the information like positions, velocities of atoms in each step by doing by doing molecular dynamics (MD) simulation. The atoms of a molecule are represented by cartesian coordinates and they have specific diameters. These atoms interact with each other through different kind of bonded and non-bonded potentials. The atoms within a molecule interact through bonded potential like bond, angle and dihedral potential. The atoms which are apart by 3 bonds interact through 1, 4 pair interaction also. The non-bonded interaction acts between two atoms which are more than 3 bonds apart and between the atoms of different molecules. The non-bonded interaction includes Lennard Jones and coulomb interactions. The atoms are considered to be interacting through the pair potential. The initial positions of atoms are obtained from the atomic coordinates of the starting structure. The initial velocities of atoms are determined from Maxwell Boltzmann distribution of velocity (Eq. 2.1) at the temperature in which MD simulation is performed.

$$f(v) = \sqrt{(m/2\pi kT)^3}\, 4\pi v^2 e^{-mv^2/2kT} \qquad (2.1)$$

The force (F_i) acting on a particle i is calculated by the derivation of potential with respect to its position and thus acceleration also obtained.

$$F_i = -\frac{\partial V(r_1, r_2, \ldots r_n)}{\partial r_i}, \quad \text{i =1...N} \qquad (2.2)$$

The positions and velocities of atoms are calculated by solving Newtonian equations of motion,

$$m_i \frac{\partial^2 r_i}{\partial t^2} = F_i\,, \quad \text{m_i is mass of the particle} \qquad (2.3)$$

We obtain position and velocities of atoms at next step by using different equation of motion integrator algorithms which are discussed in section 2.1.2.

Thus to calculate positions and velocities of atoms at each step, it's necessary to obtain interaction potential within the system accurately. There are different functional forms and parameters of bonded and non-bonded potentials which are used to calculate potential of system.

2.1.1 Force Field

Force field consists of different functional forms and set of parameters which are used to calculate the internal potential energy as a function of internal atomic coordinates. The potential energy has mainly two components- bonded and non-bonded. The bonded part comprising of four kind of potential functional forms – bonded, angle and improper dihedrals and proper dihedrals. The first three potentials among these are harmonic in nature. Proper dihedral potential is a cosine function of the dihedral angle. They contribute to the potential energies as a function of the deviation of the bond, angle and dihedral from their respective equilibrated values. Improper dihedral potential is used to prevent the planar moieties in a molecule from flipping.

$$V_{bond} = \sum_{bonds} \frac{1}{2} k_b (b - b_0)^2 + \sum_{angles} \frac{1}{2} k_\theta (\theta - \theta_0)^2 + \sum_{impropers} \frac{1}{2} k_\xi (\xi - \xi_0)^2 + \sum_{propers} (k_\phi (1 + \cos\ (n\phi - \phi_0)) \quad (2.4)$$

k_b, k_θ, k_ξ and k_ϕ are the bond, angle and improper and proper dihedral constants. Their values at time t are denoted by b, θ, ξ, ϕ, respectively. Corresponding equilibrium values are given by the same notations with subscript 0. Periodic dihedral function is defined with n multiplicity and phase value of ϕ_0.

The non-bonded interaction consists of van der Waal interaction in form Lennard-Jones (LJ) 12-6 potential and Columbic interaction.

$$V_{nb} = \sum_{atoms\,pairs} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \sum_{atoms\,pairs} \frac{q_i q_j}{4\pi\epsilon_0 \epsilon_r r_{ij}} \quad (2.5)$$

The first term of the LJ potential is repulsive interaction and varies as $1/r_{ij}^{12}$ with the inter-atomic distance (r_{ij}) between particle i and j. The second part is attractive in nature and changes less steeply ($1/r_{ij}^{6}$) with inter-atomic distance than repulsive term. ε_{ij} is the depth of potential well. The attractive part is predominant beyond a distance (~ σ_{ij}) and vanishes after ~2.5 σ_{ij} .[1] The repulsive part shows a steep wall at distance less than σ_{ij}. q_i and q_j are the partial charges of atom i and j, respectively. The absolute dielectric constant of vacuum and relative dielectric constant of medium are denoted by ϵ_0, ϵ_r, respectively. All these functional form and parameters described above are used in different force fields like OPLS[2, 3], AMBER[4], GROMOS[5], CHARMM[6]. These different force filed uses different set methodologies to derive the parameters.

The parameters for all these force fields have been optimized properly giving emphasis on proteins. The partial atomic charges of OPLS and CHARMM is based on QM calculation using HF/ 6-31G (d) while the charges in AMBER is based on fitting of RESP charges at same level of theory. Bonded parameters for both CHARMM and AMBER were derived by reproducing different experimental as well as QM data for small model compounds. It is also verified by reproducing vibrational spectra for optimization of the force constants. OPLS force field is mainly based upon liquid state simulation of organic molecules like alkanes, alkenes, alcohols, ethers, acetals, and peptides. It can well describe thermodynamic properties of liquid state[7, 8] and also protein and protein-ligand interacting systems.[9] In OPLS, parameters for bond stretching and angle bending parameters have been taken mostly from the AMBER force field. The torsional parameters are derived by fitting to torsional energy profiles of more than 50 organic molecules obtained from ab initio molecular orbital calculations at the RHF/6-31G* level of theory. The non-bonded parameters are calculated targeting liquid state properties. AMBER[10] force field is well parameterized for amino acids and nucleic acids. This force field includes parameters for mono-phosphates, thus facilitate the simulation of the nucleoside and nucleotide oligomers. It satisfactorily represents the change in conformation such as the A / B transition in duplex and triplex DNAs and have performed well in simulations of DNAs in extreme environments.

The important aspect of peptide simulation is the proper sampling of its (ψ, ϕ) backbone dihedrals. A lot of properties of protein depend upon its backbone dihedral. Generally, alanine dipeptide or tetra-peptide is used as model compound to study conformation changes

related to (ψ, ϕ) dihedral. Some researchers have found that some force fields behave well to sample dihedral angle close to QM result.[11] But, others have found that over reproduction of (ψ, ϕ) sampling data gives less accurate results for other experimentally observed properties of protein.[12] This problem has been resolved by using 2D dihedral energy grid correction map (CMAP) method with CAHRMM force field. It reproduces gas phase energy surface data of alanine dipeptide quite well.

2.1.2 Equation of motion integrator

The acting force upon each atom varies with the position of it or the other atoms with which it interacts. Therefore the motion of the all the particle is associated with the continuously changing interaction potential. This gives rise to a many-body problem which can not be solved analytically. The different algorithms have been developed to solve this. One of the methods among them is leap-frog integrator algorithm.[13]

Leapfrog algorithm

In leap-frog algorithm, acceleration is calculated at current time step from potential using Eq. 2.2. Then, the velocity at next half-time step is calculated using the acceleration at current time step and velocity at previous half-time step and it is given by the following equation (2.6).

$$v\left(t+\tfrac{1}{2}\delta t\right) = v\left(t-\tfrac{1}{2}\delta t\right) + \delta t a(t) \tag{2.6}$$

The velocity at half-time step and the current position of atoms are used to update position at next time step (Eq. 2.7).

$$r\,(t+\delta t) = r\,(t) + \delta t v\,(t+\tfrac{1}{2}\delta t) \tag{2.7}$$

The velocity at current step is obtained by averaging two adjacent velocities at previous and next half-time steps.

$$v(t) = v\left(t-\tfrac{1}{2}\delta t\right) + v\,(t+\tfrac{1}{2}\delta t) \tag{2.8}$$

To calculate kinetic energy the velocity at current time step is required.

This above mentioned protocol gives rise to the factor that position and velocity leap over each other and position remains one step ahead over velocity. The acceleration at next step

is calculated after all these steps. The first half-time step velocity, $v(\frac{1}{2}\delta t)$ is evaluated from the value of $v(-\frac{1}{2}\delta t)$ which is derived by using Eq. 2.9.

$$v\left(0-\frac{1}{2}\delta t\right) = v(0) - \frac{1}{2}\delta t a(0) \tag{2.9}$$

$v(0)$ is determined using Maxwell Boltzmann distribution of velocity (Eq. 2.1) at the temperature of performing MD simulation, and $a(0)$ is calculated from $V(0)$ using Eq. 2.2. $V(0)$ is obtained from initial configuration of system studied.

There are some advantages of using leap-frog algorithm. It calculates the velocity in more straightforward manner, and excludes the calculation of difference of large numbers.

2.1.3 Periodic boundary Condition

The objective of MD simulation is to calculate macroscopic properties which can be determined by the simulation of a very large number (~N_A) of molecules. But, this is not computationally feasible. Thus to mimic a bulk system, we perform MD simulation with a finite number of molecules enclosed in a simulation box. But, surface artefacts may arise because of the small number of molecules present in the box. In a bulk system, a very small fraction ($N^{2/3}$) of total number of atoms present near the boundaries of the box. Now, suppose if N is order of ~10 N_A, then $10^{16}/10^{24}$, i.e., 1 in 10^8 atoms only present in surface. The recent computational power can facilitate MD simulation of near upto 10^6 atoms which means 1 in 100 atoms are in vicinity of surface. It creates a major surface artefact in the calculation.

Thus to overcome this problem, the concept of periodic boundary conditions (PBC) is introduced in MD simulation. It replicates the simulation box all over the space to make the system behaving like an infinite one. There are some set of approaches by which PBC implemented. The particles in the central simulation box are called original and their copies are called periodic images. If a particle in a central box moves in particular direction then its periodic images in side boxes move in the same directions. Hence when a particle in a central box leaves the box in a certain direction its images enters into the box from the opposite direction as if there were no boundary between the boxes. The surface effect is eliminated in this way. The number of atoms in the central box is remained conserved. The particles in the central box are only considered for calculation of potential energy, not the other images. Otherwise it would have been an infinite number. When the PBC is implemented in the

system, the atoms not just interact with the atoms in the box but also with images in the nearby boxes. Thus the number of interacting pairs becomes very large. This problem is solved by taking the criteria of minimum image convention.

2.1.4 Energy minimization

Geometry optimization is the necessary step before performing the MD simulation. The initially build conformer can have overlap in between atoms and different bad connecting sites. Therefore starting of the simulation taking this structure without performing the energy minimization may lead to crush of simulation due generation of very high potential. Now, potential energy is the multi-dimensional function of atomic coordinates. This atomic coordinates can be expressed in terms of internal coordinates also. A system with N atoms has 3N cartesian coordinates or 3N-6 internal coordinates. Therefore, the energy minimization is done by calculation of change of potential energy with respect to change of those coordinates. Conformation modification is done iteratively depending upon the algorithm chosen for system of study. The way the energy of system varies with the coordinates of the atoms, is called potential energy surface (PES). There are many points corresponding to minima energies in PES which are called local minima, and the lowest one among them is called global energy minimum. In MD simulation, we are concerned about finding these minima points. There are different commonly applied algorithms like steepest descent, conjugate gradient, etc.

2.1.4.a Steepest descent Method

This method uses energy derivative with respect to coordinates to find the minima. This process leads system through the downhill of energy surface to find its local minimum near to the starting point. The sign of the derivative of energy gives information about the direction of the minima. The negative of the energy derivative is force. Therefore, this method works by decreasing the potential energy by changing the molecular conformation in response to the force acting upon them. In steepest descent method, the minimum is obtained by choosing arbitrary step approach method in which a succession of conformations is generated iteratively according to the following equation.

$$x(k) = x(k-1) + \lambda(k)F(k) \qquad (2.10)$$

x (k) is conformation at step k, k is the number of iterations, λ(k) is the step size and F(k) is force. The first step size is chosen arbitrarily. If the first iteration decreases the potential energy then step size is increased for the second iteration. If the potential energy increases then the first chosen step size is considered to be large so that it jumps over the local minima to reach into opposite side of the slope. Then the step size is decreased by some factor (e.g. 0.5) to again perform the first iteration. Thus, the continuous adjustment of step size keeps PES close to the local minima corresponding to the initial conformation. Although, the steepest descent is a very good method in searching nearest local minima but it can't find multiple local minima of PES. However, it is an efficient method for removing steric clashes between atoms and relaxing bond, angle and dihedrals.

2.1.4.b Conjugate Gradient Method

This method has the advantage over the steepest descent method that it does not show oscillatory behaviour in the narrow valley of PES. Moreover, the previous method finds minima after a large number of steps. In this method, atomic positions are updated in a direction (v_k) computing the energy gradient (g_k) of the current atomic positions and the direction (v_{k-1}) followed in the previous step.

$$v_k = -g_k + \gamma_k v_{k-1} \qquad (2.11)$$

γ_k is a scalar quantity and given by

$$\gamma_k = \frac{g_i g_j}{g_{i-1} g_{j-1}} \qquad (2.12)$$

In conjugate gradient method, gradient at each positions are orthogonal but directions are conjugate. This method is computationally more expensive and requires more storage. But, the efficient convergence to minimum is compensates the large time required for per iteration.

2.1.5 Statistical Ensemble

The macroscopic states are defined by some very basic parameters like number of molecules (N), volume (V), energy (E), and temperature (T) of the system. The various kinds of commonly used ensembles are microcanonical ensemble (N, V, E fixed), canonical ensemble

like constant-NVT ensemble, isothermal-isobaric ensemble (NPT), and grand canonical ensemble (μVT).

MD simulation is performed maintaining these conditions. There are different algorithms which maintain constant temperature or constant pressure during MD simulation.

2.1.6 Thermostat

A several thermostats are available which allow the exchange of energy between system with which MD simulation performed and surroundings to maintain constant temperature in a realistic way. These includes week coupling thermostat like Berendsen, extended ensemble coupling like Noose-Hoover, velocity rescaling method. Velocity rescaling methods is used in this thesis and discussed in the following.

Velocity rescaling method

Temperature is the defined as the average of kinetic energy of particles of the system. Therefore, temperature can be controlled by rescaling the velocities of the particles in the system. Suppose, the velocities are rescaled by a factor α to reach target temperature T_0, then α is given by, $\alpha=\sqrt{(T_0/T)}$ By, the simple rescaling method of velocity is rescaled at a predefined frequency to keep the system at target temperature. Although it is very useful to relax the system at the target temperature, but this simple rescaling does not produce the correct canonical ensemble. Therefore, instead of forcing the system into the desired temperature, a target temperature T_t is selected with stochastic procedure. T_t is determined from equilibrium distribution of kinetic energy.

$$P(T_t)dT_t \propto K_t^{\left(\frac{N_f}{2}-1\right)} e^{-\beta T_t} dT_t \qquad (2.13)$$

The system is allowed to evolve for a single-time step using Hamiltonians equation of motion. Then kinetic energy is calculated. Kinetic energy is allowed to evolve for a single-time step time using stochastic dynamics. Finally, velocity is rescaled according to new kinetic energy.

The final equation of temperature rescaling given by,

$$dK = (K_0 - K)\frac{dt}{\tau} + 2\sqrt{\frac{KK_0}{N_f}}\frac{dW}{\sqrt{\tau}} \qquad (2.14)$$

where, K is the kinetic energy, N_f is the degrees of freedom and dW a Wiener noise, τ is time scale of thermostat.

2.1.7 Barostat

As most of the experiments are performed at constant pressure, therefore in the MD simulation has been designed to maintain the same condition. It facilitates the accurate prediction of the physical properties of the system. In NVE simulation pressure is proportional product of positions of the particles and the derivative of the potential energy function. This product $r_{ij}\, dV(r_{ij})/d\, r_{ij}$ varies more than the other physical properties with position. A macroscopic system keeps its pressure fixed by constantly changing volume. An NPT –simulation also follows the same procedure by changing the cell volume. Generally, the temperature and number of molecules are also kept fixed with the temperature in the convenient MD simulation. Therefore, the volume fluctuation can be related to isothermal compressibility by,

$$\kappa = -\frac{1}{V}\left(\frac{dV}{dP}\right)_p \qquad (2.15)$$

Berendsen Barostat

This method maintains constant pressure by coupling the system with a pressure bath. The rate of the change of pressure is proportional by the pressure difference of the system.

$$\frac{dP(t)}{dt} = \frac{1}{\tau_p}\left(P_{bath} - P(t)\right) \qquad (2.16)$$

τ_p is pressure coupling constant, $P(t)$ is the actual pressure at time t, and P_{bath} is the pressure of the bath. The pressure change is implemented by scaling the volume of box by a factor λ at each time step. Thus the coordinate of each particle is scaled by $\lambda^{1/3}$.

$$r(t + \delta t) = \lambda^{1/3} r(t) \qquad (2.17)$$

where $\lambda = \left[1 - \kappa \frac{\delta t}{\tau_p}(P - P_{bath})\right]^{1/3}$.

One needs to provide value of compressibility, coupling constant, and target pressure as input parameters. Besides, the direction of applied pressure coupling can also be mentioned depending upon the system of interest.

2.1.8 Free Energy

In MD simulation, one often needs to evaluate the macroscopic properties of a system like calculation of solvation free energy, binding free energy of a particular drug with a lipid bilayer or to explore the energetic and mechanisms of conformational change. There are different free energy calculation methods available based upon MD simulation, such as thermodynamic integration[14], free energy perturbation[15, 16], potential of mean force, and umbrella sampling.[17, 18] In this thesis, umbrella sampling method has been used.

This method was developed by Torrie and Valleau in 1970 for doing the simulations of systems with high energy barriers. Umbrella sampling technique often used to calculate difference of free energy in between two states or two conformations by slowly approaching from one state to another through a definite pathway. This pathway is called reaction coordinate (R.C.). The different conformations (or windows) are generated in between the two states following the reaction coordinate. Each conformation is sampled by doing the MD simulation in presence of an external biased potential. System in which the end states are separated by the rough region of potential energy surface along the R.C. may experience poor sampling. Umbrella sampling improves the sampling of each conformation by applying a biasing umbrella potential to reach the conformational phase space which would not be feasible with the normal MD simulation.

$$E^b(r) = E^u(r) + V_i(\xi) \qquad (2.18)$$

$E^b(r)$, $E^u(r)$ are the biased and unbiased potential respectively and $V_i(\xi)$ is external umbrella potential.

The sampling in specific region of the coordinate close to ξ_0 is improved by using umbrella potential and often harmonic potential is used.

$$V_i(\xi) = \frac{1}{2}k(\xi - \xi_0)^2 \qquad (2.19)$$

This function will restrain the molecular configuration harmonically near the ξ_0.

In the umbrella sampling method, the potential of mean force (PMF) is given by the negative logarithm of unbiased probability distribution function.

$$W(r) = -k_\beta T \ln P^{(0)}(r) \qquad (2.20)$$

$$W(r) = -k_\beta T \ln P^{(b)}(r) - U^b(r) + F \qquad (2.21)$$

W(r) is PMF, $P^{(0)}(r)$, and $P^{(b)}(r)$ is unbiased and biased probability functions, respectively. $U^b(r)$ is biased potential. F is an undetermined factor which depends upon U^b (r). All the output histograms obtained from sampling of each window are combined by weighted histogram analysis method (WHAM).

The WHAM equations are the following.

$$P^0(x) = \frac{\sum_{i=1}^{N} h_i(x)}{\sum_{k=1}^{N} n_k \exp\left(\left[F_k - U^b(x)\right]/k_\beta T\right)} \qquad (2.22)$$

2.2 Dissipative particle dynamics simulation

Molecular dynamics simulation is capable of capturing motions of atoms at a very fine scale (nanoscale). The motion of the matter is represented by their atomic positions, and simulated by introducing appropriate interaction potential. The advantage of the method is that it can describe molecular motion and conformation precisely as long as the interactions potential between atoms are correctly calculated. But, atomistic simulation can't simulate system large system for longer time. The computational time to calculate the detailed interaction model in this paradigm severely limits its applicability in the region of large system. In the field of molecular biology and chemistry many important process such as membrane structuring occur in much larger time and length scales which is beyond the scope of molecular dynamics simulation. Dissipative particle dynamics (DPD) simulation is being very effective to overcome this limitation. The DPD technique simulates at a scale which is more relevant to those processes, by ignoring the degrees of freedom which are not immediately important to describe the system at the studied level.

DPD is a proper simulation method to perform mesoscale simulation of soft matter. Many properties of soft matter depend upon aggregate of molecules or supra-molecular domains instead of individual molecules. DPD has been widely applied to study self-assembly of systems such as polypeptides and lipid membranes, and biopolymers.[19-21] In DPD modelling, hydrophilic and hydrophobic groups are represented by different type beads. The beads are connected by bonds of certain force constant.

In the DPD simulation, the interaction potential between beads is soft in nature. Two beads with distance (r_{ij}) less than a particular cut-off value (r_c) interact via short-ranged forces.

$\vec{F}_{ij}^{C}$ is a conservative force exerted on particle i by particle j, $\vec{F}_{ij}^{D}$ is a dissipative fore, and $\vec{F}_{ij}^{R}$ is random force. The conservative force is repulsive in nature, acting up to cut-off radius which is 1.

$$\vec{F}_{ij}^{C} = \begin{cases} a_{ij}\left(1-\frac{r_{ij}}{r_c}\right)\hat{r}_{ij} & (r_{ij} < 1) \\ 0 & (r_{ij} \geq 1) \end{cases} \quad (2.23)$$

Where a_{ij} is maximum repulsion between particles i and j; $\vec{r}_{ij} = \vec{r}_i - \vec{r}_j$where $\hat{r}_{ij} = \vec{r}_{ij} \big/ \left|\vec{r}_{ij}\right|$

The drag force and random force work as heat sink and heat source, respectively. Together they maintain constant temperature of the system, and they are given by,

$$\vec{F}_{ij}^{D} = -\gamma\omega_d(r_{ij})\left(\hat{r}_{ij}\cdot\vec{v}_{ij}\right)\hat{r}_{ij} \quad (2.24)$$

$$\vec{F}_{ij}^{R} = \sigma\omega_r(r_{ij})\theta_{ij}\hat{r}_{ij} \quad (2.25)$$

where, $\vec{v}_{ij} = \vec{v}_i - \vec{v}_j$ is relative velocity and γ is friction coefficient, and are r-dependent weight-function and vanishes when r> r_c, θ_{ij} is a random variable with zero mean and variance of one. The noise amplitude σ is fixed to 3 following the work of Groot and Warren. According to the fluctuation-dissipation theorem for the DPD two weight functions are related to each other by $\omega_D = \omega_R^2$relation, and two amplitudes are related by $\sigma^2 = 2k_BT\gamma$

DPD simulation uses units of time and length scale which is higher compare to MD simulation. The length unit in DPD is equivalent to multiplication of 1/3 th power of density of system in DPD unit and radius of one DPD bead in real unit. Generally a bead is represented by 3 water molecules and density of system is considered 3 DPD unit. The volume of one water molecule is 30 Å. Therefore, 1 DPD unit equivalents to 6.46 Å. The time unit of DPD is equivalent to 25.7 $N_m^{5/3}$ in real unit (ps). When the one DPD unit is represented by 3 water molecules, the DPD unit of time become 160.4 ps. DPD simulation is carried out using NVT ensemble. Temperature parameter is considered as $k_BT = 1$. The mass of each bead is considered to be 1 dpd unit of mass.

References

[1] Allen, M. P., and Tildesley, D. J., *Computer Simulation of Liquids* (Oxford Univesity Press, New York, 1989).

[2] Jorgensen, W. L., Maxwell, D. S., and Tirado-Rives, J., "Development and Testing of the Opls All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids," Journal of American Chemical Society 118, 11225-11236 (1996).

[3] Kaminski, G. A., Friesner, R. A., Tirado-Rives, J., and Jorgensen, W. L., "Evaluation and Reparametrization of the Opls-Aa Force Field for Proteins Via Comparison with Accurate Quantum Chemical Calculations on Peptides," The Journal of Physical Chemistry B 105, 6474-6487 (2001).

[4] Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., and Kollman, P., "A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations," Journal of Computational Chemistry 24, 1999-2012 (2003).

[5] Oostenbrink, C., Villa, A., Mark, A. E., and Van Gunsteren, W. F., "A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: The Gromos Force-Field Parameter Sets 53a5 and 53a6," Journal of Computational Chemistry 25, 1656-1676 (2004).

[6] Brooks, B. R., Brooks, C. L., Mackerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., and Karplus, M., "Charmm: The Biomolecular Simulation Program," Journal of Computational Chemistry 30, 1545-1614 (2009).

[7] Fowler, J. E., Alberts, I. L., and Schaeferlll, H. F., "Mechanistic Study of the Electrocyclic Ring-Opening Reaction of Thiirane," Journal of American Chemical Society 113, 4768-4776 (1991).

[8] Kaminski, G., Duffy, E. M., Matsui, T., and Jorgensen, W. L., "Free Energies of Hydration and Pure Liquid Properties of Hydrocarbons from the Opls All-Atom Model," The Journal of Physical Chemistry 98, 13077-13082 (1994).

[9] Essex, J. W., Severance, D. L., Tirado-Rives, J., and Jorgensen, W. L., "Monte Carlo Simulations for Proteins: Binding Affinities for Trypsin−Benzamidine Complexes Via Free-Energy Perturbations," The Journal of Physical Chemistry B 101, 9663-9669 (1997).

[10] Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A., "A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules," Journal of the American Chemical Society 117, 5179-5197 (1995).

[11] Beachy, M. D., Chasman, D., Murphy, R. B., Halgren, T. A., and Friesner, R. A., "Accurate Ab Initio Quantum Chemical Determination of the Relative Energetics of Peptide Conformations and Assessment of Empirical Force Fields," Journal of the American Chemical Society 119, 5908-5920 (1997).

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

[13] Hockney, R. W., Goel, S. P., and Eastwood, J. W., "Quiet High-Resolution Computer Models of a Plasma," Journal of Computational Physics 14, 148-158 (1974).

[14] "Thermodynamic Integration of the Free Energy Along a Reaction Coordinate in Cartesian Coordinates," The Journal of Chemical Physics 112, 7283-7292 (2000).

[15] Shivakumar, D., Williams, J., Wu, Y., Damm, W., Shelley, J., and Sherman, W., "Prediction of Absolute Solvation Free Energies Using Molecular Dynamics Free Energy Perturbation and the Opls Force Field," Journal of Chemical Theory and Computation 6, 1509-1519 (2010).

[16] Roux, B., Nina, M., Pomès, R., and Smith, J. C., "Thermodynamic Stability of Water Molecules in the Bacteriorhodopsin Proton Channel: A Molecular Dynamics Free Energy Perturbation Study," Biophysical Journal 71, 670-681.

[17] Torrie, G. M., and Valleau, J. P., "Monte Carlo Free Energy Estimates Using Non-Boltzmann Sampling: Application to the Sub-Critical Lennard-Jones Fluid," Chemical Physics Letters 28, 578-581 (1974).

[18] Torrie, G. M., and Valleau, J. P., "Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling," Journal of Computational Physics 23, 187-199 (1977).

[19] Shillcock, J. C., and Lipowsky, R., "Tension-Induced Fusion of Bilayer Membranes and Vesicles," Nat Mater 4, 225-228 (2005).

[20] Gao, L., Shillcock, J., and Lipowsky, R., "Improved Dissipative Particle Dynamics Simulations of Lipid Bilayers," The Journal of Chemical Physics 126, 015101 (2007).

[21] Kranenburg, M., and Smit, B., "Phase Behavior of Model Lipid Bilayers," The Journal of Physical Chemistry B 109, 6553-6563 (2005).

Chapter 3

Propensity of Self-assembled Leucine-Lysine Diblock Co-polymeric α-Helical Peptides to Remain in Parallel and Anti-parallel Alignments in Water

3.1 Introduction

Amphiphilic molecules self-assemble and give rise to a wide range of structures. Individual hydrophobic and hydrophilic blocks of any amphiphilic molecule assemble near themselves and form ordered structures. Self-assembly of block copolymers is a well studied example. Block copolypeptides consisting of hydrophobic and hydrophilic blocks self-assemble to form various structures such as vesicles [1, 2], lamellae, rods [3], and many others [4−6]. Being biocompatible, the assembled structures formed from peptide block copolymers have potential applications in tissue engineering and drug delivery [7]. Molecular self-assembly of biomolecules has attracted plenty of attention because of easy tunable design that can lead to desirable nanostructures.

Block copolypeptides have been designed in different ways to obtain self-assembled structures [8, 9]. Bellomo et al. reported diverse self-assembled structures with diblock copolypeptides made up of leucine and poly(ethylene glycol)-substituted lysine blocks. They also varied the chain lengths of the block copolypeptides and the fraction of hydrophobic residues [10]. Vauthey et al. have designed surfactant-like peptides made of alternative D and L α-peptide and cyclic β-peptide to form nanotubes and nanovesicles [11]. The design principle which they have followed was based on strong cross-strand attractive forces such as electrostatic and hydrophobic interaction between the side chains. Hartgerink et al. designed bolaamphiphilic peptide conjugates that formed well ordered nanofibers. pH-responsive block copolypeptides have also been designed so that they selfassemble into different structures with variation in pH. Rodriguez-Herna´ndez et al. studied self-assembly of poly(L-glutamic acid)-poly(L-lysine) as a function of pH and observed reversible formation of schizophrenic vesicles [12]. Also, polybutadiene-block-poly(L-glutamate) copolymers form pH-dependent secondary structures and hence self-assemble into different morphologies depending upon pH [13, 14].

Further, theoretical and modeling studies have been performed to gain structural insight and quantification of interactions in the self-assembled structures. Lee et al. investigated the interactions responsible for stabilizing a cylindrical nanofiber in water using molecular dynamics (MD) simulations [15]. The influence of hydrogen bonds (H-bonds) on the self-assembly of polyamphiphilic peptides was studied by Cruz et al. using a simplified coarse grain model [16]. Coarse grain MD simulations have also been used to model self-assembly of the laminin-1 sequence (Ile-Lys-Val-Ala-Val) leading to the formation of fiber structures

[17]. Thota et al. have studied the self-assembly of Ala-Phe-His-Lys in water using the MARTINI coarse-grained model [18]. They have reported the formation of micelles with hydrophobic amino acids present at the core, hydrophilic amino acids at the periphery, and amphiphilic amino acids at the interface of hydrophobic and hydrophilic regions. They have also investigated the importance of hydrophilic and hydrophobic residues in self-assembly by varying the number of hydrophobic as well as hydrophilic residues [19]. They found that with an increase in hydrophilic residues, assembly decreases and smaller micelles form. The substitution of alanine by more hydrophobic phenylalanine results in faster self-assembly with an increase in the size of micelles. The driving factor for self-assembly is primarily the affinity of hydrophobic and hydrophilic moieties to order near themselves. An extensively studied example is the self-assembly behavior of lipids. Lipids consist of a hydrophilic headgroup and long hydrophobic tails. When they self-assemble to form different structures, the hydrophilic head groups and hydrophobic tails arrange near themselves [20, 21]. Also, assembly of block copolymers follows a similar trend [22−24]. However, structural insight into the self-assembly of block copolypeptides is lacking. This is maybe because of the complexity in the nanostructures formed by the peptides. What happens when a hydrophilic block of copolypeptide comes near a hydrophobic block during self-assembly? Is there a possibility for hydrophilic moieties to fall within the aggregated hydrophobic moieties or vice versa in the self-assembled structures obtained from block copolypeptides? All these questions remain unanswered.

In the present study, we have investigated the propensity of capped α-helical diblock copolypeptides, consisting of leucine (LEU) and lysine (LYS) blocks, to remain in the lamellar structure. Wimley and White suggested a hydrophobicity scale for different amino acids by measuring partitioning free energy in water−lipid and water−oil interfaces [25, 26]. According to this scale, LEU and LYS fall into two extremes, i.e., LEU is hydrophobic, and LYS is much less hydrophobic (hydrophilic). The reported free energies for transfer of the whole residue of LEU and LYS from water to n-octanol are −1.25 ± 0.11 and 2.80 ± 0.11 kcal/mol, respectively. According to other hydrophobicity scales defined for amino acids, LEU is again designated as more hydrophobic than LYS [27, 28]. This indicates that there should be phase separation between LEU and LYS blocks if they are in copolymeric form or in mixtures. However, in the presence of water, phase-separated nanostructure may disintegrate because of favorable water penetration into the LYS phase. Water may also invade the LEU phase because of local water affinity of the peptide linkage.

The helical copolymers considered in this work consist of ten residues of LEU (N-terminal) and ten residues of LYS (C-terminal) each (Figure 3.1). The helical copolymer chains were taken as capped, i.e., N-terminal and C-terminal are capped by $COCH_3$ (ACE) and $NHCH_3$ (NMA), respectively. The side chain of LEU contains the hydrophobic isopropyl group ($CH_2CH(CH_3)_2$), whereas LYS contains the n-butylamine group ($(CH_2)_4NH_2$). The mismatch in hydrophobicity is the basis of our choice of the LEU block as hydrophobic and the LYS block as hydrophilic.

Holowka et al. have reported preparation and characterization of a series of poly(L-lysine)-b-poly(L-leucine) block copolypeptides in water [29]. They have shown the formation of stimuli-responsive vesicles, which are responsible for the construction of the preassembled lamellar phase. Prabhu et al [30] have studied self-assembly of lionelic acid sophrolipid (LNSL) in water. They found that LNSL molecules aggregate to form spherical structure. In this present study, side chains of LYS were considered in the neutral form (NH2). Two different lamellar structures were constructed such that, in one case, the helical chains remain parallel to each other (LEU block near the LEU block), whereas in the other case the chains remain antiparallel to each other (LEU block near the LYS block). We have taken preassembled starting structures because of known preferential interactions between hydrophobic and hydrophilic blocks of the peptide chains. All atomistic self-assembly simulation of large peptides chains is computationally demanding, and often the system gets stuck in local minima. Therefore, we have performed MD simulations with these preassembled parallel and antiparallel lamellar structures in water and realized that both of these structures remain as stable lamella. We have envisaged the structural basis of the stability which is responsible for the favorable interactions of these lamellar structures even with hydrophobic mismatches. The present study provides structural insight into self-assembled diblock copolypeptides along with interactions, which favor the stability of the structure. Further, we have performed hydrogen bond analysis to identify the reason for the stability of these self-assembled structures. We found hydrogen bond network paths in different directions. Water plays a role of mediator for the formation of these hydrogen bond networks. Finally, we have addressed the stability of these preassembled peptide structures by performing free energy calculations.

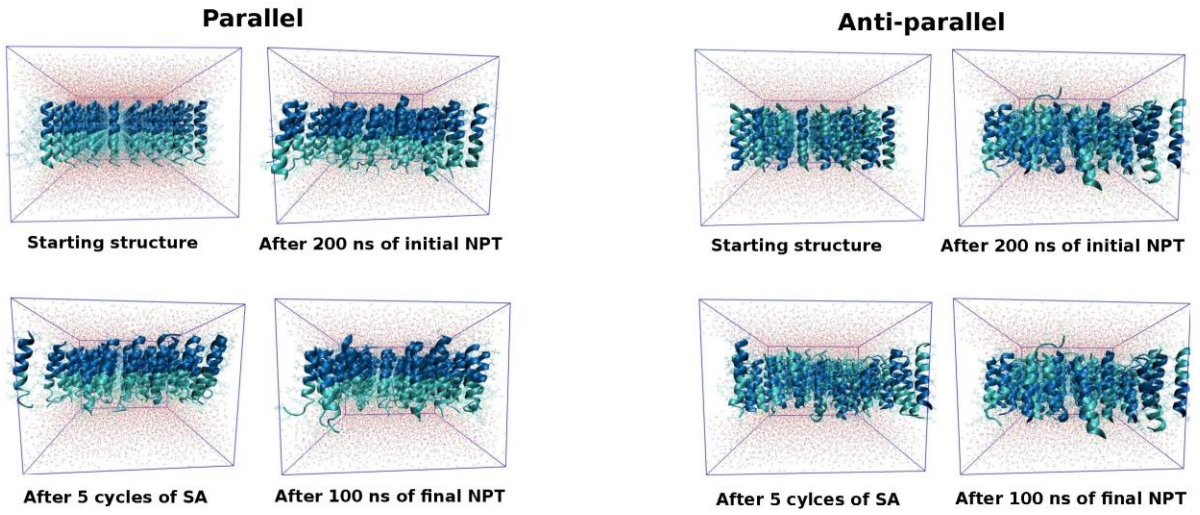


Figure 3.1: Snapshots of parallel and antiparallel alignments at different times. As indicated in the picture, blue denotes the LEU block and cyan denotes the LYS block. LEU and LYS blocks are near the N- and C-terminus, respectively.

3. 2 Computational Methods

All the simulations were performed with GROMACS-4.6.3.31 The Φ and Ψ angles for preparation of initial α-helical secondary structure of the capped α-helical diblocks copolymeric (ACE-LEU_{10}-LYS_{10}-NMA) chain is considered as −64.0° and −43.0°, respectively. The diblock copolymeric α-helical peptide consisting of ten residues of LEU and LYS each was constructed and translated along the XY-plane (8 × 8) to build the parallel and antiparallel lamellar structures. All atomistic OPLS force field [32−34] was used for the study. Water (TIP3P-model [35]) was added on both sides of the lamellar structure such that the water−lamella interface was created along Z-axis, and no water molecules were present inside the parallel or antiparallel lamella in the starting structure shown in Figure 3.1. The water content was maintained to be ∼50% by weight for both alignments (∼9000 water molecules). The initial system was constructed with simulation cell dimension of 9.4 × 9.4 × 6.8 nm^3 for both the systems. The constructed lamellar structures with water on both sides were energy minimized using the steepest descent algorithm to relax the bonds, angles, and dihedrals. Following this, a 200 ns NPT simulation, simulated annealing steps, and a final 100 ns simulation were performed. The three-step simulation protocol was followed to make

sure that the system is not trapped in a metastable state and that the lamellar alignments are stable.

3.2.1 NPT Run of 200 ns

A 200 ns isothermal−semi-isotropic ensemble (NPT) simulation was performed for both systems. Pressure was kept constant at 1 bar using Berendsen Barostat.36 Semi-isotropic pressure coupling was used with separate coupling to the XY-plane and Z-direction (normal to monolayer−water interface) for both systems. The temperature was maintained at 300 K using a V-rescale thermostat [37]. Bonds were not constrained, and a time-step of 1 fs was used. Long range electrostatic interactions were subjected to the particle mesh Ewald (PME) method [38] with a real space cutoff of 1.0 nm; also for the Lennard−Jones interaction, a cutoff of 1.0 nm was used. Trajectories were updated every 1 ps for both systems. Potential energy (PE) for the 200 ns simulation is depicted in Figure 3.2. To understand the time-scale of relaxation of peptide chains, the autocorrelation of the end-to-end distance of peptide chains was calculated for the initial 200 ns and final 100 ns NPT simulation before and after simulated annealing (SA), respectively. The details of the calculation are given in following section, and autocorrelations as a function of time for parallel and antiparallel alignments are shown in Figure 3.3a and Figure 3.3b, respectively. The plots were fitted to the Kohlrausch−Williams−Watts (KWW) [39] stretched exponential. The timescales of relaxations are found to be very high, 2.67×10^4 for parallel alignment and 3.822×10^3 for antiparallel alignment.

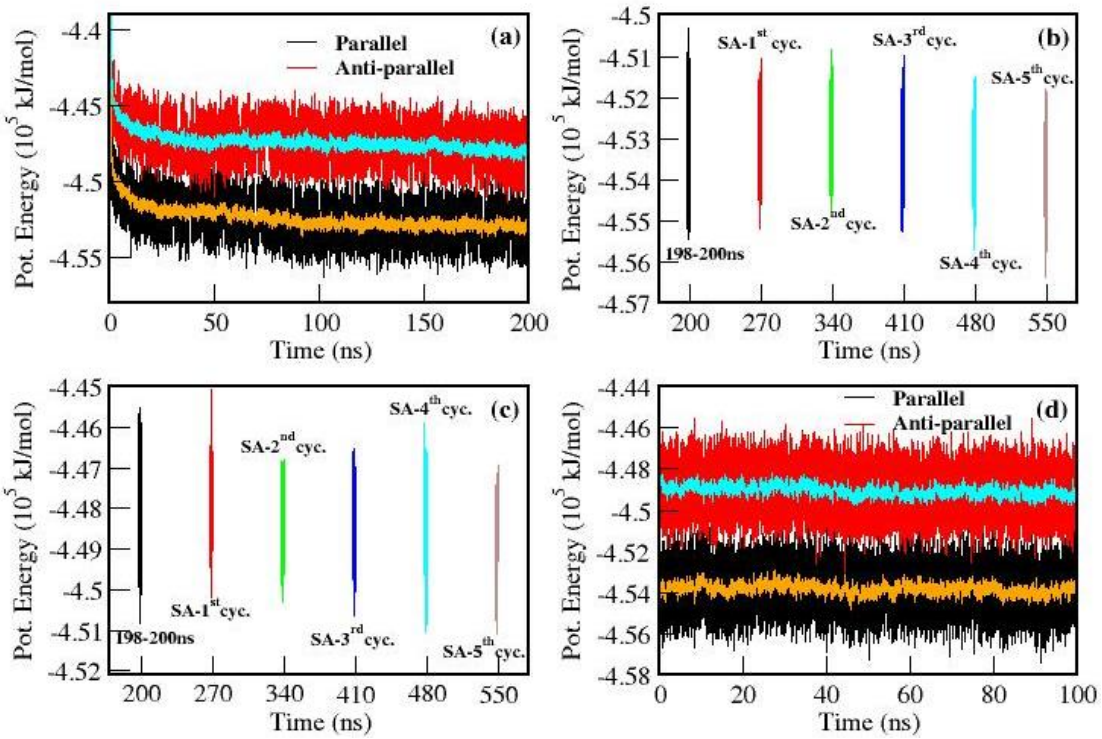


Figure 3.2: (a) Plot of potential energy with initial 200 ns of semi-isotropic NPT run for parallel and anti-parallel alignment are shown by black and red lines respectively and their running averages are shown by orange and cyan colour respectively. Plot of potential energy with time for 2ns of equilibration time at temperature 300 K after completion of each SA cycles for **(b)** parallel alignment and **(c)** anti-parallel alignment. **(d)** Potential energy plot for parallel and anti-parallel alignment for final 100ns of semi-isotropic NPT run after completion of SA cycles are shown by black and red lines respectively and their running averages are shown by orange and cyan colour respectively.

3.2.2 Calculation of end-to-end distance auto-correlation function

Autocorrelation function is defined as the amount of similarity of a property observed between a given time series and its lagged version as a function of time gap between them. We have taken a collection of 64 α-helical polypeptide chains (8*8) in assembled form for our study. The end-to-end distance is an important dynamical property of α-helical polypeptide as it varies with the progresses of simulation of monolayer systems. The distance between C atoms of two terminal $–CH_3$ groups of a peptide chain is defined as the end-to-end distance. To understand the relaxation time scale of our system of study, the auto-correlation

function of end-to-end distance is calculated over all the 64 peptide chains of parallel and anti-parallel alignments.

The autocorrelation function, C(t) of end-to-end distance is given by,

$$C(t) = \frac{\langle R(\tau) R(\tau + t) \rangle}{\langle R^2 \rangle}$$

,

where $R(\tau)$ is end-to-end distance of the di-block co-polypeptide chain at any time moment τ and after time t its value becomes $R(\tau + t)$. $\langle R^2 \rangle$ is the mean-square average of end-to-end distance. The last 80ns of NPT simulation trajectory before and after performing SA were considered for calculation. To calculate relaxation times (τ_R), these plots were fitted to KWW stretching potential of the form,

$$C(t) = \exp\left[-\left(\frac{t}{\tau_{KWW}} \right)^{\beta} \right],$$

where β is stretch exponent which considers the deviation from ideal behaviour and τ_{KWW} is KWW relaxation time. The values of β and τ_{KWW} were used to calculate τ_R using the following equation,

$$\tau_R = \frac{\tau_{KWW}}{\beta} \Gamma\left(\frac{1}{\beta} \right)$$

The values of β, τ_{KWW}, and τ_R for parallel and anti-parallel alignments before and after SA are given in Table 3.1.

Table 3.1: Table showing the values of β, τ_{KWW}, τ_R for parallel and anti-parallel alignments before and after SA

	Systems	β	τ_{KWW} (ns)	τ_R (ns)
Before SA	Parallel	0.2436	948.336	2.67×10^{4}
	Anti-parallel	0.176	17883.9	7.127×10^{6}
After SA	Parallel	0.1747	8.843	3.822×10^{3}
	Anti-parallel	0.3636	468.8	2.153×10^{3}

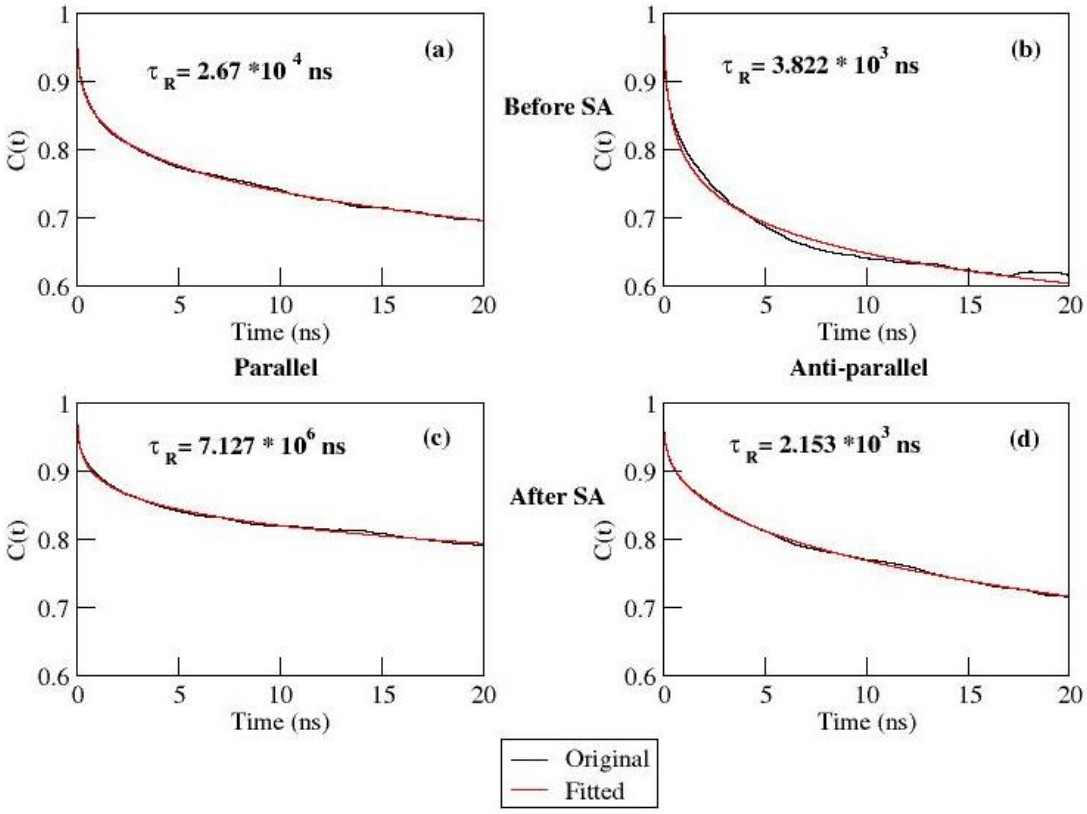


Figure 3.3: (a) End-to-end distance auto-correlation function vs time plot calculated over last 80 ns of NPT simulation trajectory before simulated annealing for parallel alignment, **(b)** for anti-parallel alignment. **(c)** End-to-end distance auto-correlation function vs time plot calculated over last 80 ns of NPT simulation trajectory after simulated annealing for parallel alignment, **(d)** for anti-parallel alignment.

3.2. 3 Simulated Annealing Steps

After 200 ns of the isothermal−semi-isotropic NPT run, SA was performed for both types of monolayer. SA was carried out in order to check if the system is trapped in any metastable state. Also, SA would allow the systems to scan any other lower energy conformation if present. The SA procedure was carried out using the same parameter as that used during the initial 200 ns NPT run. All the other simulation parameters were unchanged. The temperature of the monolayer systems were periodically heated and cooled in a temperature range of 300−348 K with an interval of 8 K. A time duration of 2 ns was used to increase the temperature in each step which was followed by 3 ns of equilibration at each temperature.

Potential energies at 300 K after completion of each SA step for parallel and antiparallel alignments are provided in Figure 3.2b and Figure 3.2c, respectively.

3.2.4. NPT Run of 100 ns

After completion of five cycles of SA, again the semi-isotropic NPT run was performed for 100 ns using the same simulation protocol for the initial 200 ns simulation. PE for this final 100 ns simulation is shown in Figure 3.2d. The last 10 ns from the 100 ns trajectory after the SA steps were considered for further analysis. The end-to-end distance autocorrelation function as a function of time for parallel and antiparallel alignments are shown in Figure 3.3c and Figure 3.3d, respectively. The relaxation times of the end-to-end distance have also been calculated from the NPT simulation trajectory. From all these time-correlation plots it is evident that the structures of the peptide chains do not change much even after SA steps. In the case of a system with parallel alignment of peptide chains there is an increase in relaxation time, which shows that the SA steps help in stabilizing the chains and therefore self-assembly. The end-to-end distance autocorrelation function as a function of time plots also suggests that there is not much change in helicity of the peptide chains. We have further investigated the helicity of chains by calculating different order parameters. The calculation of order parameter is described in the following section. Equilibration was confirmed by convergence of potential energy (Figure 3.2d), penetration of water into the monolayer (Figure 3.4b), helicity of α-helical copolypeptide chains (Figure 3.5), for both systems. Methods for calculating helical order parameter and water penetration are explained in following section and in Results and Discussion of the main text, respectively. Trajectories were analyzed further to calculate partial density along the peptide−water interface. The tetrahedral ordering of water in different regions of monolayer has been evaluated. Distance distributions were calculated among different atoms in the peptide helices. Interhelical H-bonds and H-bonded network structures were calculated and analyzed. This H-bond network length was further analyzed in different directions of the peptide monolayer. The analysis procedures are described in Results and Discussion. Finally, free energy calculations were performed to evaluate the affinity of the helical diblock copolypeptide chains to remain in parallel and antiparallel alignments.

3.2.5 Calculation of helical order parameters

The H2, H3, H4 helical order parameter does not depend upon any predefined criterion for helix. To calculate these helical parameters the centre of mass (COM) of backbone atoms for each of residues in alpha helical peptide chains is calculated. At the COM of each residue a vector is defined ($\vec{u_i}$) by the cross product of two connected bond vectors.

$$\vec{u_i} = (\vec{r_i} - \vec{r}_{i-1}) \times (\vec{r}_{i+1} - \vec{r_i}) / \sin\theta ,$$

$\vec{r_i}$ is the position of COM of i-th monomer and θ is angle between two connected bond vectors. The terminal capping residues are excluded from calculation of bond vectors. The local helicity is calculated by taking dot product of u_i vectors of conjugated residues and denoted as H_2.

$$H2 = \frac{1}{N-3} \sum_{i=2}^{N-2} \vec{u}_i \cdot \vec{u}_{i+1}$$

The other helical parameters are defined as,

$$H3 = \frac{1}{N-2} \sum_{i=2}^{N-1} \vec{u}_i \cdot \vec{u}_{mid} ,$$

u_{mid} is bond vector of middle monomer (N/2) of peptide chain.

$$H4 = \left(\frac{1}{N-2} \sum_{i=2}^{N-1} \vec{u}_i \right)^2$$

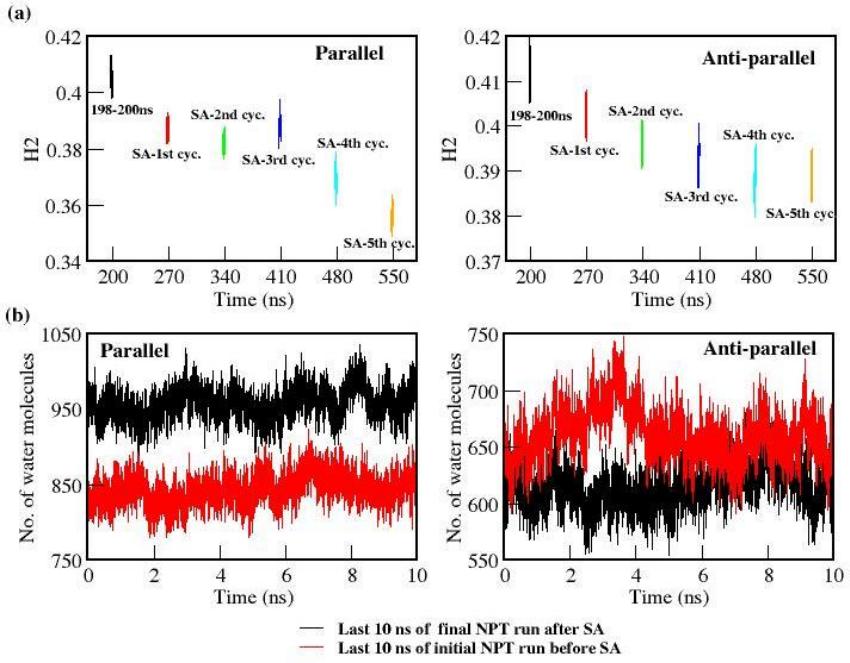


Figure 3.4: (a) Plot of H2 helical parameters with time after different SA cycles at 300 K for parallel (upper left-hand side graph) and anti-parallel alignment (upper right-hand side graph). The H2 helical parameter for 198-200 ns semi-isotropic NPT run and following 5 cycles of SA are shown by black, red, green, blue, cyan and orange colour respectively **(b)** Plot of number of water molecules inside peptide monolayer for last 10 ns of initial 200 ns of NPT run before performing SA steps (black line) and last 10 ns of 100ns NPT after SA (red line) for parallel (lower left-hand side graph) anti-parallel alignment (lower right-hand side graph).

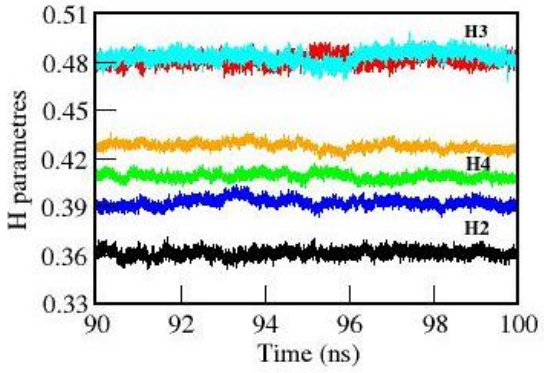


Figure 3.5: Plot of average H2, H3, H4 helical parameters with time for alpha-helical chains for parallel and anti-parallel monolayer. Black, red and green line represents H2, H3, H4 parameters of parallel alignment respectively. The same parameters for anti-parallel are shown by blue, cyan and orange lines respectively.

3.2.6 Free Energy Calculation

We have carried out free energy calculations to understand the energetics of the interaction of two peptide chains in their different orientations and also of a single peptide chain to a preassembled system of peptide chains in two different alignments as discussed above. To mimic all the interactions between two peptide chains of the equilibrated system, two chains oriented in parallel fashion were selected from the parallel monolayer alignment. One chain was kept fixed and another chain was rotated up to 180° with 30° of intervals with respect to the X-axis passing through its center of geometry. In this way, parallel alignment of peptide chain was transformed to antiparallel alignment. Thus, seven initial configurations were generated and each was solvated in the same amount of water. The force fields for peptide−water models were considered the same throughout this work. All the initial configurations were first energy minimized using the steepest descent method. It was followed by 1 ns of NPT simulation by positionally restraining the C_α atoms of the peptide chains using a force constant (f.c.) of 1000 kJ/mol/nm^2 so that water can well equilibrate around the peptide and the system can attain equilibrium density. Then for each initial angle, one chain was pulled along the X-direction of the box (perpendicular axis to the reference chain) at a velocity of 5 nm/ns by applying a harmonic potential to its center of mass (COM) with a f.c. of 300 kJ/mol/nm^2. During the pulling simulation, the C_α atoms of the reference chain were positionally restrained in all three directions using a f.c. of 1000 kJ/mol/nm^2 whereas the other chain was positionally restrained only in Y- and Z-directions so that the initial angle of orientation (between two chains) remained intact throughout the pulling. Snapshots (configurations) from these pulled trajectories were then used to perform umbrella sampling simulations. All these configurations were equilibrated initially for 200 ps of the NPT run, keeping all the C_α atoms of both peptide chains positionally restrained in all directions. It was followed by 1 ns of an umbrella sampling production run by applying a harmonic potential with a f.c. of 150 kJ/mol/nm^2. During these production runs, positional restraint on the C_α atoms were removed along the X-axis so that it could move only along the X-axis, but the angle of orientation between pulled and reference chain remained fixed. Then the weighted histogram method (WHAM) [40] was used to calculate the potential of mean force (PMF). The PMFs as a function of COM distance between two peptide chains at different angles of orientation are shown in Figure 3.6. In these PMF plots each minimum represents the bound state at that particular angle between two peptide chains. The free energy difference between the bound and reference state is the free energy of binding. The

free energy of binding at different angles of orientation between reference and pulled chains is depicted in Figure 3.15. From the plot it is evident that the free energies of binding are comparatively more favorable for parallel and antiparallel alignments, which supports our choice of these two initial configurations in their preassembled form.

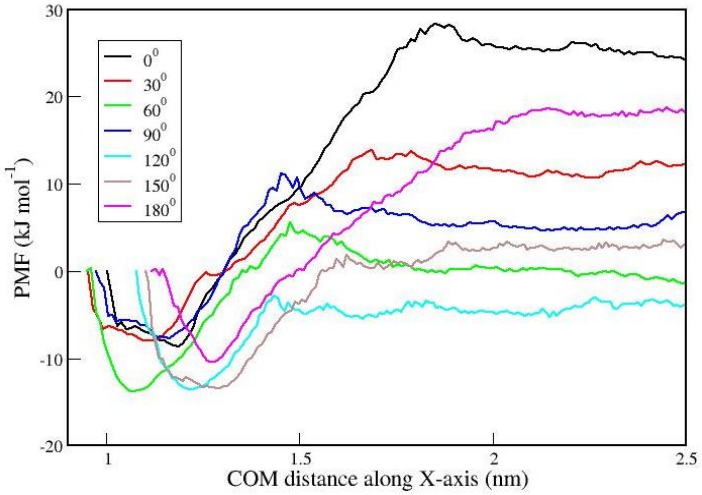


Figure 3.6: Plot of PMF with COM distance between reference and pulled chain at different angle of orientation between them**.**

Further, we have performed free energy calculations more realistically for monolayers of peptides in parallel and antiparallel alignments. For this we have considered a collection of 16 peptide chains (2 × 8 in the XY-plane) in water (which was extracted from the 95th ns frame of the last 100 ns of NPT run after SA). The simulation box dimensions were kept the same. A sufficient amount of water was added to fill the void space created after removal of the other peptide chains. The system was first energy minimized by the steepest descent method. Then 5 ns of NPT simulation was performed, keeping the peptide chains positionally restrained. It was followed by 2 ns of the NPT run by positionally restraining only the C_α atom. of the LEU residue, which is located at the middle (11th residue) of the peptide chain. Then one peptide chain from the middle was pulled along the X-direction at a velocity of 5 nm/ns by applying a harmonic potential to its COM with a f.c. of 150 kJ/mol/nm^2 . The value of f.c. was chosen such that it causes minimal disruption of helicity of that peptide chain during pulling. The Cα atoms of the 11th LEU residue of all other chains were kept positionally restrained during the pulling simulation, and these were collectively considered as the reference group. The different frames extracted from pulled trajectories at certain

intervals of the COM distance between pulled and reference chains were first equilibrated by 100 ps of the NPT run, which was followed by 5 ns of an umbrella sampling production run by applying the same harmonic potential used during pulling. The PMF vs COM distance was extracted in the same method as discussed above. The statistical error was calculated using a bootstrap analysis method.

3.3. Results & Discussion

3.3.1 Helicity of Peptide Chains

The α-helical diblock copolypeptides, arranged in parallel and antiparallel arrangements with water on both sides, retained their lamellar structure after the initial 200 ns of NPT simulations, followed by five simulated annealing cycles and 100 ns of NPT simulation. Snapshots after finishing the different parts of the simulation are depicted in Figure 3.1 for both parallel and antiparallel alignment. For both alignments, though the helical structures of the individual helices seem to be partly disrupted visually, the lamellae as a whole remain intact. To quantify the degree of helicity of the α-helical copolypeptide chains H2, H3, and H4 helical parameters are plotted as a function of time for the last 10 ns of the final 100 ns NPT run (Figure 3.5). The details of the calculation procedure of these helical parameters [41] are discussed in section 3.2.5. In Figure 3.5, from the values of H2 and H4 parameters it is evident that a higher degree of helicity persists for antiparallel alignment in comparison to parallel alignment. As H3 indicates helical ordering of the middle residue with the other residues of the peptide chain, therefore it is found to be similar for both types of alignment. Further, H2, H3, and H4 parameters from the last 10 ns is compared with the H2, H3, and H4 parameters of the energy-minimized structure and is shown in Table 2. From the comparison it appears that even after 200 ns of the NPT run, followed by SA steps and a final 100 ns of the NPT run, helicity retained for parallel alignment is ∼80% and for antiparallel alignment ∼86%. This retention in helicity and constancy in PE during the three steps of the simulation protocol indicates the stability of the lamellar structure and also that the system is not trapped in a metastable state. The H2 parameters are also calculated after completion of each SA cycle at 300 K for both alignments (Figure 3.4a). The decreases in average helicities after completion of the SA cycles with respect to the structure obtained after completion of 200 ns of the initial semi-isotropic NPT run are ∼12.5% and ∼9% for parallel alignment and antiparallel alignment, respectively.

Table 2: Table showing conservation of percentage helicity with respect to energy minimized structure by evaluation of H2, H3, H4 helical parameters for parallel and anti-parallel alignment

System	Helical parameters	Helicity averaged over last 10ns of NPT after SA	Std. Dev.	Helicity of energy minimized structure (EM)	Helicity conserved wrt EM structure (%)	Average helicity conserved wrt EM structure (%)
Parallel	H2	0.361	± 0.003	0.463	78	79.8
	H3	0.482	± 0.003	0.588	82	
	H4	0.409	± 0.002	0.515	79.4	
Anti-parallel	H2	0.393	± 0.003	0.463	84.88	85.6
	H3	0.483	± 0.004	0.54	89.94	
	H4	0.428	± 0.002	0.522	82	

3.3.2 Water Penetration in Lamellar Structure

Water molecules penetrated both the parallel and the antiparallel peptide monolayer. Thus, to understand the degree of penetration of water molecules, plots of partial density as a function of simulation box dimension along the Z-axis (monolayer−water interfacial axis) are depicted in Figure 3.7a and 3.7b. Partial density plots are averaged over 5 ns in different intervals of time. There is almost no change in partial density of water observed during the last 100 ns of simulation. This indicates that the monolayer is well equilibrated and there is no chance of water flooding and disintegration of the monolayer. The small amount of water molecules that penetrated the monolayer is only due to the water affinity of the amide linkages. For parallel alignment, the LEU and LYS block regions can clearly be distinguished (Figure 3.7). Water density decreases on moving from bulk to the middle of the monolayer. Finite density of water in the middle region of the monolayer is indicative of the penetration of water molecules into the monolayer structure. LYS blocks get more water density than the LEU,

which is expected from a hydrophobicity scale. For antiparallel alignment, there is finite density of water in the middle region of the monolayer. To verify the above findings and to quantify them, we have plotted the number of water molecules per unit volume in different parts of the peptide monolayer. The peptide monolayer region was identified from 95 to 100 ns interval of time in the partial density plots (Figure 3.7a and 3.7b) as 1.2 to 4.4 nm for parallel and 1.6 to 4.6 nm for antiparallel alignment along the Z-axis. To calculate the number of water molecules per unit volume in these peptide monolayer regions, they were divided 8 × 8 × 3 grids along the X-, Y-, and Z-axis (the Z-axis is the peptide−water interfacial axis). The number of water molecules falling into each grid were counted and divided by the volume of the grid. These numbers were converted into a histogram, and distribution was plotted (Figure 3.7c), averaged over the last 10 ns of the trajectory. For parallel alignment, the three parallel grids along the Z-axis were chosen as follows: from Figure 3.7a, the intersection point of partial densities of water and the LEU block (1.2 nm) to the point where density of the LYS block becomes zero (2 nm), 2 nm to the point where density of the LEU block becomes zero (3.25 nm), and 3.25 nm to intersection of partial densities of the LYS block and water (4.4 nm). There are three more regions in the parallel monolayer, the LEU−water interface region (0.7−1.2 nm), the LYS−water interface (4.6−5.2 nm), and the bulk, which are also considered for calculating the three-body order parameter of water. Furthermore, distinction between LEU and LYS blocks is not present for antiparallel alignment (Figure 3.7b). Thus, the peptide monolayer region (1.6−4.6 nm) was divided into three equal parts, i.e., 1.6−2.6 nm, 2.6−3.6 nm, and 3.6− 4.6 nm. The two outer regions in the peptide monolayer are referred to as the periphery, and the middle region as the core. In Figure 3.7c, the two peripheral regions (LEU and LYS sides) and core are considered and plotted separately for parallel alignment. The peptide−water interface region is 1.0−1.5 nm and 4.6−5.1 nm, which are considered during calculation of the order parameter of water. The numbers of water molecules per unit volume at the periphery are more than that of the core, which shows that two interfaces are created between water and peptide blocks and that water density slowly decreases toward the core of the monolayer. The interface toward the LYS side shows more water molecules per unit volume. The average number of waters per unit volume in the small grids of LEU, LYS, and core region of the parallel monolayer was found to be 4.73, 6.25, and 0.65, respectively. The number of waters per unit volume present at the core of the system is similar for parallel alignment and antiparallel alignments. This observation is also consistent from the partial density plots (Figure 3.7a and 3.7b). Thus, it is

evident that a similar amount of water molecules penetrates the parallel and antiparallel alignments, and still the lamellar structure remains intact.

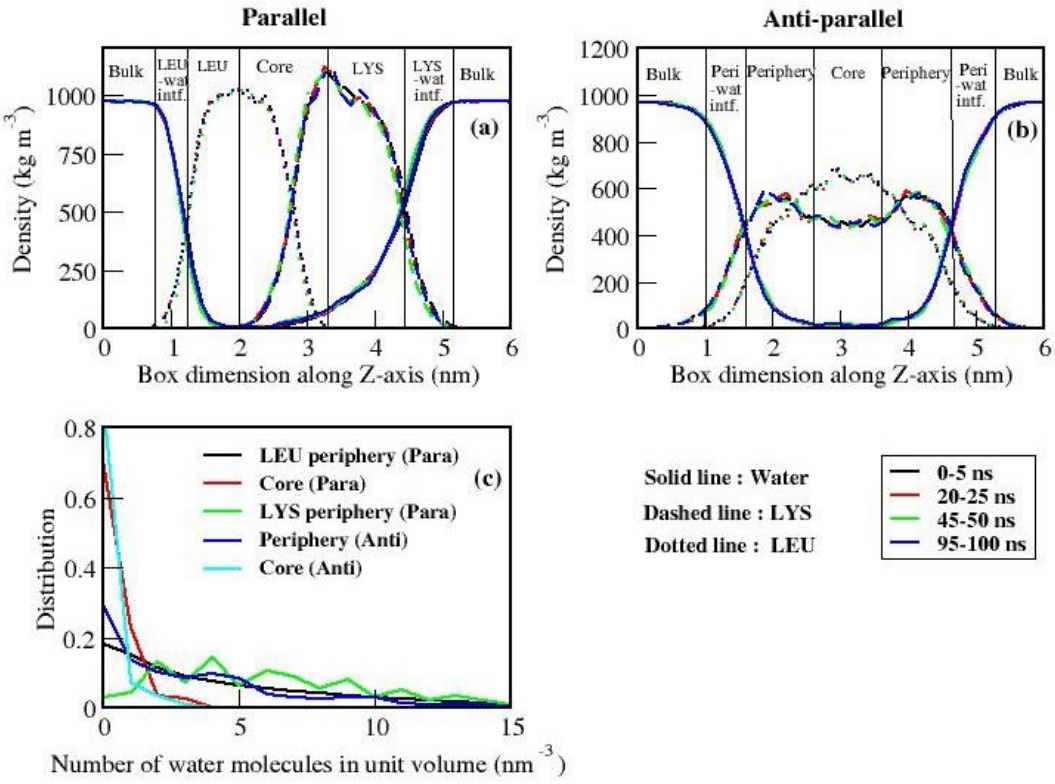


Figure 3.7: Partial densities of leucine (dotted line), lysine (dashed line) and water (solid line) along normal of peptide-water interface consisting of peptide lamella in **(a)** parallel alignment and **(b)** anti-parallel alignment. Black, red, green, blues lines represent partial density averaged over 0-5 ns, 20-25 ns, 45-50 ns, 95-100 ns respectively. **(c)** Distribution of number of water molecules per unit volume penetrating different region the lamellar structure.

3.3.3 Structural Arrangements

Realizing the penetration of water molecules into the monolayer structures, we found it interesting to investigate the arrangement of water molecules that have penetrated the monolayer structure. Thus, distance distributions between water oxygen (O) and side-chain nitrogen (N) for both parallel and antiparallel alignment are depicted in Figure 3.8a from the last 10 ns of the final 100 ns NPT run. Also, distance distributions between water O and water O for the water molecules that have penetrated the lamella are also shown in the same

figure. Distances between the respective atoms were computed in each frame of the last 10 ns of the final 100 ns trajectory and converted into a histogram. Distributions were plotted, normalizing the histograms with respect to the number of frames and average number of water molecules penetrating the monolayer structure (Figure 3.7c) in the last 10 ns trajectory. Minima of the first peak of distance distributions between side-chain N and water O appear in the same position (~0.35 nm) for both the parallel and the antiparallel alignment. The minima of water O−water O distance distribution appear at ~0.33 nm for both alignments. The peak height of the water O−water O distribution was found to be higher (Figure 3.8a). In the parallel alignment the hydrophilic LYS blocks are packed close to each other, but in the antiparallel alignment hydrophilic LYS blocks are packed next to hydrophobic LEU blocks. Thus, in the case of antiparallel alignment, water molecules penetrating the lamellar region would feel the interaction from both LEU and LYS blocks irrespective of its presence at the periphery or core (Figure 3.7c). This is not the case for parallel alignment, but the appearance of peak heights and the peak minimum at similar positions in Figure 4a suggests similar arrangement of water molecules in the two cases. To better understand the arrangement of water molecules, the tetrahedral order parameter [42] (q_{order}) is calculated (Figure 3.8b) in different regions as depicted in the plots of partial densities (Figure 3.7a and 3.7b). The value of the order parameter of water in bulk is ~0.55. It is lowered in the LEU−water interface region (~0.46) and further decreases in the LEU region (~0.38). The order parameter of water in the core region has not been calculated, as it requires at least four neighboring water molecules within minima of the distance distribution of water O−water O. The values of qorder are higher in the hydrophilic LYS region (~0.45) and the LYS−water interface region (~0.51) than in the other regions of parallel alignment. The value of the order parameter is slightly higher in the periphery−water interface region (~0.47) than in the periphery (~0.45) of antiparallel alignment.

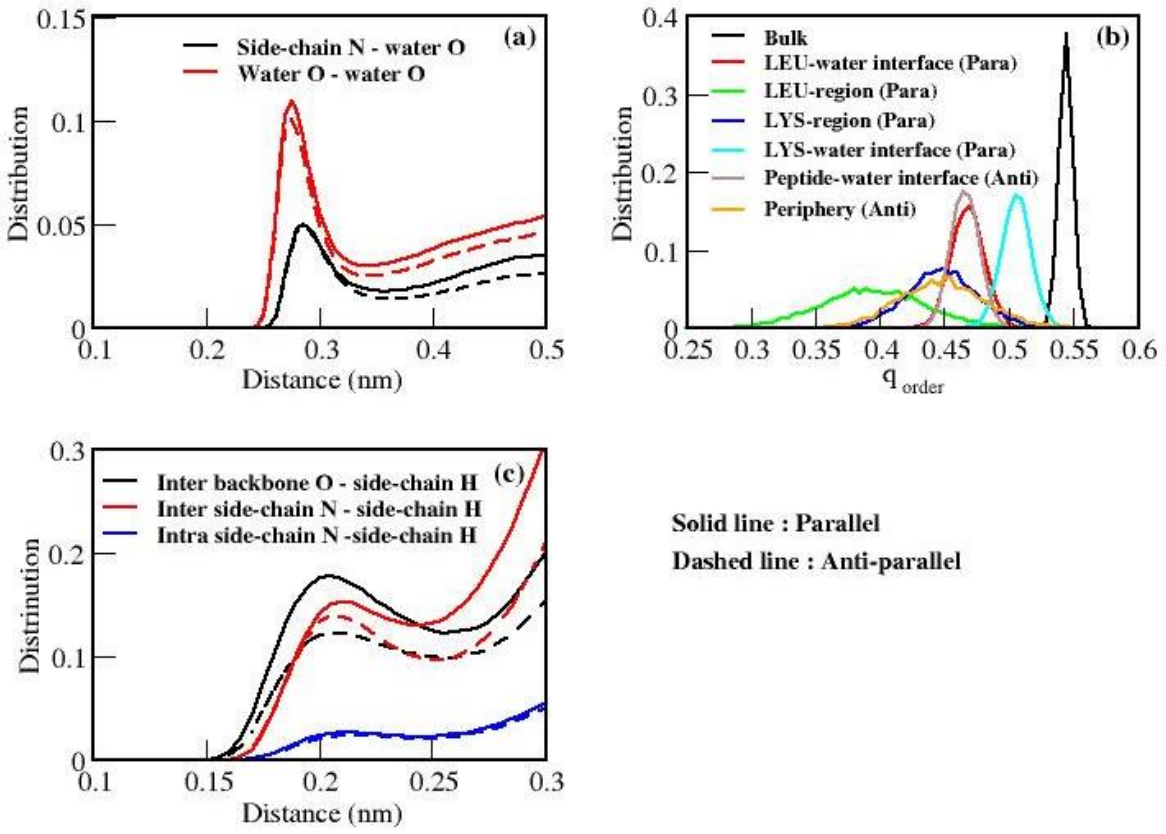


Figure 3.8: (a) Water O – water O, and water O – side-chain N distance distribution. Solid line and dashed represents these distribution of parallel and anti-parallel monolayer respectively. **(b)** Distribution of tetrahedral order parameter of water in different region of parallel and anti-parallel monolayer. **(c)** Inter-helical side-chain H – side-chain N, side-chain H – backbone O, and intra-helical side-chain H – side-chain distance N distribution of parallel and anti-parallel monolayer.

Further, to understand how the side chains arrange and how eventually inter- as well as intramolecular H-bonds are formed, distance distributions are plotted among side chains, and side chain with main chain (Figure 3.8c). Histograms were defined similar to those in Figure 4a. Distributions were normalized with the number of frames appearing in the last 10 ns of the final 100 ns NPT run and the number of peptide helices in the monolayer. The interhelical distance distributions are plotted between carbonyl oxygen (O) of the peptide backbone and hydrogen (H) of the side-chain amine on the LYS residues, and the side-chain amine N and the amine H. Also, intrahelical distance distribution between the side-chain amine N and H is calculated and shown in Figure 4b. Distribution between backbone O and side-chain amine H

shows a peak near ~0.26 nm for both parallel and antiparallel alignment. This indicates that the side chains of neighboring helical chains come in close proximity to the main chains. This is possible only if the side chains of the neighboring peptide helices penetrate each other in both alignments. To further visualize the interpenetration of the side chains, snapshots extracted from real time trajectory are depicted in Figure 3.9. The interpenetrating residues depicted in Figure 3.9 are recognized from distance distribution analysis code. Residues for which backbone O and side-chain H distance was <0.26 nm were determined for the 95 ns snapshot from the code for both alignments. Out of them some residues are depicted for a single frame in Figure 3.9 for clarity.

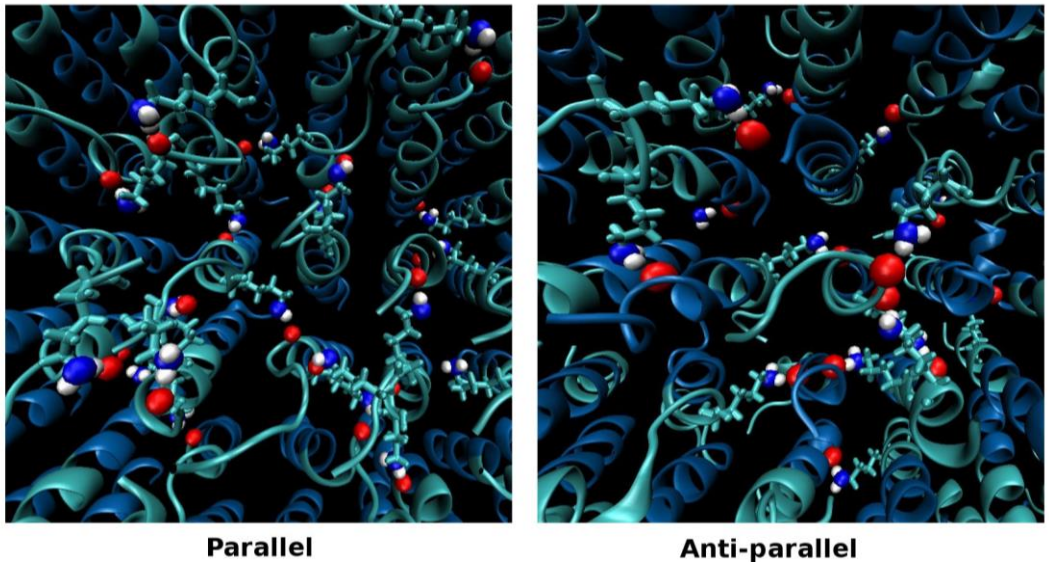


Figure 3.9: Snapshots extracted from real time trajectory showing inter penetration of side-chains. LEU and LYS blocks are represented by blue and cyan colours respectively, side-chains of LYS-blocks are represented by cyan. Main chain O and side chain amine H are represented as red and white sphere respectively.

In parallel and antiparallel alignments, although the peak heights of distance distribution side-chain N side-chain H are similar, peak minima appear at a slightly higher distance for antiparallel alignment (Figure 3.8c). In parallel alignment, all the side chains of the LYS block are in close proximity to each other, whereas in antiparallel alignment, the side chains of LYS are only close in a diagonal direction. Therefore, in antiparallel alignment, a side-chain−side-chain interaction occurs only through the diagonal direction which shows a location of peak minima at a slightly higher distance compared to that of parallel alignment.

3.3.4 Origin of Stability of Lamellar Structures

It can be clearly interpreted from the distance distributions (Figure 3.8a) that water molecules may arrange near themselves, as well as near the interactions sites of peptide helices. Also, interpenetration of the side chains suggests stabilizing interactions between the α-helical chains. It is important to investigate the interactions among the helices in order to understand the guiding factor for stability of the lamellar structures. In parallel alignment, the LYS blocks are packed close to each other. The side-chain NH_2 groups on the LYS blocks are capable of engaging in interhelical H-bonding interactions among themselves and with the main chain due to interpenetration of side chains. For antiparallel alignment, NH2 groups of the diagonal LYS blocks may also undergo H-bonding (Figure 3.9). As the nature of the interaction sites are the same in both systems, i.e., parallel and antiparallel arrangements, quantification of the number of H-bonds will lead us toward the origin of the integrity of the lamellar structure. In Figure 3.11a, we show the distribution of the number of H-bonds between side-chain NH2 groups among themselves and with main-chain amide groups for both parallel and antiparallel alignment. H-bonds among donor−H−acceptor were calculated, taking into account both distance and angle criterion. The distance cutoff between H and the acceptor pair was defined near the minima of the first peak of the distance distributions presented in Figure 3.8c. Acceptors falling within this distance of H attached to donor were identified and the donor−H−acceptor angle was calculated (plot of donor−H−acceptor angle distribution is shown in Figure 3.10). If the angle exceeded 140°, H-bonding was considered. The numbers of such H-bonds were counted in each frame of the last 10 ns of the final 100 ns NPT run and converted into a histogram. These histograms are normalized with the number of frames in the last 10 ns of the trajectory. The number of interhelical H-bonds among side-chain NH_2 groups is higher than the number of H-bonds between side-chain H of NH_2 groups and main-chain carbonyl O for both parallel and antiparallel alignment (Figure 3.11 a). This is due to the fact that in parallel alignment all NH_2 groups are in close proximity to each other, which increases the probability of formation of sidechain−side-chain H-bonds compared to side-chain−main H-bonds. In the case of antiparallel alignment, the number of sidechain−main-chain H-bonds as well as side-chain−side-chain H-bonds is less than that of parallel alignment. In antiparallel alignment, as the LYS blocks are packed close to each other only in the diagonal direction (Figure 3.9), the number of side-chain−side-chain H-bonds is less than in parallel alignment (Figure 3.11a). In parallel alignment, the number of side-chain− side-chain H-bonds is lower than the number of main-chain− side-chain H-bonds

whereas the counts of these two types of H-bonds are similar in antiparallel alignment. The number of intrahelical side-chain−side-chain H-bonds is much lower than that of interhelical H-bonds for both parallel and antiparallel alignment. From Figure 3.11a, it is evident that the total number of H-bonds (main-chain−side-chain and side-chain−sidechain) in antiparallel alignment is lower than in parallel alignment. However, a large number of H-bonds among the helical chains is present for both arrangements, which preserves the integrity of the lamellar structure.

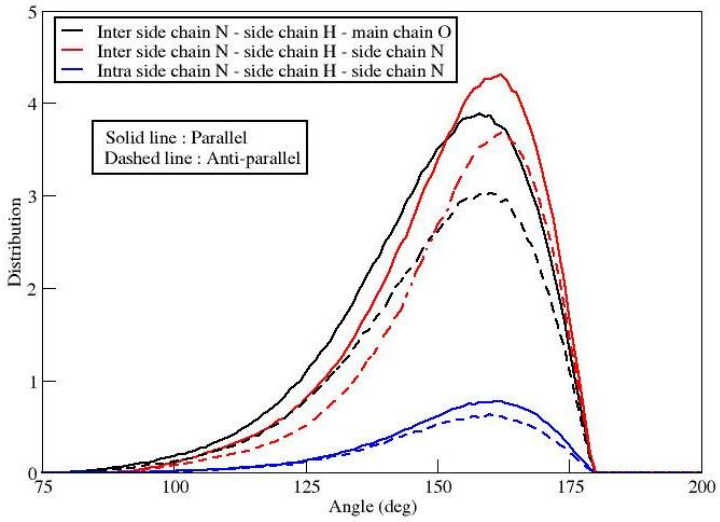


Figure 3.10: Plot of distribution for hydrogen bond donor – hydrogen - hydrogen bond acceptor angle for parallel (solid line) and anti-parallel alignment (dotted line). The inter side chain N – main chain H – main chain O, inter side chain N – main chain H – main chain, intra side chain N – main chain H – main chain bond angles are shown by black, red and blue colour respectively.

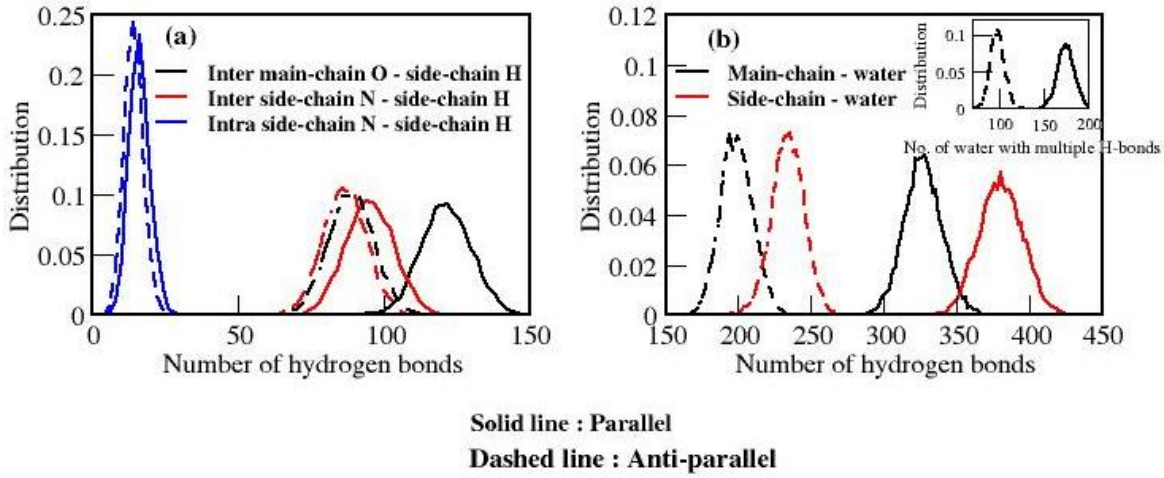


Figure 3.11: (a) Distribution of number of intra and inter-helical H-bonds. **(b)** Distribution of number of H-bonds between peptide and water. Inset shows the distribution of number of water molecules forming more than one H-bond with peptide chain. The distribution plots for parallel and anti-parallel alignments are shown by solid and dashed line respectively.

We observed penetration of water in the preassembled peptide structure (Figure 3.7). We showed that the number of waters is different at different regions of the monolayer. In peptide self-assembly, water may play a role of further stabilizing the self-assembled structure. To understand the contribution of water molecules in stabilizing the lamellar structures, we computed the number of H-bonds among water and peptide chains. The number of H-bonds was calculated only when the water molecules penetrated the monolayer structure. The number of H-bonds between side chains and water is higher than between main chains and water in both parallel and antiparallel alignment (Figure 3.11b). In parallel alignment, the total number of peptide−water H-bonds is higher than in antiparallel alignment. Water may also act as a bridge between two peptide chains. So, the number of water molecules forming more than one H-bond (bridging water molecules) simultaneously with peptide chains (inset of Figure 3.11b) has been calculated. Almost 50% of the water−peptide H-bonds are bridging H-bonds for both systems. A water-mediated H-bond network is shown in Figure 3.12 for both alignments which is extracted from real time trajectory. This large count of peptide−water bridging H-bonds is indicative of ordering of water inside the lamella and stabilizing the lamellar structure of the self-assembly.

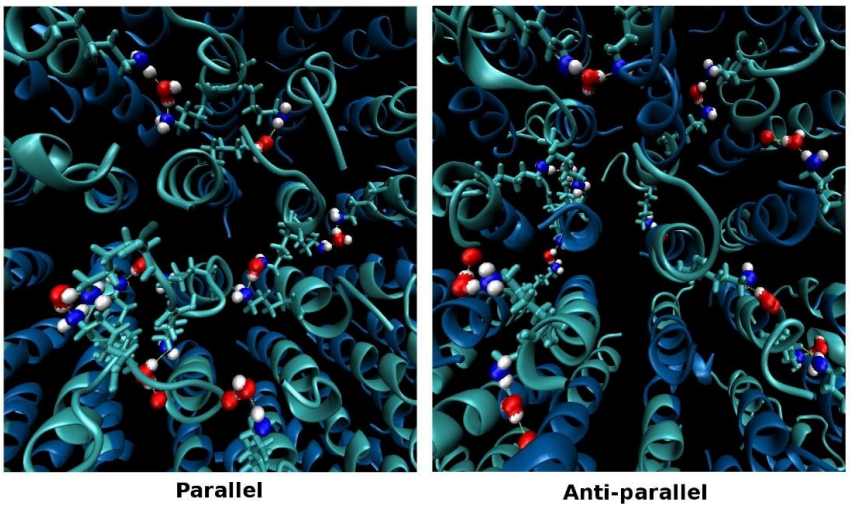


Figure 3.12: Snapshots extracted from real time trajectory showing water mediated H-bonded network. LEU and LYS blocks are represented by blue and cyan respectively, side-chains of LYS-blocks are represented as cyan coloured stick. Main chain O, side chain amine H and side chain N are represented by red, white and blue colours respectively. Water O and H are represented by red and white colours respectively.

3.3.5 Hydrogen Bonded Network

A self-assembled structure remains stable when interconnected H-bonds are formed in three dimensions and a network of H-bonds is present. We have observed a large number of H-bonds in pure peptide-derived self-assembled structure. However, it remains unclear whether there is a formation of a H-bonded network in three dimensions. Therefore, we have further investigated whether these H-bonds act in a cooperative manner to give rise to a Hbonded network, i.e., translated along the consecutive peptide chains. If a network is present, then to what extent? For this, we first computed the number of interhelical H-bonds along different directions (parallel, perpendicular, and diagonal) in the plane of the lamella (XY-plane). These three directions were determined from the distribution of angles between vectors joining the center of masses of two H-bonded peptide chains with the Y-axis (Figure 3.13). The vector was considered to be parallel if its angle with the Y-axis fell within 0−25°, diagonal if 25−65°, and perpendicular if greater than 65°. The contribution toward stability from the two types of Hbond in these different directions is obtained from calculation of total H-bond energies between peptide chains. In the calculation of H-bond energy, it is considered as purely electrostatic interaction between hydrogen and H-bond acceptor. The energies of

each type of H-bonds were calculated and converted into distributions for individual contributions. Distributions are plotted (Figure 3.14a and 3.14b) by normalizing the histogram by the number of frames in the last 10 ns of the final 100 ns NPT run. For parallel alignment, the contribution of side-chain−side-chain H-bonds is higher than that of side-chain−main-chain H-bonds (Figure 3.14a). The contribution toward H-bonded energy of these two types of H-bonds is found to be higher along parallel and perpendicular directions with respect to the Y-axis compared to a diagonal direction. In antiparallel monolayers, the contribution toward hydrogen bond energies of the two types of H-bonds in different directions is found to be similar except for a very high contribution of side-chain−side-chain H-bonds in a diagonal direction (Figure 3.14b). This shows that side-chain−side-chain H-bonds along a diagonal direction keeps the integrity of antiparallel monolayer structure.

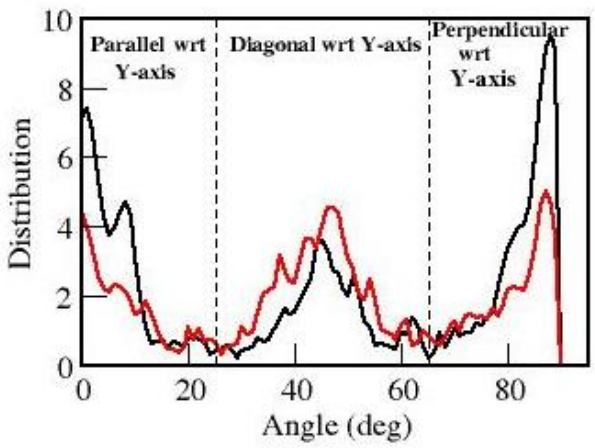


Figure 3.13: Plot for distribution of angle between line connecting centre of mass of hydrogen bond forming peptide chains and Y-axis for parallel (black line) and anti-parallel alignment (red line). The three distinct direction of formed H-bonds are clearly appears from distribution plot and shown by vertical dotted lines.

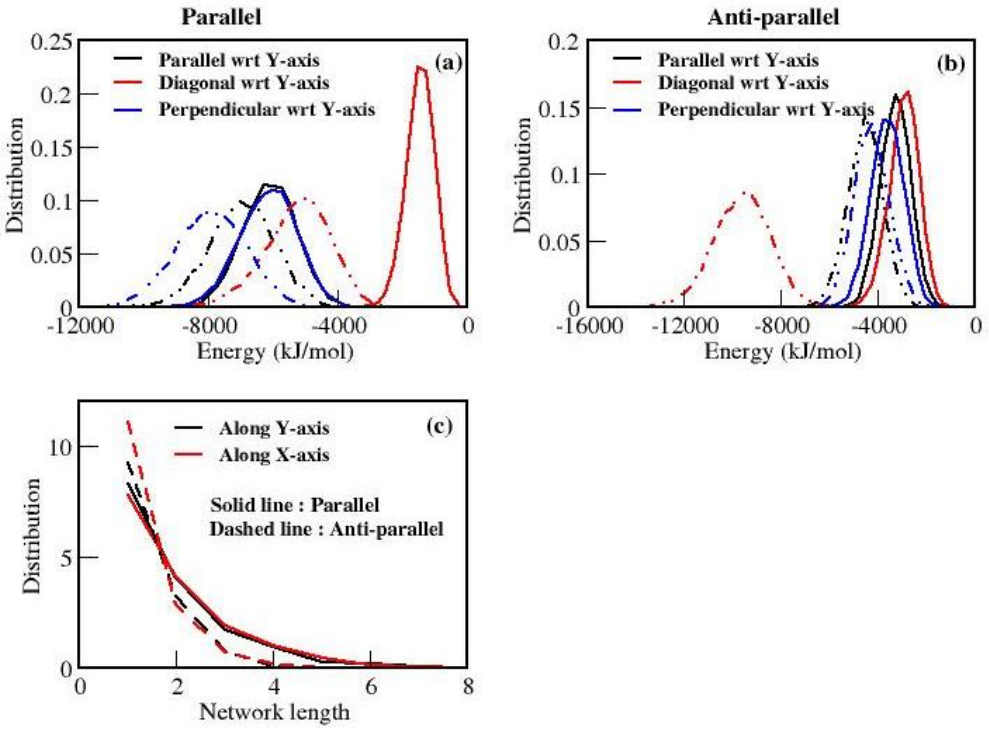


Figure 3.14: Distribution of two different types H-bond energy in different direction with respect to Y-axis for **(a)** Parallel monolayer, **(b)** Anti-parallel monolayer. In figure a and b, mainchain-sidechain H-bonds and sidechain-sidechain H-bonds are represented by solid and dashed-dotted lines respectively. **(c)** Distribution of H-bonded network length along X and Y direction for both types of monolayer. The network lengths for parallel and anti-parallel alignments are shown by solid and dashed lines respectively.

The number of H-bonds along the different directions definitely suggests the possibility of H-bonded network structure in the plane of the peptide lamella. Formation of a H-bonded network can further stabilize the lamella. Thus, the H-bonded network length is envisaged. That is, how many consecutive peptide chains are H-bonded to give rise to a network? To calculate this, the plane of peptide lamella (XY-plane) was first divided into 1 nm × 1 nm grids. The consecutive H-bonded (side-chain−side-chain and main-chain−side-chain both considered) and peptide chains were counted along the grids in X- and Y-directions. A network break was considered if two consecutive peptide chains were not H-bonded. Network lengths were counted in each frame of the last 10 ns of the trajectory and converted into a histogram. Occurrences of the network length are plotted in Figure 3.14c, normalizing over the number of frames. The network lengths are similar in the Y-direction and in the X-

direction for both parallel and antiparallel alignment. For parallel alignment, a network length up to 6 is found, which suggests that out of 8 peptide chains in the Y-direction, 6 are in a continuous network. The network length is found to be higher for parallel alignment than for antiparallel alignment (network length up to 4).

Thus, we may infer that the interhelical H-bonded network in the plane of peptide lamellae arises due to interaction among side chains. This network structure is responsible for the stability of lamellar structures constructed from diblock copolypeptides aligned in parallel and antiparallel fashion.

3.3.6 Quantification of Stability: PMF Calculations

The PMF with COM distance between reference and pulled group has been calculated for different angles of orientation between them (Figure 3.6). The corresponding free energy of binding has been evaluated from the plot for all of these angles of orientations and is shown in Figure 3.15. The free energy of binding is found to be the lowest (−34 kJ/mol) at a 0° angle of orientation, and it is the highest (−8.78 kJ/mol) at 120° and again reaches a lower value (−28.2 kJ/mol) at 180° (antiparallel alignment). This shows that binding of two diblock polypeptide chains are more favorable in their parallel and antiparallel alignments.

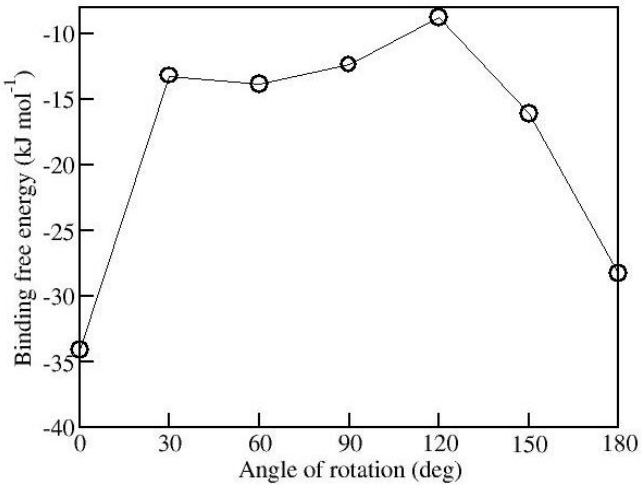


Figure 3.15: Free energy of binding at different angle of orientation between pulled and reference chain.

The simulation of copolypeptide monolayers suggested that the integrity of parallel and antiparallel lamellar structures remain intact even after the extensive three-step simulation protocol. Stabilities of the lamellar structures are found to arise because of H-bonded network structure. Further, the stabilities of parallel and antiparallel monolayers can be quantified and compared more realistically by taking account of the interactions between peptide chains which are present diagonally. Thus, PMF calculations are performed by pulling one helical chain from a bunch of 16 chains along the X-direction for both parallel and antiparallel alignment. The detailed procedure for PMF calculation is discussed in Computational Methods. A plot of PMF as a function of distance between COM of the pulled chain and reference group is provided in Figure 3.16a with error bars. Snapshots at different points of COM pulling are depicted in Figure 3.16b. Minima of the PMF appear at a similar position (∼0.7 nm) for both parallel and antiparallel alignment. A peak is observed near 1 nm of COM distance which is prominent for parallel alignment (4.83 kJ/mol). This peak represents short-range attractive interaction between the pulled chain and reference group. This attractive interaction can be visualized in the snapshot near the peak minima where the LEU block is pulled away and the LYS block still stays near the lamella (Figure 3.16b). The barrier, represented by the peak, needs to be crossed to overcome this short-range attractive interaction. For antiparallel alignment, a corresponding small hump of −1.47 kJ/mol appears near 1 nm. Following this, almost no difference is observed in the PMF−COM distance plot for both alignments (Figure 10a). This indicates that the long-range interaction between reference and pulled chain is similar for parallel and antiparallel alignments. Binding free energies obtained from the PMF calculations are −71.84 kJ/mol and −66.55 kJ/mol for parallel and antiparallel systems, respectively.

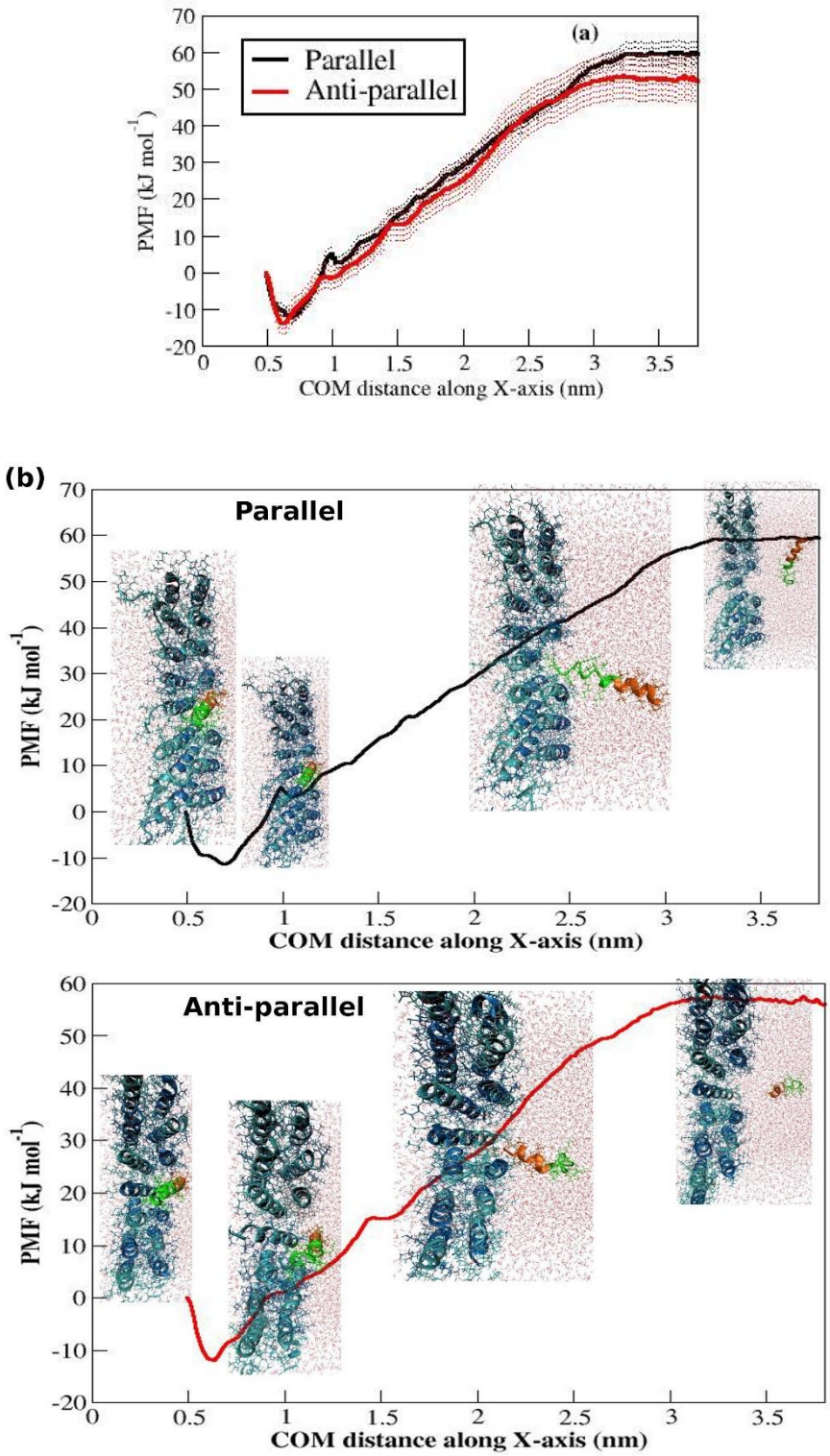


Figure 3.16: (a) PMF for unbinding of one di-block α-helical from rest of the system along pulling axis (X-axis) for parallel and anti-parallel alignment. (b) Snapshots after umbrella sampling simulation at different position along pulling axis (X-axis) for parallel and anti-parallel alignment.

3.4 Conclusions

MD simulations have been performed at 300 K using an NPT ensemble for diblock copolypeptides with hydrophobicity mismatch preassembled in lamellar structures. The copolypeptides were arranged in parallel (LYS blocks packed close to each other) and antiparallel (LYS blocks packed close to LEU blocks) alignments to construct the lamellas. Water molecules were added on both sides of the lamella. Performing a threestep simulation protocol for the lamella−water systems shows that both the parallel and antiparallel alignment are not disrupted. That is, the lamella remains stable for both parallel and antiparallel alignment of copolypeptides. We have further analyzed the systems to gain structural insight into the system. We have concluded that primarily the interhelical H-bonding is responsible for the stability of the lamellar structures in both alignments. Bridged H-bonds also play a role in further stabilizing structure. We have further envisaged the H-bonded network structures formed in both alignments, which are again responsible for higher stability. Stabilities of the parallel and antiparallel alignments were further compared by free energy calculations. They indicate that the degrees of stability for both systems are similar.

References

[1] ChÃ©cot, F., Lecommandoux, S., Klok, H. A., and Gnanou, Y., "From Supramolecular Polymersomes to Stimuli-Responsive Nano-Capsules Based on Poly(Diene-B-Peptide) Diblock Copolymers," The European Physical Journal E 10, 25-35 (2003).

[2] Cai, C., Zhang, L., Lin, J., and Wang, L., "Self-Assembly Behavior of Ph- and Thermosensitive Amphiphilic Triblock Copolymers in Solution: Experimental Studies and Self-Consistent Field Theory Simulations," The Journal of Physical Chemistry B 112, 12666-12673 (2008).

[3] Cai, C., Lin, J., Chen, T., Wang, X.-S., and Lin, S., "Super-Helices Self-Assembled from a Binary System of Amphiphilic Polypeptide Block Copolymers and Polypeptide Homopolymers," Chemical Communications 2709-2711 (2009).

[4] Naik, S. S., Ray, J. G., and Savin, D. A., "Temperature- and Ph-Responsive Self-Assembly of Poly(Propylene Oxide)-B-Poly(Lysine) Block Copolymers in Aqueous Solution," Langmuir 27, 7231-7240.

[5] Breedveld, V., Nowak, A. P., Sato, J., Deming, T. J., and Pine, D. J., "Rheology of Block Copolypeptide Solutions:Â€‰ Hydrogels with Tunable Properties," Macromolecules 37, 3943-3953 (2004).

[6] Nowak, A. P., Breedveld, V., Pine, D. J., and Deming, T. J., "Unusual Salt Stability in Highly Charged Diblock Co-Polypeptide Hydrogels," Journal of the American Chemical Society 125, 15666-15670 (2003).

[7] Gazit, E., "Self-Assembled Peptide Nanostructures: The Design of Molecular Building Blocks and Their Technological Utilization," Chemical Society Reviews 36, 1263-1269 (2007).

[8] Cui, H., Webber, M. J., and Stupp, S. I., "Self-Assembly of Peptide Amphiphiles: From Molecules to Nanostructures to Biomaterials," Peptide Science 94, 1-18.

[9] Reches, M., and Gazit, E., "Molecular Self-Assembly of Peptide Nanostructures: Mechanism of Association and Potential Uses," Current Nanoscience 2, 105-111 (2006).

[10] Bellomo, E. G., Wyrsta, M. D., Pakstis, L., Pochan, D. J., and Deming, T. J., "Stimuli-Responsive Polypeptide Vesicles by Conformation-Specific Assembly," Nat Mater 3, 244-248 (2004).

[11] Vauthey, S., Santoso, S., Gong, H., Watson, N., and Zhang, S., "Molecular Self-Assembly of Surfactant-Like Peptides to Form Nanotubes and Nanovesicles," Proceedings of the National Academy of Sciences 99, 5355-5360 (2002).

[12] RodrÃ-guez-HernÃ¡ndez, J., and Lecommandoux, S. b., "Reversible Insideâˆ’out Micellization of Ph-Responsive and Water-Soluble Vesicles Based on Polypeptide Diblock Copolymers," Journal of the American Chemical Society 127, 2026-2027 (2005).

[13] Chécot, F., Lecommandoux, S., Gnanou, Y., and Klok, H.-A., "Water-Soluble Stimuli-Responsive Vesicles from Peptide-Based Diblock Copolymers," Angewandte Chemie International Edition 41, 1339-1343 (2002).

[14] Kukula, H., Schlaad, H., Antonietti, M., and Forster, S., "The Formation of Polymer Vesicles or Â€Œpeptosomesâ€ by Polybutadiene-Block-Poly(L-Glutamate)S in Dilute Aqueous Solution," Journal of the American Chemical Society 124, 1658-1663 (2002).

[15] Lee, O.-S., Stupp, S. I., and Schatz, G. C., "Atomistic Molecular Dynamics Simulations of Peptide Amphiphile Self-Assembly into Cylindrical Nanofibers," Journal of the American Chemical Society 133, 3677-3683.

[16] Velichko, Y. S., Stupp, S. I., and de la Cruz, M. O., "Molecular Simulation Study of Peptide Amphiphile Self-Assembly," The Journal of Physical Chemistry B 112, 2326-2334 (2008).

[17] Lee, O.-S., Cho, V., and Schatz, G. C., "Modeling the Self-Assembly of Peptide Amphiphiles into Fibers Using Coarse-Grained Molecular Dynamics," Nano Letters 12, 4907-4913.

[18] Thota, N., Luo, Z., Hu, Z., and Jiang, J., "Self-Assembly of Amphiphilic Peptide (Af)6h5k15: Coarse-Grained Molecular Dynamics Simulation," The Journal of Physical Chemistry B 117, 9690-9698.

[19] Thota, N., and Jiang, J., "Self-Assembly of Amphiphilic Peptide (Af)6h5k15 Derivatives: Roles of Hydrophilic and Hydrophobic Residues," The Journal of Physical Chemistry B 118, 2683-2692.

[20] Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P., and de Vries, A. H., "The Martini Force Field:Â€‰ Coarse Grained Model for Biomolecular Simulations," The Journal of Physical Chemistry B 111, 7812-7824 (2007).

[21] Rodgers, J. M., SÃ¸rensen, J., de Meyer, F. d. r. J. M., SchiÃ¸tt, B., and Smit, B., "Understanding the Phase Behavior of Coarse-Grained Model Lipid Bilayers through Computational Calorimetry," The Journal of Physical Chemistry B 116, 1551-1569.

[22] Crossland, E. J. W., Kamperman, M., Nedelcu, M., Ducati, C., Wiesner, U., Smilgies, D. M., Toombes, G. E. S., Hillmyer, M. A., Ludwigs, S., Steiner, U., and Snaith, H. J., "A Bicontinuous Double Gyroid Hybrid Solar Cell," Nano Letters 9, 2807-2812 (2008).

[23] Stadler, R., Auschra, C., Beckmann, J., Krappe, U., Voight-Martin, I., and Leibler, L., "Morphology and Thermodynamics of Symmetric Poly(a-Block-B-Block-C) Triblock Copolymers," Macromolecules 28, 3080-3097 (1995).

[24] Khandpur, A. K., Foerster, S., Bates, F. S., Hamley, I. W., Ryan, A. J., Bras, W., Almdal, K., and Mortensen, K., "Polyisoprene-Polystyrene Diblock Copolymer Phase Diagram near the Order-Disorder Transition," Macromolecules 28, 8796-8806 (1995).

[25] Wimley, W. C., Creamer, T. P., and White, S. H., "Solvation Energies of Amino Acid Side Chains and Backbone in a Family of Hostâ^'Guest Pentapeptidesâ€ " Biochemistry 35, 5109-5124 (1996).

[26] Wimley, W. C., and White, S. H., "Experimentally Determined Hydrophobicity Scale for Proteins at Membrane Interfaces," Nat Struct Mol Biol 3, 842-848 (1996).

[27] Kyte, J., and Doolittle, R. F., "A Simple Method for Displaying the Hydropathic Character of a Protein," Journal of Molecular Biology 157, 105-132 (1982).

[28] Argos, P., Rao, J. K. M., and Hargrave, P. A., "Structural Prediction of Membrane-Bound Proteins," European Journal of Biochemistry 128, 565-575 (1982).

[29] Holowka, E. P., Pochan, D. J., and Deming, T. J., "Charged Polypeptide Vesicles with Controllable Diameter," Journal of the American Chemical Society 127, 12423-12428 (2005).

[30] Dhasaiyan, P., Pandey, P. R., Visaveliya, N., Roy, S., and Prasad, B. L. V., "Vesicle Structures from Bolaamphiphilic Biosurfactants: Experimental and Molecular Dynamics Simulation Studies on the Effect of Unsaturation on Sophorolipid Self-Assemblies," Chemistry – A European Journal 20, 6246-6250 (2014).

[31] Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E., "Gromacs 4:Â€‰ Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation," Journal of Chemical Theory and Computation 4, 435-447 (2008).

[32] Jorgensen, W. L., and Gao, J., "Monte Carlo Simulations of the Hydration of Ammonium and Carboxylate Ions," The Journal of Physical Chemistry 90, 2174-2182 (1986).

[33] Jorgensen, W. L., Maxwell, D. S., and Tirado-Rives, J., "Development and Testing of the Opls All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids," Journal of the American Chemical Society 118, 11225-11236 (1996).

[34] Kaminski, G. A., Friesner, R. A., Tirado-Rives, J., and Jorgensen, W. L., "Evaluation and Reparametrization of the Opls-Aa Force Field for Proteins Via Comparison with

Accurate Quantum Chemical Calculations on Peptidesâ€ " The Journal of Physical Chemistry B 105, 6474-6487 (2001).

[35] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L., "Comparison of Simple Potential Functions for Simulating Liquid Water," The Journal of Chemical Physics 79, 926-935 (1983).

[36] Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R., "Molecular Dynamics with Coupling to an External Bath," The Journal of Chemical Physics 81, 3684-3690 (1984).

[37] Bussi, G., Donadio, D., and Parrinello, M., "Canonical Sampling through Velocity Rescaling," The Journal of Chemical Physics 126, 014101 (2007).

[38] Patra, M., Karttunen, M., Hyvönen, M. T., Falck, E., Lindqvist, P., and Vattulainen, I., "Molecular Dynamics Simulations of Lipid Bilayers: Major Artifacts Due to Truncating Electrostatic Interactions," Biophys. J. 84, 3636-3645 (2003).

[39] Williams, G., and Watts, D. C., "Non-Symmetrical Dielectric Relaxation Behaviour Arising from a Simple Empirical Decay Function," Transactions of the Faraday Society 66, 80-85 (1970).

[40] Hub, J. S., de Groot, B. L., and van der Spoel, D., "G_Whamâ€"a Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates," Journal of Chemical Theory and Computation 6, 3713-3720.

[41] Kemp, J. P., and Chen, J. Z. Y., "Helical Structures in Proteins," Biomacromolecules 2, 389-401 (2001).

[42] Errington, J. R., and Debenedetti, P. G., "Relationship between Structural Order and the Anomalies of Liquid Water," Nature 409, 318-321 (2001).

Chapter 4

Transferability of Different Classical Force Fields for Right and Left Handed α-Helices Constructed from Enantiomeric Amino Acids

4.1 Introduction

Chirality is common in nature at the molecular level. Molecular chirality arises from the lack of symmetry elements in the molecular structures. Building block of proteins, that is amino acids, are chiral and contain the chiral centre at the alpha carbon atom (except glycine) giving rise to possibility of D and L enantiomers. Further, in addition to molecular chirality, chirality in shape is also common in nature. Left and right handed helical configuration contain axis of chirality instead of chiral centre. In nature, the A and B forms of DNA form right handed helices, Z form of DNA forms left handed helices. Also, among the various secondary structure of proteins, alpha helix is the most commonly encountered secondary structure [1]. Right and left handed helices are defined depending upon positive or negative rotational translation of sequence of atoms in the peptide chain with respect to the helical axis [2]. Interestingly, naturally occurring L-alpha amino acids form right handed helices more preferentially due to steric repulsion between side chain and main chain carbonyl group [3]. Also, left-handed helices from D-amino acids are energetically more favorable over right handed helices [4]. Although, left handed helices are not common in nature, they contribute in stability of some proteins, and also in some cases play a role in ligand binding [5]. Also, in peptide design, addition of D-amino acid to the peptide sequence makes it less susceptible to enzymatic cleavage, resulting in enhanced biocompatibility of the designed peptide [6, 7].

Conformational preferences of peptide chains constructed of D and L enantiomers of amino acids have drawn plenty of research interest. Colonna-Cesari et al computed conformational energies of several poly-D-L-peptide helices [8]. Nanda et al reported a Monte Carlo simulation based method to design scaffolds for interacting with left handed helices [9]. Novotny et al have surveyed left handed helices in different protein structures [5]. Brown et al have shown that oligomer of 2- aminoisobutyric acid (a synthetic amino acid) adopts left or right handed 3_{10} helical conformation depending upon the natural amino acids present at the terminal of the peptide chain [10]. Left and right handed helices from peptidic foldamers have also been observed and studied [11]. To model protein structure and dynamics, quantification of conformational energies are important.

Conformational energies can be computed by considering interactions between atoms. For this, quantum chemical calculations are accurate and useful. Coupled-cluster method (CCSD(T) [12]) and the second-order Moller-Plesset perturbation (MP2) methods are often

considered for correctly accounting van der Waals (non-bonded) and all other interactions, but these methods are computationally very expensive at the scale of $O(N^7)$ and $O(N^5)$ respectively. Density functional theory (DFT) methods using B3LYP [13] and PBE [14] functional are also used but they are not the best for accounting non-bonded interactions. Zhao et al developed M06 suite of functional which is a set of four meta-hybrid GGA DFT functionals, and they recommend usage of M06 functional for better estimation of non-covalent interactions [15, 16]. They have calculated conformational energies of proteins using different functional like MO6, BLYP, B3LYP, PBE. Their study revealed that the M06 suite of functionals provide best accuracy for non-covalent interaction [17]. The M06 suite of functional has been performed for calculating non-covalent interactions in biomolecules and model systems of hydrogen bonding interactions [18-20]. Jiang et al studied 20 tetra peptides in 5 conformations (right-handed helix, left handed helix, beta-sheet, antiparallel beta-sheet and polyproline II) as structural model to investigate the relative conformation energies at the MP2/cc-pVTZ/B3LYP/6-31G** level [21]. They found that the energetic patterns are similar for amino acids of same class (i.e., hydrophobic, aromatic, polar), and the patterns alter when they belong to different class. They used MP2 energy as a standard to statistically evaluate the overall performance of other different functionals like MO5-2X, PBE, B3LYP and different classical force fields. They reported that the MO5-2x functional are better than PBE, B3LYP functional. Quantum chemical calculations have also been helpful in revealing the effect of thermodynamic contributions, solvents effects, and charge of the side chain on conformational energies of small peptide molecules [22-24].

All atomistic molecular dynamics (MD) simulation technique is one of the major tools to decipher structure, dynamics, and function of bio-molecules. Structural level information on the fluctuations and conformational changes of biological molecules have been successfully derived from this technique. However, one of the major limitations of MD simulations is the choice of reliable parameters for potential energy function (force field). For studying biological systems the commonly used force fields are OPLS [25, 26], AMBER [27], CHARMM [28] and GROMOS [29, 30]. Although the energy functions used in different force fields are very similar, their parameters differ due to different ways of parameterization to achieve the values of target properties. Different force fields have preferences to favor certain types of secondary structures [31, 32]. A good amount of effort have been invested for validation of force-fields. Best et al. explored the Ramachandran space sampled by Ala_5

peptides using Amber03, Amber99SB, AMBERGS [33], OPLS/AA, GROMOS53a6 [30] and GROMOS43a1, CHARMM27/CMAP [34], and by comparing with J-coupling NMR experiments they found that GROMOS43a1, AMBER03, OPLS/AA and CHARMM27/CMAP are good force fields [35]. Lange et al found that AMBER99SB is the best force field among OPLS/AA, AMBER03, AMBER99SB, GROMOS43a1, GROMOS53a6 and CHARMM22 after comparing results from microsecond simulations of ubiquitin and the gb3 domain of protein G with their NMR data [36]. Cino et al simulated a β-hairpin forming peptides using 10 different force fields and concluded that GROMOS-53a6 and GROMOS-43a1p perform better than the other force fields [37]. Lindroff-Larsen et al systematically validated Amber03, Amber03* [38], OPLS/AA, CHARMM22, CHARMM22* [38], AMBER99SB*-ildn [39], CHARMM27, AMBER99SB-ildn protein force fields against experimental data and concluded that AMBER99SB*-ildn or CHARMM22* performs better than other force fields [40]. Unwinding mechanisms of poly-L-ALA, poly-L-LEU, poly-L-LYS, and poly-L-GLU helical peptides in water have been investigated in the past as a function of terminal groups using OPLS/AA force field [41, 42]. Distinctions observed in the unwinding mechanisms from OPLS/AA force field were also validated from AMBER force field. All of these studies validate different force fields and are guide to choose particular force field to address a problem of interest. But, most of these studies have been performed on naturally occurring L-amino acids. Although most of the α-helices are right-handed and made up of L-amino acids, other conformations can also exist. These are right-handed α-helices ($\alpha_{R)}$ consisting of D-amino acids, left-handed α-helices ($\alpha_{L)}$ consisting of D-amino acids as well as L-amino acids. Therefore, it is essential to investigate the robustness of these recent force fields to explain other possible conformations.

Further, peptide block copolymers can self-assemble to form variety of structures - vesicles, lamella, rods [43-47]. Since the peptide block copolymeric chains are biocompatible, they have useful application in tissue engineering and drug delivery [48]. We have reported the structural insights into the self-assembled configuration of block copolypeptides in our previous study [49]. We observed that, block copolypeptides consisting of hydrophobic and hydrophilic blocks may assemble into both parallel (similar blocks face to face), anti-parallel (dissimilar blocks face to face). However, we only considered α_R from L-amino acids. Pati et al. have shown that it is possible to control the helical structures obtained from polymerizing racemic glycopeptides [50]. They have shown the variation of the helical conformations using CD spectral analysis. Their study on the racemic glycopeptides can further be extended

to racemic peptide chains, i.e. peptide chains constructed of combinations of D- and L- amino acids. Molecular simulations could be useful at this to decipher the structural features that guide the self-assembled helical or non-helical conformation. But before investigating that, it is imperative to realize, which of the commonly used force field performs best in representing helical conformation obtained from enantiomeric amino acids.

In the present study, we have investigated the energetic and structural features of α_R and α_L homomeric polypeptides of two different chain lengths (5-mer and 10-mer) consisting of enantiomeric (D, L) amino acids. Homeric helical chains of poly-alanine (ALA), poly-leucine (LEU), poly-lysine (LYS) and poly-glutamic acid (GLU) are considered for this study. These four residues belong to different classes (hydrophobic or hydrophilic) of amino acids. ALA and LEU both are the hydrophobic amino acids with different length of side chains (methyl, t-butyl groups respectively). These are taken because of their higher helix forming propensity compare to other amino acids with hydrophobic alkyl side chain [51]. LYS and GLU both are the hydrophilic amino acids. The side chains of LYS and GLU are considered in their neutral forms. We have focused mainly on the four secondary structures of peptides - α_R from L-isomer (R-H-L), α_R from D-isomer (R-H-D), α_L from L-isomer (L-H-L), and α_L from D-isomer (L-H-D). We have studied total 32 α-helical polypeptides (4 amino acids × 4 conformations of 5-mer and 10-mer).

All these helical conformations are optimized quantum mechanically (QM) using M06/6-311g (d, p) level of theory in vacuum and in presence of water (dielectric constant, ε=78) using polarizable continuum model (PCM) [52]. Further, we have performed MD simulation of all these conformations in vacuum, and in water using OPLS/AA [26], CHARMM27/CMAP [53] and AMBER03 [54] force fields. This way, using QM calculations as a reference, we have investigated the ability of different force fields to reproduce the structural features of different helical conformations constructed of enantiomeric amino acids. The primary objective of the present study is to evaluate the force field that can describe all the helical conformations constructed of enantiomeric amino acids. To our best of knowledge, this is the first time when α_R and α_L homomeric polypeptides from enantiomeric (D, L) amino acids have been studied using density function theory (DFT) as well as the MD simulation. We have analyzed simulation trajectories for structure and energetics to provide a conclusion on the best performing force field for peptides with different chirality. For evaluation of structural properties we have considered phi and psi angles, Hydrogen bonds

and different long range helix order parameters. In this work we have performed simulations in large numbers to represent out results with statistical errors.

4.2 Computational Details

Amino acid residues (NH_2–Cα(H)(R)–COOH, R = side chain) can have two enantiomers (D,L) due to the presence of the chiral Cα atom. In the present study, L- and D-isomers of ALA, LEU, LYS and GLU were considered to construct capped homomeric α_R and α_L polypeptides. The initial polypeptides were built by taking backbone dihedral angles (Φ,Ψ) -64.0° and -43.0° respectively [55] for α_R conformer, and at +57.0° and +47.0° respectively for α_L conformer. The C-terminal and N-terminal of α-helical polypeptide chains were capped with $–NHCH_3$ and $–COCH_3$ groups respectively. The side chains of LYS and GLU amino acid residues were taken as neutral, that is $–CH_2$-$(CH_2)_3$-NH_2 and $–CH_2$-CH_2-COOH respectively. 5- and 10-mer homopolymeric right and left handed helical chains were generated from all the enantiomeric amino acids. The χ_1 torsions (N-C_α-C_β-C_γ) taken for the LEU, LYS, GLU residues are 180°. All the above mentioned α-helical polypeptides chains were subjected to QM optimization, followed by classical MD in water, and in vacuum separately. The QM optimizations were performed using M06/6-311g (d, p) level of theory and the classical MD simulations were performed using OPLS/AA, CHARMM27/CMAP and AMBER-03 force fields. Details of the QM calculations and the classical MD simulations are as follows.

4.2.1 QM optimization of helical peptides

All the α-helical polypeptide chains (5- and 10-mer of R-H-L, R-H-D, L-H-L, and L-H-D) were QM optimized using M06 hybrid functional and 6-311G (d, p) basis set in vacuum and in presence of water. The tight convergence criterion of 10^{-5} a.u. was used for optimization. All the QM optimizations were performed using Gaussian 09 software package [56]. No constraints were used during their QM optimization. The effect of water was included by using polarizable continuum model (PCM) [57]. The optimized geometries of α-helical polypeptides were verified by calculating vibrational frequencies. No imaginary frequency was found which indicated that all the optimized conformers were in local minima of the potential energy surface. The zero-point energy correction and basis set superposition error

(BSSE) were also evaluated for all the α-helical polypeptides. The use of small basis-set with any type of DFT functionals overestimates intermolecular interaction. Therefore, the correction of BSSE is required to calculate accurate energy. The BSSE corrections have been implemented using Boys-Bernardi CP method[58] where each amino acid residues including capping ones are considered as a fragment.

4.2.2 Details of Classical MD Simulations

Following QM optimization of all the α-helical polypeptide chains (5- and 10-mer of R-H-L, R-H-D, L-H-L, and L-H-D), all-atomistic MD simulations were performed using Gromacs-4.5.5 (double precision) code.[59] All the α-helical polypeptide chains were simulated separately in vacuum and in water using TIP3P water model.[60] All helical chains (5- and 10-mer of R-H-L, R-H-D, L-H-L, and L-H-D) were modeled individually using OPLS/AA, CHARMM27/CMAP and AMBER-03 force fields. In vacuum, all the α-helical polypeptides were initially energy minimized using steepest descent algorithm which was followed by conjugate gradient method to optimize initial geometry. The convergence criterion of 0.001 kJ $mol^{-1}nm^{-1}$ was used for energy minimization in vacuum. No cut-off value for Lennard-Jones (LJ) as well as electrostatic interaction was considered so that complete intramolecular non-bonded interactions were applied during the simulation in vacuum. Also, the simulation in vacuum was performed in absence of periodic boundary condition to eliminate the possibility of interaction with periodic image. MD simulation run for 1 ns was performed for all the α-helical peptides (5- and 10-mer of R-H-L, R-H-D, L-H-L, and L-H-D). The temperature was maintained at 300 K by using V-rescale thermostat applying 0.1 ps coupling constant.

Further, MD simulations were carried out in presence of water for 5- and 10-mer of R-H-L, R-H-D, L-H-L, and L-H-D separately. To generate the starting conformation for MD simulation, the α-helical polypeptide chain was placed in simulation box maintaining certain minimum distance with the walls of the simulation box. This minimum distance varied across the various peptide chains such that its value is 0.7 nm plus the end-do-end distance of polypeptide chain. This was done to take into account the whole intramolecular non-bonded interaction within the polypeptide chain during the MD simulation. The distance between nitrogen of C-terminal capping ($–NHCH_3$) group and carbonyl carbon of N-terminal capping ($–COCH_3$) group was considered as the end-to-end distance. The minimum distance between wall of simulation box and polypeptide chain, and end-to-end distances of 10-mer and 5-mer

of ALA, LEU, GLU and LYS are listed in Table 4.1, 4.2 respectively. Then polypeptide chains were solvated with TIP3P water followed by energy minimized using steepest descent algorithm. 1ns of MD simulation was carried out using isothermal –isobaric (NPT) ensemble keeping backbone atoms of the peptide position restrained to relax the side chain atoms and to equilibrate water molecules around peptide chain. After removing the position restraint of the backbone, final NPT simulation was performed for 1ns. Long range electrostatic interactions were taken care of using particle mesh Ewald (PME) method [61]. The value of cut-off lengths for electrostatic and LJ interactions were taken as greater than the value of end-do-end distance so that it can take care of whole intramolecular non-bonded interactions within the peptide chain. The values of the cut-off lengths taken during MD simulation of 10-mer and 5-mer polypeptides in water are provided in Table 4.1, 4.2, respectively. The minimum distances with periodic image was checked for each simulation and found to be greater than the corresponding cut-off distances throughout the simulation which confirms no interaction between polypeptide chain and its periodic image. Temperature was maintained at 300 K by V-rescale thermostat using coupling constant of 0.1 ps. Pressure was kept constant at 1 bar using Berendsen barostat [62] with coupling constant of 1 ps. All the bonds were kept unconstrained for simulation in both vacuum and in water. The time step of 1 fs was used for all simulations. Trajectories were saved at every 1 ps. All the MD simulations were repeated for three times in vacuum as well as in water medium.

Table 4.1: End to end distance, cut-off's distance for Lennar-Jones and electrostatic interaction, and minimum distance between peptide and the boundary of simulation box for 10-mer of ALA, LEU, LYS and GLU in water.

Systems	Conformers	End to end distance (nm)	Cut-off's (nm) (rcoulomb and rvdw)	Minimum distance between peptide and the box (nm)
ALA-10-mer	R-H-L	1.466	2.1	2.2
	R-H-D	1.458	2.1	2.2
	L-H-L	1.482	2.1	2.2
	L-H-D	1.478	2.1	2.2
LEU-10-mer	R-H-L	1.443	2.1	2.2

LEU-10-mer	R-H-D	1.436	2.1	2.2
	L-H-L	1.452	2.1	2.2
	L-H-D	1.443	2.1	2.2
LYS-10-mer	R-H-L	1.442	2.2	2.3
	R-H-D	1.421	2.2	2.3
	L-H-L	1.454	2.2	2.3
	L-H-D	1.452	2.2	2.3
GLU-10-mer	R-H-L	1.443	2.2	2.3
	R-H-D	1.431	2.2	2.3
	L-H-L	1.470	2.2	2.3
	L-H-D	1.463	2.2	2.3

Table 4.2: End to end distance, cut-off's distance for Lennar-Jones and electrostatic interaction, and minimum distance between peptide and the boundary of simulation box for 5-mer of ALA, LEU, LYS and GLU in water.

Systems	Conformers	End to end distance (nm)	Cut-off's (nm) (rcoulomb and rvdw)	Minimum distance between peptide and the box (nm)
ALA-5-mer	R-H-L	0.778	1.3	1.4
	R-H-D	0.756	1.3	1.4
	L-H-L	0.779	1.3	1.4
	L-H-D	0.787	1.3	1.4
LEU-5-mer	R-H-L	0.775	1.4	1.5
	R-H-D	0.760	1.4	1.5
	L-H-L	0.756	1.4	1.5

LEU-5-mer	L-H-D	0.776	1.4	1.5
LYS-5-mer	R-H-L	0.772	1.4	1.5
	R-H-D	0.782	1.4	1.5
	L-H-L	0.783	1.4	1.5
	L-H-D	0.781	1.4	1.5
GLU-5-mer	R-H-L	0.786	1.3	1.4
	R-H-D	0.769	1.3	1.4
	L-H-L	0.795	1.3	1.4
	L-H-D	0.788	1.3	1.4

4.3 Results & Discussion

The present study aims to decipher which force field(s) best describe different helical conformations of polypeptides. Depending upon structural properties computed from QM optimized structures, and that computed from classical MD force fields conclusions are drawn regarding performance of the different force fields. For this, 5-mer and 10-mer homopolymeric helical chains in R-H-L, R-H-D, L-H-L, and L-H-D conformations from ALA, LEU, GLU and LYS are studied in water and vacuum separately. All the chains were first QM optimized in water and vacuum separately, followed by classical MD simulations. The QM optimized structures are used as reference to compute the structural properties from classical MD simulations. Depending upon these structural properties, conclusions are drawn regarding performance of the different force fields.

4.3.1 Relative Energies of the Helical Conformers from QM Optimization

Energies of QM optimized structures of the R-H-L, R-H-D, L-H-L and L-H-D helical conformers of 5-mer and 10-mer polypeptide chain from ALA, LEU, LYS and GLU are compared. The objective is to understand the distinctions in conformational energies of the different helical conformations for the helices constructed from a particular amino acid. For example, differences in conformational energies of R-H-L, R-H-D, L-H-L and L-H-D helices from 5-mer poly-ALA chain. Absolute energy from quantum chemical calculations of each conformer was calculated as the sum of total electronic energy, zero-point energy and basis

set superposition error correction values. Then the relative energies of the conformers were computed with respect to R-H-L, and are reported in Table 4.3. Relative energies are reported separately for 5-mer and 10-mer polypeptide chains from ALA, LEU, LYS and GLU. Positive values indicate less favorable conformation compared to R-H-L, and negative value the reverse. Furthermore, to better understand the distinctions in energies of the different conformations, the relative energies are plotted in Figure 4.1 across different systems. The relative energies of the different conformations from 5- and 10-mer of ALA, LEU, LYS and GLU are plotted along left to right in Figure 4.1.

Table 4.3: Relative energies (kJ/mol) of different conformers of 5-mer and 10-mer homo-polypeptides of ALA, LEU, LYS, GLU with respect to their R-H-L conformer of all the systems at the M06/6-311G (d, p) level.

Systems	Medium	R-H-D	L-H-L	L-H-D
ALA-5-mer	In Vacuum	16.348	16.348	0.000
	In Water	25.018	25.018	0.001
ALA-10-mer	In Vacuum	51.753	52.684	-3.330
	In Water	57.932	53.461	-0.212
LEU-5-mer	In Vacuum	58.741	54.298	38.089
	In Water	60.149	54.542	32.666
LEU-10-mer	In Vacuum	170.760	116.700	69.568
	In Water	132.744	118.042	64.152
LYS-5-mer	In Vacuum	29.159	30.535	-1.313
	In Water	33.223	32.914	-0.279
LYS-10-mer	In Vacuum	66.511	90.255	-1.704
	In Water	88.644	95.435	1.489
GLU-5-mer	In Vacuum	36.338	37.157	7.516
	In Water	18.484	20.723	10.581
GLU-10-mer	In Vacuum	68.425	99.318	22.282
	In Water	110.108	68.819	46.651

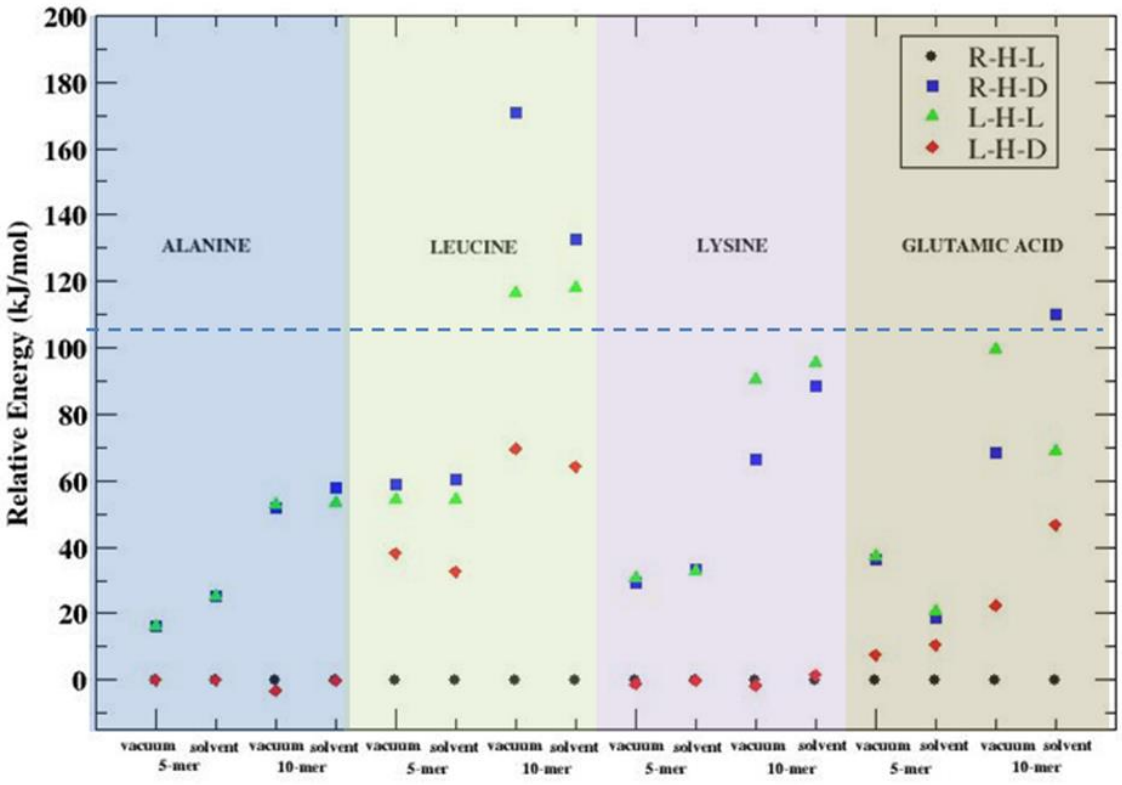


Figure 4.1: Relative energies (kJ/mol) with reference to the R-H-L of 5-mer and 10-mer of ALA, LEU, LYS and GLU of all the conformers in vacuum and in water at the M06/6-311G (d, p) level.

From Figure 4.1, it can be observed that R-H-D and L-H-L are less favorable conformations with respect to R-H-L across the different helices constructed from ALA, LEU, LYS and GLU. This observation remains consistent for both 5- and 10-mer chains across helices constructed from different amino acids. L-H-D conformations have energies similar to R-H-L conformation for most of the systems. Exceptions appear for 5- and 10-mer of LEU (in both vacuum and water), and for 10-mer of GLU (in both vacuum and water). In these cases, energies of L-H-D conformation are higher compared to R-H-L (Figure 4.1). But, for 5- and 10-mer of LEU, and 10-mer of GLU energies of R-H-D and L-H-L conformation are even higher. As a result, R-H-L and L-H-D conformations can be inferred to be more favorable conformation compared to R-H-D and L-H-L conformations (Figure 4.1). The QM optimized structures of all the conformers of 10-mer polypeptides in water are depicted in Figure 4.2.

The conformational preference of all the helical chains for R-H-L and L-H-D can be attributed to the steric repulsion between main chain C=O moiety and Cα-C1 bond.[3] In R-H-

D and L-H-L conformers, main chain C=O group and Cα-C1 bond remain in syn orientation (same side) to each other. On the other hand, in R-H-L and L-H-D conformers, main chain NH moiety remains in syn orientation relative to Cα-C bond. As a result, in R-H-D and L-H-L conformers, repulsion between main chain C=O group and Cα-C1 bond contribute to the higher energy of these helical conformations. Orientations mentioned above for all the four α-helical polypeptide conformers of ALA are depicted in Figure 4.3. The steric hindrance between side chain atoms with main chain –C=O moiety is more compared to that with main chain –NH moiety. Such interactions are responsible for the differences in energy for the different conformations. Jiang et al[21] have also obtained similar trend in relative energies of α_R and α_L conformation of amino acids belonging to hydrophobic and polar classes.

Further, considering relatively higher energy conformers, i.e. R-H-D and L-H-L of different α-helical polypeptides, some interesting trends in energetics are observed. These observations provide useful insight on the correlation of the energies of α-helical polypeptides with their nature (hydrophobicity, hydrophilicity) and lengths of amino acid side chains. For α-helical polypeptides of same chain lengths (10-mer and 5-mer), relative energies of R-H-D and L-H-L with respect to R-H-L conformer is much higher for LEU helix than ALA helix (Table 4.3, and Figure 4.1). This is because of higher repulsion between bulky LEU side chain with main chain, and higher side chain–side chain repulsion among bulkier LEU residues. Next, although side chain length of LEU and LYS are comparable, the relative energies of R-H-D and L-H-L conformers of LYS are lower than that of LEU. The relative energies of R-H-D and L-H-D conformers of LYS-10-mer are 88.64 kJ/mol and 95.4 kJ/mol in water which are very less than that values found for R-H-D (132.74 kJ/mol) and L-H-L conformers (118 kJ/mol) of LEU-10-mer (Table 4.3). This trend is observed for 10-mer in vacuum, and 5-mer in vacuum and in water (Table 4.3, and Figure 4.1). This can be attributed to favorable interaction among the hydrophilic side chains in LYS. For GLU, energies of R-H-D and L-H-D conformers are close to each other for 5-mer chains in both vacuum and water, but their energies are different for 10-mer chains. It can further be observed that although for almost all of the cases, R-H-L conformers have lowest energies irrespective of their chain lengths. Some distinctions in this trend are observed for L-H-D from 10-mer of ALA, 5-mer from LYS in vacuum and in water, and 10-mer from LYS in vacuum have lower energy among all other conformers (Table 4.3). However, the difference in energy from R-H-L is small.

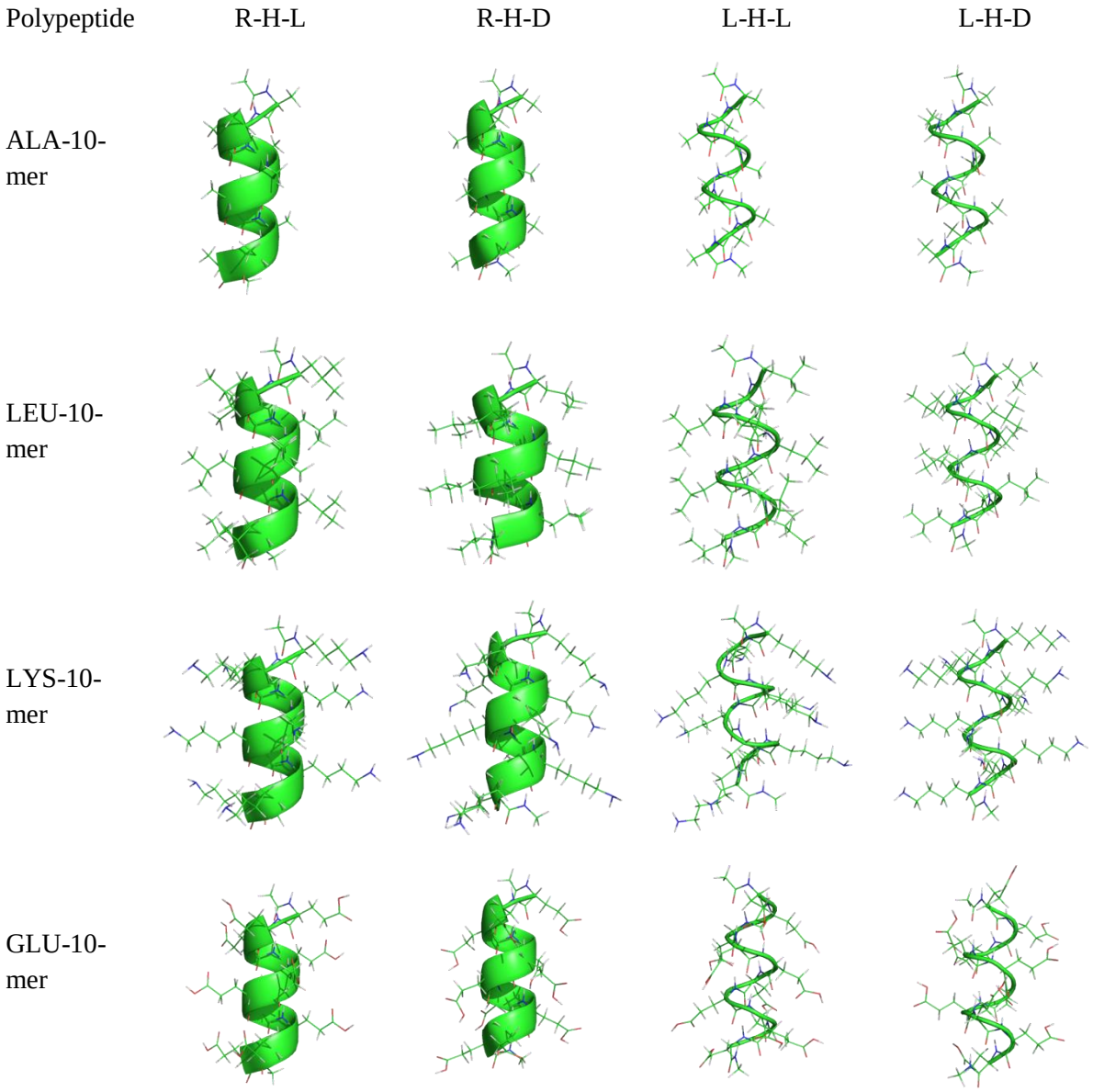


Figure 4.2: QM optimized structures of α-helical polypeptides of 10-mer of ALA, LEU, LYS, GLU amino acids in water at the M06/6-311G (d, p) level.

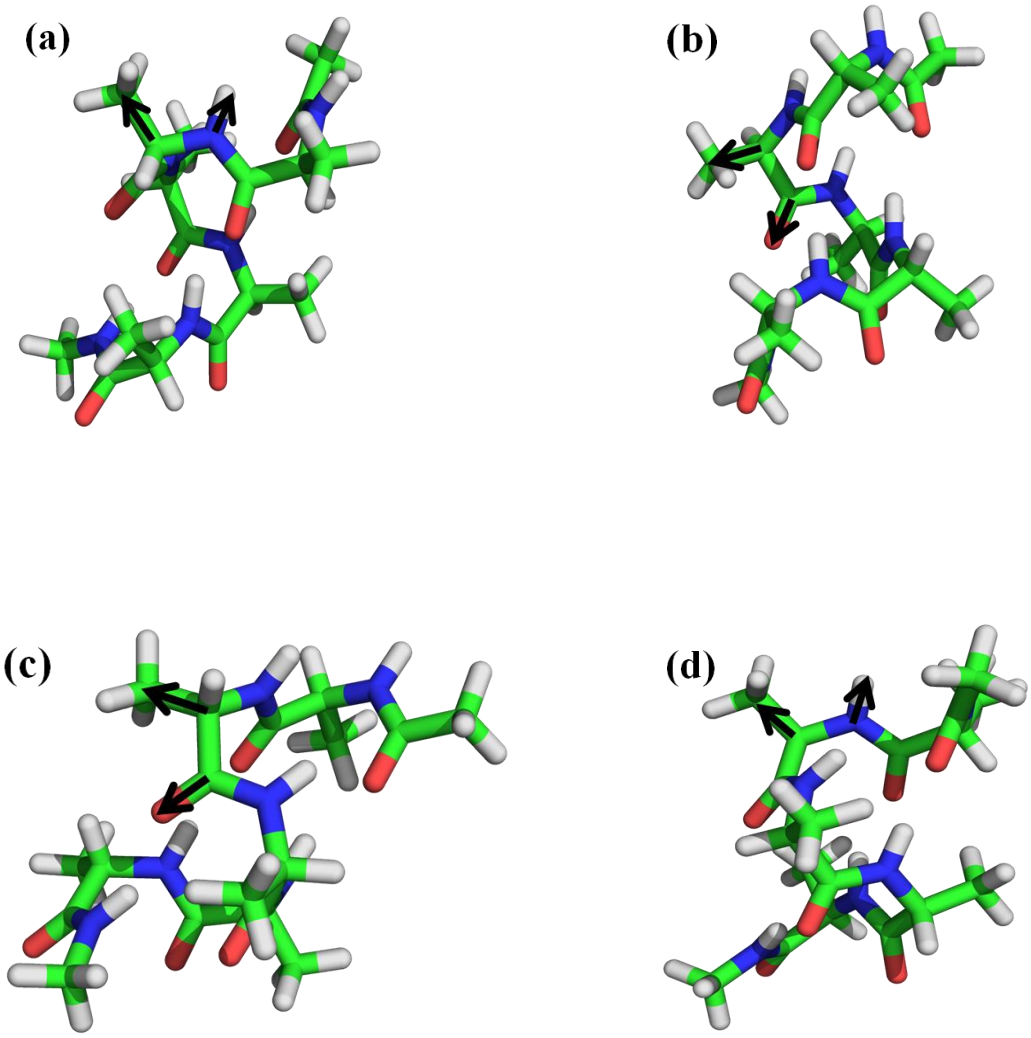


Figure 4.3: Orientation of side chain and main chain atoms of (a) R-H-L, (b) R-H-D, (c) L-H-L, (d) L-H-D conformers of ALA amino acid residue are shown by black arrows.

4.3.2 Structural Properties of Helices - Comparison Between QM Optimized and Classical MD Structures

Helical conformation of polypeptide chains are defined depending upon backbone Φ and Ψ dihedral angles of peptide chain backbone, and H-bonding between -C=O of the i^{th} and -NH group of the $i+4^{th}$ residue. Values of Φ and Ψ dihedral angles and their sign define α_R and α_L helices. For constructing starting structure for QM optimization Φ and Ψ angles for α_R and α_L helices were considered to be -64.0° and -43.0°, and +57.0°, +47.0° respectively. Further, in helical conformations all the –C=O of the i^{th} and –NH group of the $i+4^{th}$ residue remain H-

bonded to each other. In the present study, all the helical chains are capped at the terminal (–$NHCH_3$ and –$COCH_3$ groups at C- and N- terminal, respectively). These terminal capped groups also participate in *i-i+4* H-bonding.

Another important criterion to define helicity is the helical order parameter (defined and explained later in the text). Helical order parameter provides direct measurement of helical configuration and does not depend upon any predefined criterion for helicity.

The primary objective of this study is to realize that among the commonly used force fields, which one can better represent all the helical conformations of homo-polymeric peptides. To achieve that, it is important to investigate structural properties of 5- and 10-mer of R-H-L, R-H-D, L-H-L, and L-H-D helices obtained after performing MD simulation using different classical force fields. The above mentioned structural properties is used to define the helicity of peptide chains are computed, and QM optimized structure of the different helical conformations are used as benchmark for comparison. That is, the structural properties are first calculated from the QM optimized structures, and then compared separately with that obtained from classical MD simulation using each of the force fields. Depending upon comparison of these structural properties with QM optimized structures, performance of the different force fields to represent the different helical conformations are analyzed.

4.3.2.a Comparison depending upon Φ and Ψ dihedral angles and i-i+4 H-Bonding

First, Φ and Ψ angles and number of H-bonds between -C=O of the i^{th} and -NH group of the $i+4^{th}$ residue in all the α-helical peptide chains were calculated from QM optimized structures in vacuum and in water. Then, average of eight Φ and Ψ angles for 10-mer and average of three Φ and Ψ angles for 5-mer were considered for comparing the angles obtained from classical MD simulations using different force fields. For this averaging, Φ and Ψ angles at the terminals were discarded. This is because deviation of the Φ and Ψ angles from the values considered for construction of the helical chains for QM optimization were relatively higher for C- and N-terminals than the middle of the polypeptide chain. Further, there is possibility of formation of eight *i-i+4* H-bonds for 10-mer and three for 5-mer helical chains considering the capping residues. But as the distance and angle criterion for the formation of *i-i+4* H-bond by capped residues deviates most from the other residues, the terminal capped residues were not considered. Thus, six intra-helical H-bonds were considered for 10-mer and one for 5-mer. The averages of distances and angles obtained from 10-mer were used as cut-off for the calculation of number of intra-helical H-bonds in the structures, which are obtained at

different time frames from MD simulation trajectories using the above mentioned force-fields. All the average Φ and Ψ angles, and average H-bond distances and angles for 10-mer α-helical polypeptides, calculated from QM optimized structures are summarized in Table 4.4.

Table 4.4: Φ, Ψ angle values, distance cut-off and angle cut-off values for H-bond calculated from the optimised structures of 10-mer of α-helical peptides in vacuum and in water.

Systems	Conformers	Medium	Φ	Ψ	H-bond distance (nm)	H-bond angle (deg.)
ALA-10-mer	R-H-L	In Vacuum	-62	-39	<0.24	>134
		In Water	-61	-42	<0.19	>150
	R-H-D	In Vacuum	-67	-40	<0.22	>131
		In Water	-53	-47	<0.198	>157
	L-H-L	In Vacuum	+57	+46	<0.21	>165
		In Water	+53	+47	<0.198	>157
	L-H-D	In Vacuum	+62	+39	<0.31	>109
		In Water	+62	+42	<0.198	>157
LEU-10-mer	R-H-L	In Vacuum	-60	-41	<0.22	>144
		In Water	-60	-43	<0.2	>158
	R-H-D	In Vacuum	-53	-46	<0.21	>156
		In Water	-61	-53	<0.22	>121
	L-H-L	In Vacuum	+52	+47	<0.21	>153
		In Water	+57	+49	<0.194	>150
	L-H-D	In Vacuum	+58	+46	<0.20	>155
		In Water	+59	+46	<0.192	>161
	R-H-L	In Vacuum	-59	-43	<0.21	>152

LYS-10-mer		In Water	-60	-44	<0.197	>159
	R-H-D	In Vacuum	-51	-47	<0.232	>148
		In Water	-52	-48	<0.2	>157
	L-H-L	In Vacuum	+51	+47	<0.22	>156
		In Water	+52	+47	<0.2	>153
	L-H-D	In Vacuum	+59	+43	<0.215	>147
		In Water	+60	+44	<0.2	>158
GLU-10-mer	R-H-L	In Vacuum	-62	-39	<0.21	>147
		In Water	-60	-43	<0.197	>160
	R-H-D	In Vacuum	-55	-30	<0.38	>115
		In Water	-53	-48	<0.195	>160
	L-H-L	In Vacuum	+55	+46	<0.21	>159
		In Water	+53	+47	<0.2	>154
	L-H-D	In Vacuum	+62	+39	<0.22	>138
		In Water	+60	+43	<0.198	>156

Further, the 1 ns production run trajectories obtained from classical MD simulations were analyzed for calculating Φ and Ψ angles and number of H-bonds. Φ and Ψ angles were first computed in each frame of last 1 ns of production run. The range of φ, ψ angles (+/-10) for MD simulation analysis was obtained from literature. Barlow et al [55] have shown in Thermolysin protein where all the four types of α-helices are present consisting of ALA, LEU, LYS, GLU, their φ, ψ vary upto ±10° from their mean values. If the Φ and Ψ angles fell within ±10° of the average Φ and Ψ angle calculated from QM optimized structure, the corresponding residue was considered to be forming helix. This count of helix forming residues was then divided by eight (for 10-mer and three for 5-mer), which is the number of residues considered for averaging Φ and Ψ angle from QM optimized structure, thus obtaining the fraction of residues in helical conformation. This fraction was averaged over whole 1 ns to obtain average fraction of residues in helical conformation. Similarly, average fraction of *i-i+4* H-bonds was computed. For this, first the number of *i-i+4* H-bonds was

calculated in each frame of last 1 ns trajectory. Then this count was divided by six (for 10 mer and one for 5-mer), obtaining fraction of number of *i-i+4* H-bonds. This fraction was averaged over whole 1 ns to obtain average fraction *i-i+4* H-bonds. Finally, the average of averages of fraction of residues in helical conformation (obtained from Φ and Ψ angles) and fraction *i-i+4* H-bonds was calculated. This number is plotted in Figure 4.4a and 4.4b for 10-mer chains in water, and in vacuum respectively. The number is computed and plotted in Figure 4a and b for all the helical conformation and the different force fields considered in the present study. This way, equal contribution from Φ and Ψ angles, and *i-i+4* H-bond was considered for determining the helicity of different polypeptide chains. Same number obtained from QM optimized structure is also depicted in Figure 4.4a and 4.4b, in water and in vacuum respectively. Closer the number obtained from classical MD calculations to that from DFT, better is the helicity.

For 10-mer chains (Figure 4.4a and 4.4b) in water and in vacuum, the number depicted is closest to the number obtained from QM optimized structure for CHARMM-27/CMAP force field. Only exception appears for R-H-D conformer from ALA in vacuum which is best explained with OPLS/AA. We also observed that AMBER-03 force-field gives average values more closer to QM optimized structure than obtained from OPLS/AA force-field in water. Similar trend is obtained in water too, with the exception of R-H-D conformer in which OPLS/AA gives better result than AMBER-03. Therefore, it can inferred that CHARMM-27/CMAP shows the best trend with QM data among the three force fields, and AMBER-03 is better than OPLS/AA to describe all the conformers from all α-helical polypeptides considered in this study.

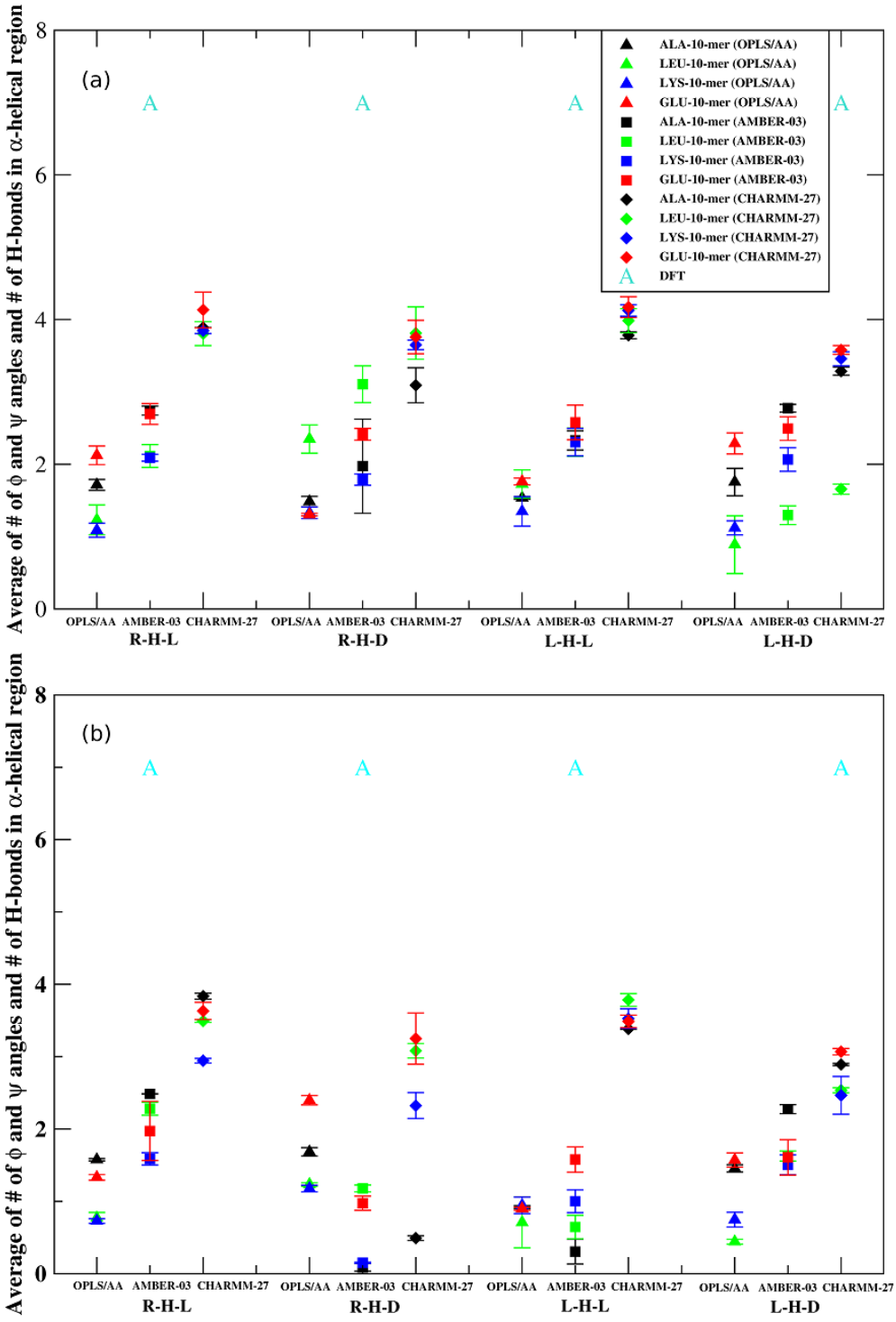


Figure 4.4 (a) Average of number (#) of Φ, Ψ angles and H-bond with equal contribution of 10-mer of ALA, LEU, LYS and GLU in OPLS/AA, AMBER-03 and CHARMM-27/CMAP force-fields in TIP3P water, **(b)** Average of number (#) of Φ, Ψ angles and H-bond with equal contribution of 10-mer of ALA, LEU, LYS and GLU in OPLS/AA, AMBER-03 and CHARMM-27/CMAP force-fields in vacuum.

Further, the conditions for Φ and Ψ angles and number of *i-i+4* H-bonds were segregated, and performances of the different force fields to represent these criteria were checked. Φ and Ψ angles were calculated from 1 ns of classical MD trajectory and converted into histogram. Distributions, normalized over number of frames in 1 ns, are depicted in Figure 4.5-4.8 in water. It can be observed clearly that CHARMM-27/CMAP force filed gives narrower distribution of Φ and Ψ dihedral angles relative to the other two force fields. It implies that CHAMM27/CMAP force field keeps the polypeptide chains in α-helical conformation more compared to the other two force fields both in vacuum and in water. Few helical conformers are observed to scan regions other than α-helix in water (Figure 4.5-4.8). The R-H-D conformers of all the polypeptides in water sample regions other than α-helix using CHAMM27/CMAP force field. Similar behavior is observed for L-H-L conformers of 10-mer polypeptides from ALA and LYS in water using AMBER-03 force field (Figure 4.5-4.8). OPLS/AA force field samples R-H-L conformers from LYS and GLU in other regions also along with the α-helical region in vacuum. In summary, Φ and Ψ angles are the best represented in the helical region by CHAMM27/CMAP force field for the helices considered in the present study.

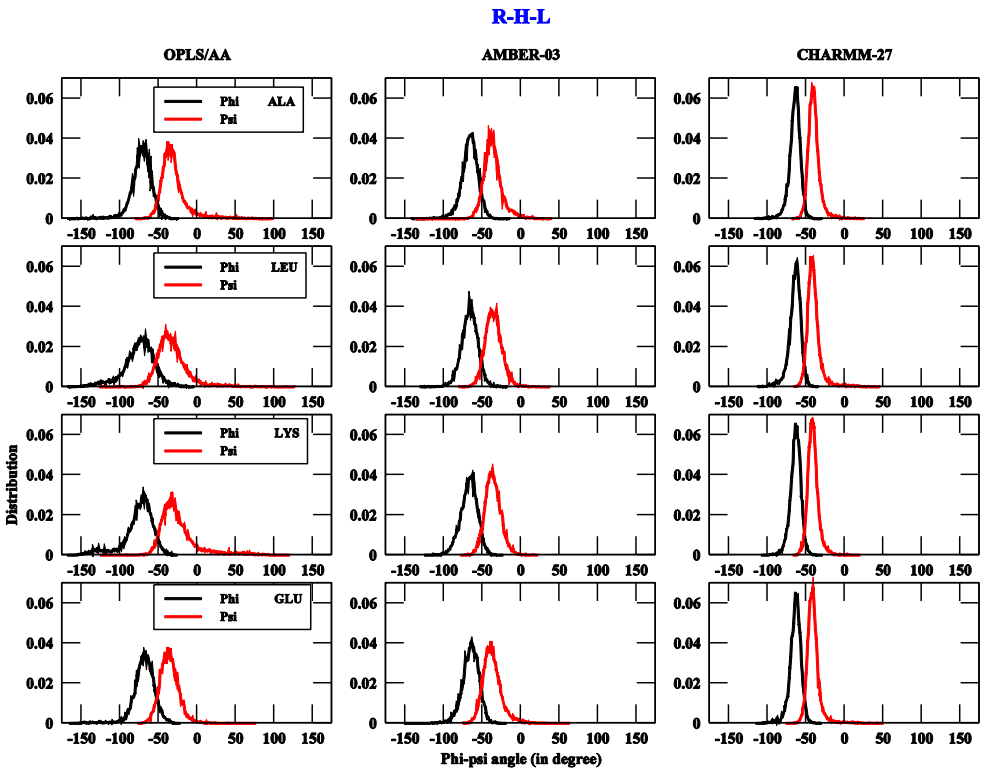


Figure 4.5: Distribution of Φ, Ψ angles of 10-mer R-H-L of ALA, LEU, LYS and GLU in OPLS/AA, AMBER-03 and CHARMM-27/CMAP force-fields in TIP3P water.

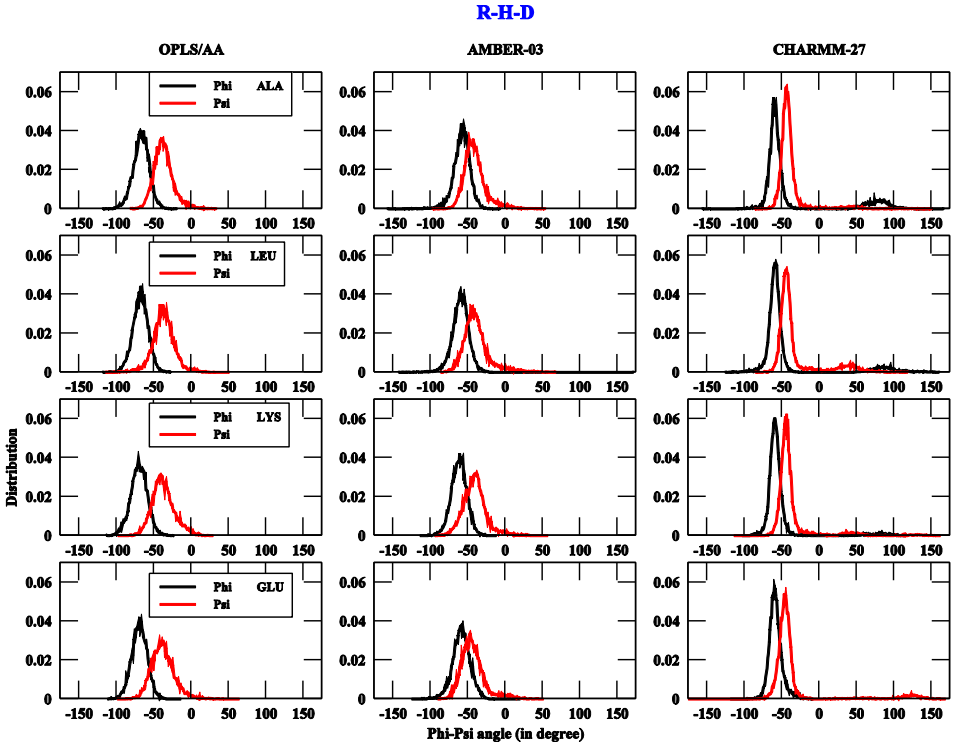


Figure 4.6: Distribution of Φ, Ψ angles of 10-mer R-H-D of ALA, LEU, LYS and GLU in OPLS/AA, AMBER-03 and CHARMM-27/CMAP force-fields in TIP3P water.

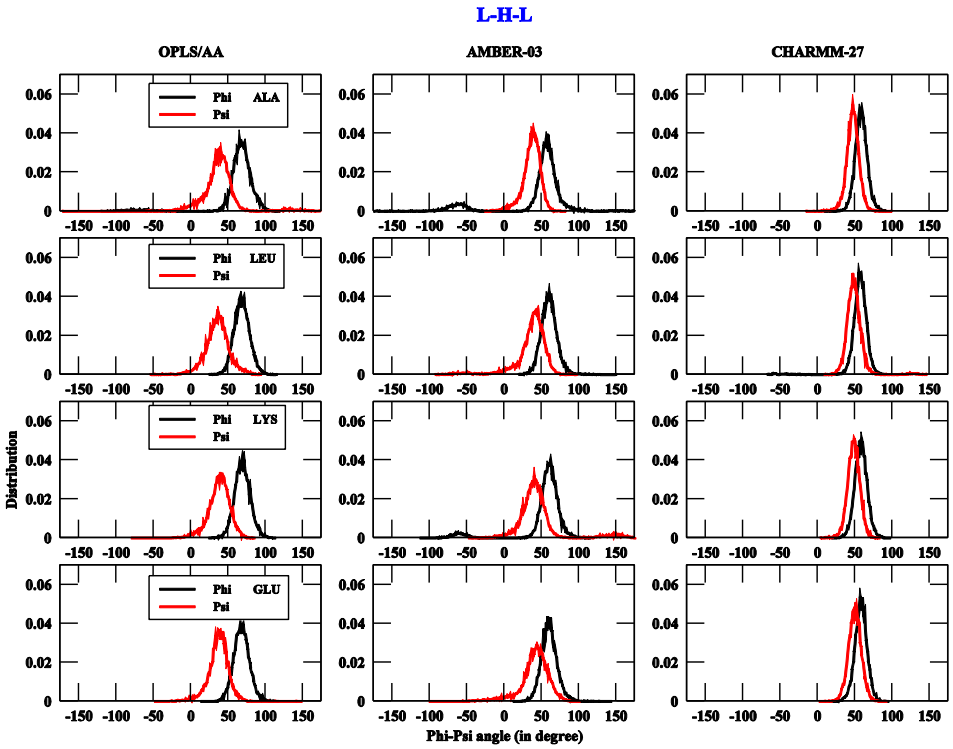


Figure 4.7: Distribution of Φ, Ψ angles of 10-mer L-H-L of ALA, LEU, LYS and GLU in OPLS/AA, AMBER-03 and CHARMM-27/CMAP force-fields in TIP3P water.

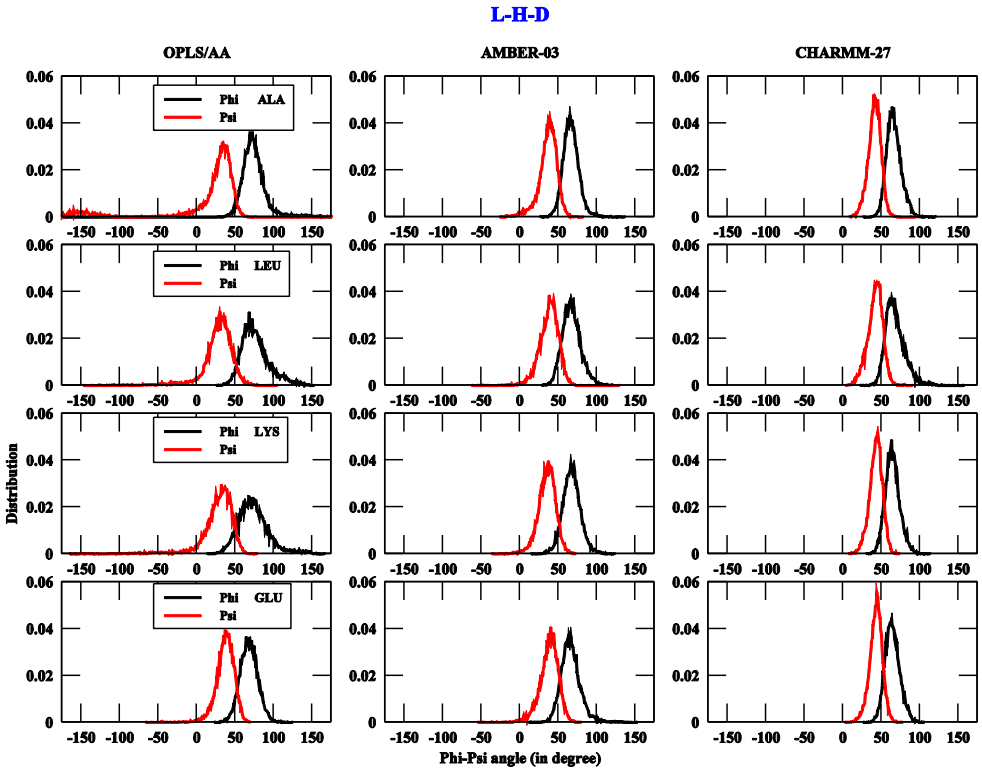


Figure 4.8: Distribution of Φ, Ψ angles of 10-mer L-H-D of ALA, LEU, LYS and GLU in OPLS/AA, AMBER-03 and CHARMM-27/CMAP force-fields in TIP3P water.

To analyze the performance of the different force fields to represent *i-i+4* H-bonding, H-bonding distance (between Hydrogen and acceptor), H-bonding angle (H-bond donor-Hydrogen-acceptor) were calculated for the 1 ns production run. H-bonding angle is plotted as a function of H-bonding distance for all the numbers obtained from 1 ns trajectory. These H-bonding angle-distance population plots for 10-mer polypeptide chains are depicted in ESI Figures 4.9-4.12 in water. Here also, CHARMM- 27/CMAP force field gives narrower H-bond distance-angle population compared to other force fields for all the helical conformers except R-H-D in vacuum and in water. It indicates that the helical structures fluctuate less using CHARMM-27/CMAP force field.

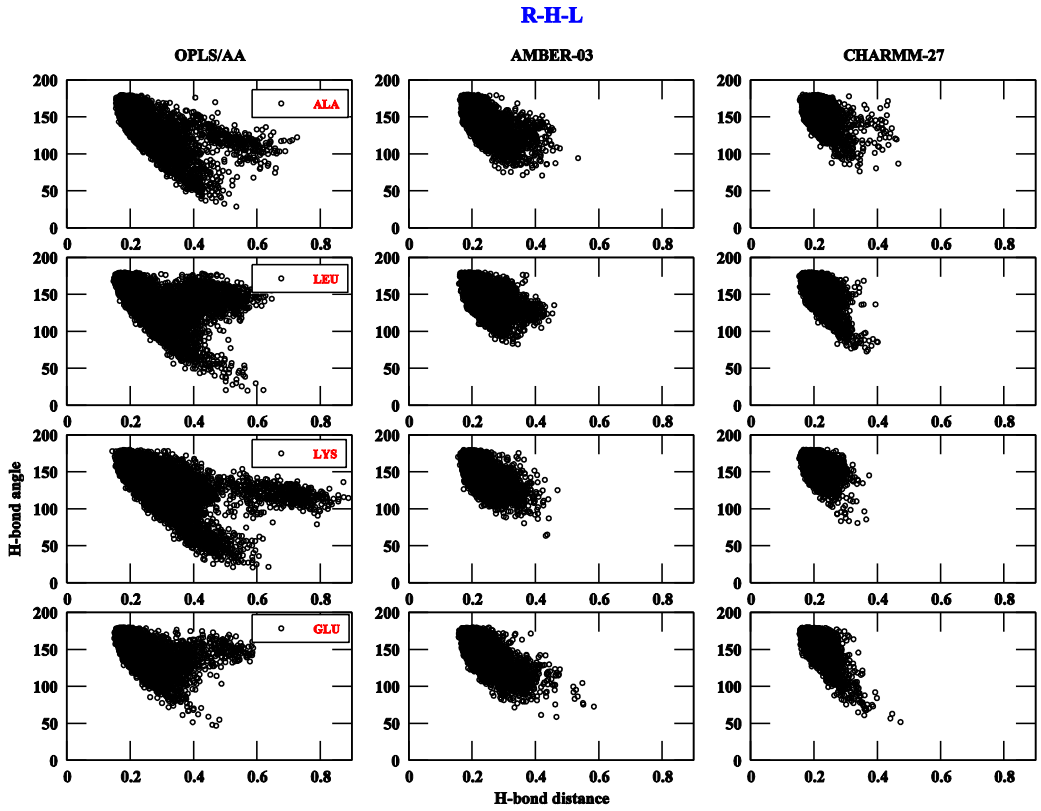


Figure 4.9: H-bond distance vs. H-bond angle of 10-mer R-H-L of ALA, LEU, LYS and GLU in OPLS/AA, AMBER-03 and CHARMM-27/CMAP force-fields in TIP3P water.

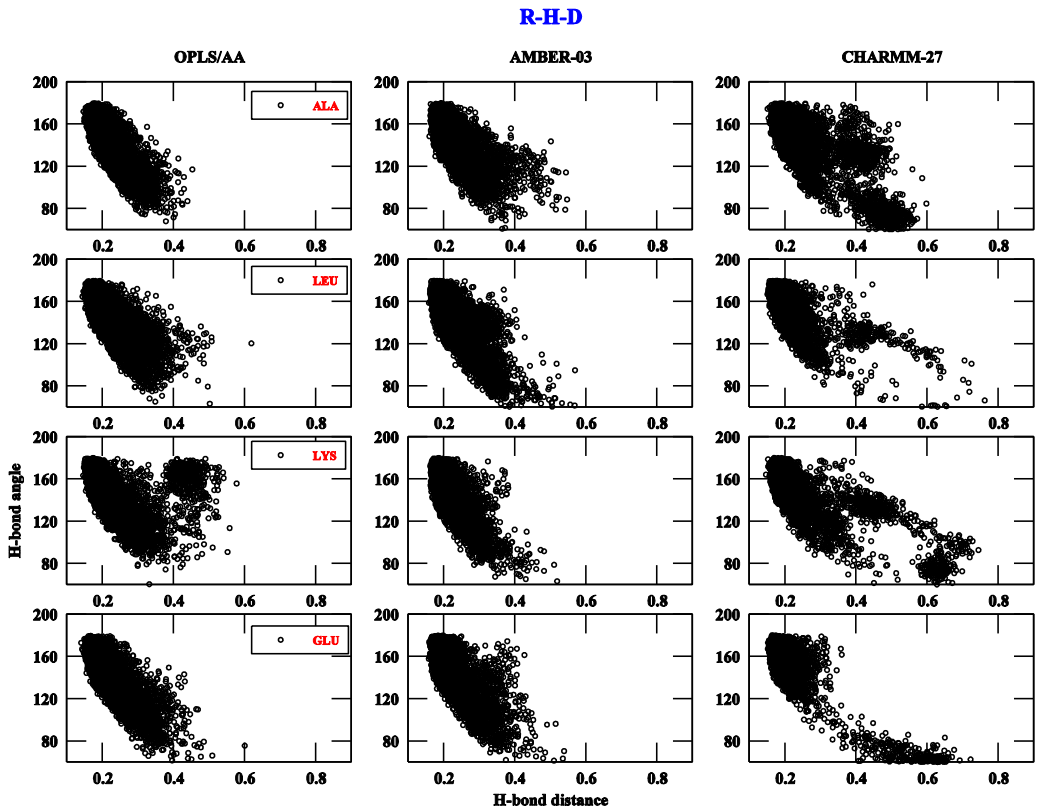


Figure 4.10: H-bond distance vs. H-bond angle of 10-mer R-H-D of ALA, LEU, LYS and GLU in OPLS/AA, AMBER-03 and CHARMM-27/CMAP force-fields in TIP3P water.

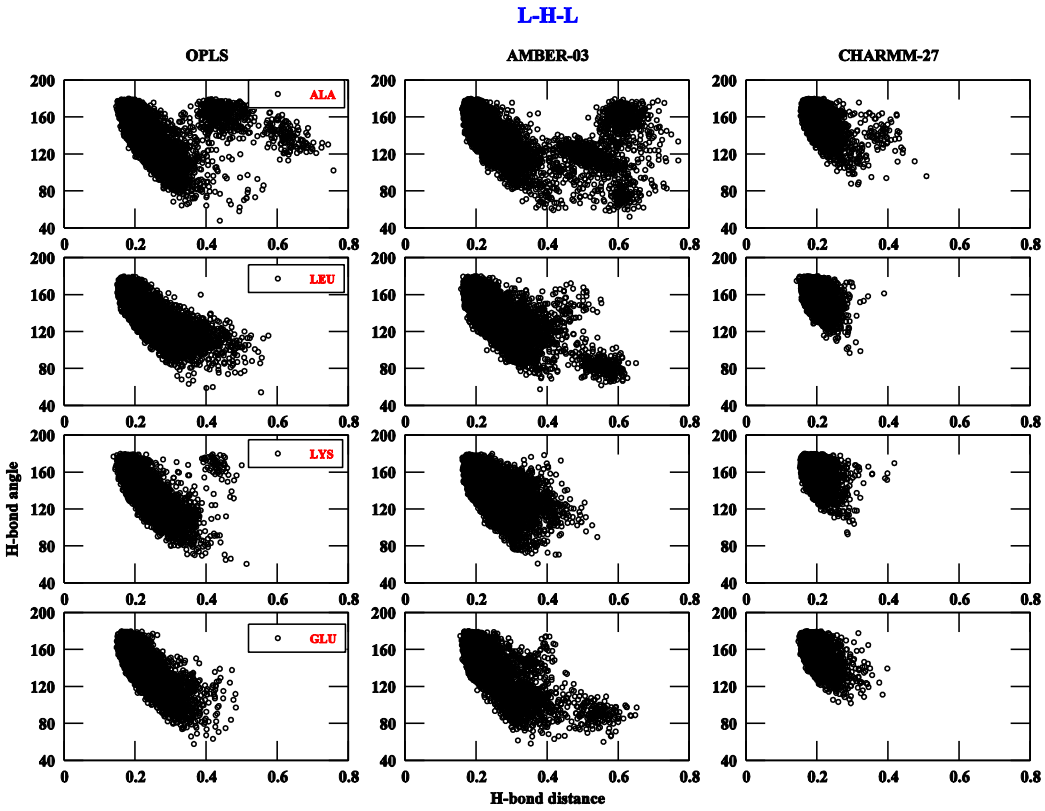


Figure 4.11: H-bond distance vs. H-bond angle of 10-mer L-H-L of ALA, LEU, LYS and GLU in OPLS/AA, AMBER-03 and CHARMM-27/CMAP force-fields in TIP3P water.

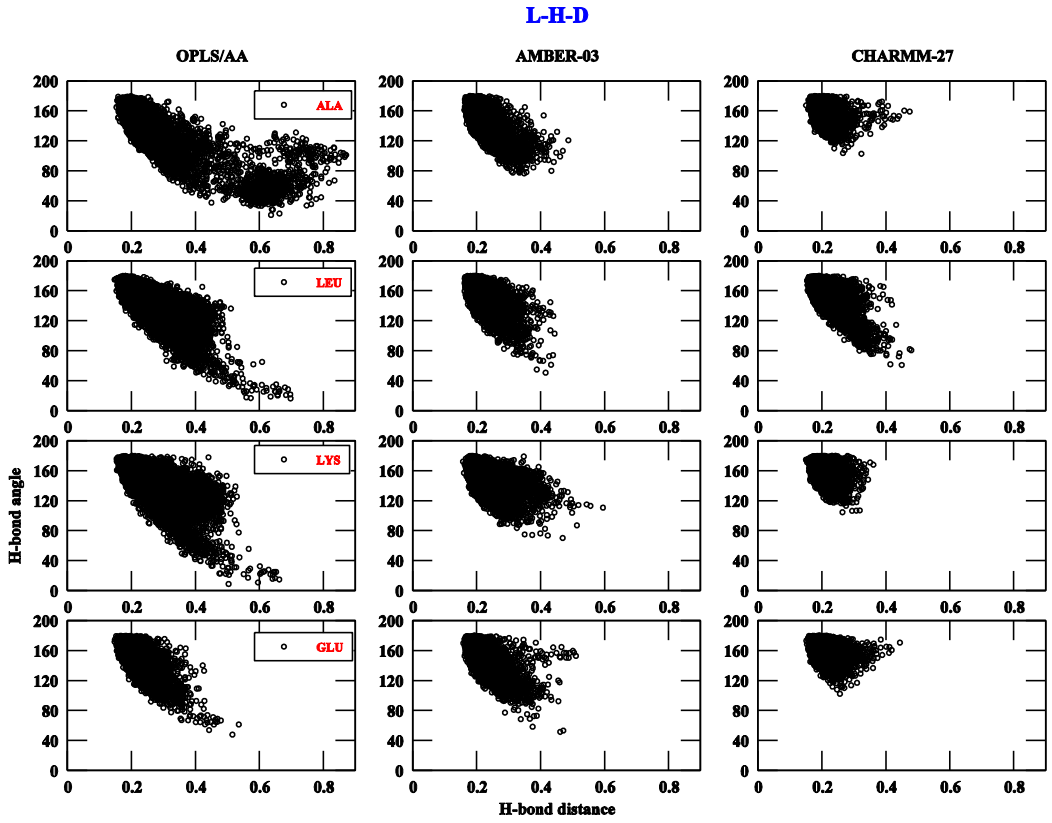


Figure 4.12: H-bond distance vs. H-bond angle of 10-mer L-H-D of ALA, LEU, LYS and GLU in OPLS/AA, AMBER-03 and CHARMM-27/CMAP force-fields in TIP3P water.

4.3.2.b Comparison depending upon helical order parameter

Three helical order parameters, namely H_2, H_3, and H_4 were calculated to analyze the distinctions in helicity of the different helical conformations of polypeptide chains. Helical order parameters are direct quantification of helicities. Helical order parameters were calculated for all the structures obtained from MD simulations using different force fields, and compared with that obtained from QM optimized structures. To calculate the helical parameters, centre of mass (COM) of backbone atoms for each of the residues in α-helical peptide chain was determined. At the COM of each residue a vector was defined ($\vec{u_i}$) by the cross product of two connected bond vectors (line connecting two consecutive amino acid residues). That is, $\vec{u_i} = (\vec{r_i} - \vec{r}_{i-1}) \times (\vec{r}_{i+1} - \vec{r_i}) / \sin\theta$, where $\vec{r_i}$ is the position vector of COM of i^{th} monomer, and θ is the angle between two connected bond vectors. The terminal capping residues were excluded from calculation of bond vectors. The local helicity, H_2 is finally calculated by taking dot product of u_i vectors of conjugated residues.

$$H_2 = \frac{1}{N-3}\sum_{i=2}^{N-2}\vec{u}_i \cdot \vec{u}_{i+1}$$

Similarly, H_3, and H_4 helical parameters are calculated as,

$$H_3 = \frac{1}{N-2}\sum_{i=2}^{N-1}\vec{u}_i \cdot \vec{u}_{mid} \ ,$$

u_{mid} is bond vector of middle monomer (N/2) of peptide chain.

$$H_4 = \left(\frac{1}{N-2}\sum_{i=2}^{N-1}\vec{u}_i\right)^2$$

Now, before calculating H_2, H_3, and H_4 helical order parameters for the different helical conformers considered in the present study, it is important to know the range of H_2, H_3, H_4 for α_R and α_L helices. For this, total 441 structures were generated for left-handed (Φ_L=57°±10°, Ψ_L=47°±10°) and right-handed α-helices (Φ_R=-64°±10°, Ψ_R=-43°±10°) by randomly chosen values of Φ and Ψ dihedral angles within 10° range from perfect helix value. From these randomly generated structures, distributions of H_2, H_3, H_4 helical order parameters are plotted in Figure 4.13a, and 4.13b. From the distributions it can be observed that H_2, H_3, H_4 for right-handed helices vary in the range of 0.35-0.65, 0.45-0.72 and 0.46-0.68 respectively, whereas for left-handed helices they vary in the range of 0.32-0.58, 0.42-0.7 and 0.42-0.65. The peak of the H_2, H_3 and H_4 distributions for the right handed α-helix appear at 0.5, 0.59 and 0.56, respectively, whereas for the left handed α-helix the peak values appear at 0.46, 0.56, and 0.53, respectively.

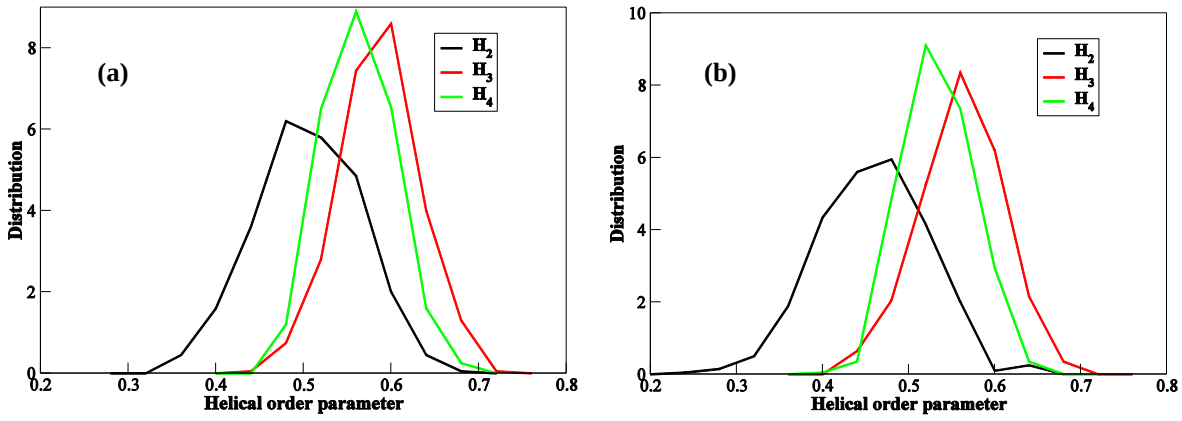


Figure 4.13: (a) Distribution of H_2, H_3 and H_4 helical order parameters of initially build

ALA-10 mer (a) α_{R-} helix helices (Φ_R=-64°±10°, Ψ_R=-43°±10°) by varying (Φ_R, Ψ_R) angles with an interval of 1°. (b) α_L helix helices (Φ_L=57°±10°, Ψ_L=47°±10°) by varying (Φ_L, Ψ_L) angles with an interval of 1°.

Next, H_2, H_3, H_4 helical order parameters were calculated from QM optimized structures, and from last 1ns of production run of classical MD simulation. Distribution of H_2 (Figure 4.14), H_3 (Figure 4.15), H_4 (Figure 4.16), and values obtained from QM optimized structure (dotted blue line in the plots) are plotted. The plots are depicted for 10-mer chains in water. Values obtained from QM optimized structures indicate that the different conformers of polypeptides mostly remain in the α-helical region (determined from distributions in Figure 4.13). But, the exception is observed for energetically higher R-H-D conformers of polypeptides constructed from LEU. The relative energy of the same conformer from LEU is also found to be higher in comparison to the other conformers (Figure 4.1). Further, the peaks of H_2 distribution for the different helical conformers fall within 0.50-0.55, except the scattered distributions for the R-H-D conformers from ALA and LEU when CHARMM-27/CMAP force field is used (Figure 4.14). Similar trend is also observed for H_3 and H_4 distribution plots using CHARMM-27/CMAP force field (Figure 4.15, 4.16). The peak of H_2 distribution comes within 0.43-0.49 for all the conformers using AMBER-03 force field (Figure 4.14). H_2 distribution is scattered for L-H-L conformer from ALA and GLU, and R-H-D conformers of LYS. The peaks of H_3 distribution appear near 0.59 for right handed helices using AMBER-03 force field, whereas for left-handed helices the peak comes near 0.6. The scattered H_3 distribution is observed for L-H-L conformer from ALA. The similar trend is observed for H_4 distribution plot using AMBER-03 force field. The OPLS/AA force field shows more scattered plots in comparison to other two force fields. Thus from the H_2, H_3, H_4 distribution from last 500 ps of MD simulation it is evident that CHARMM-27/CMAP is able to maintain most of the conformers of different polypeptides in α-helical region.

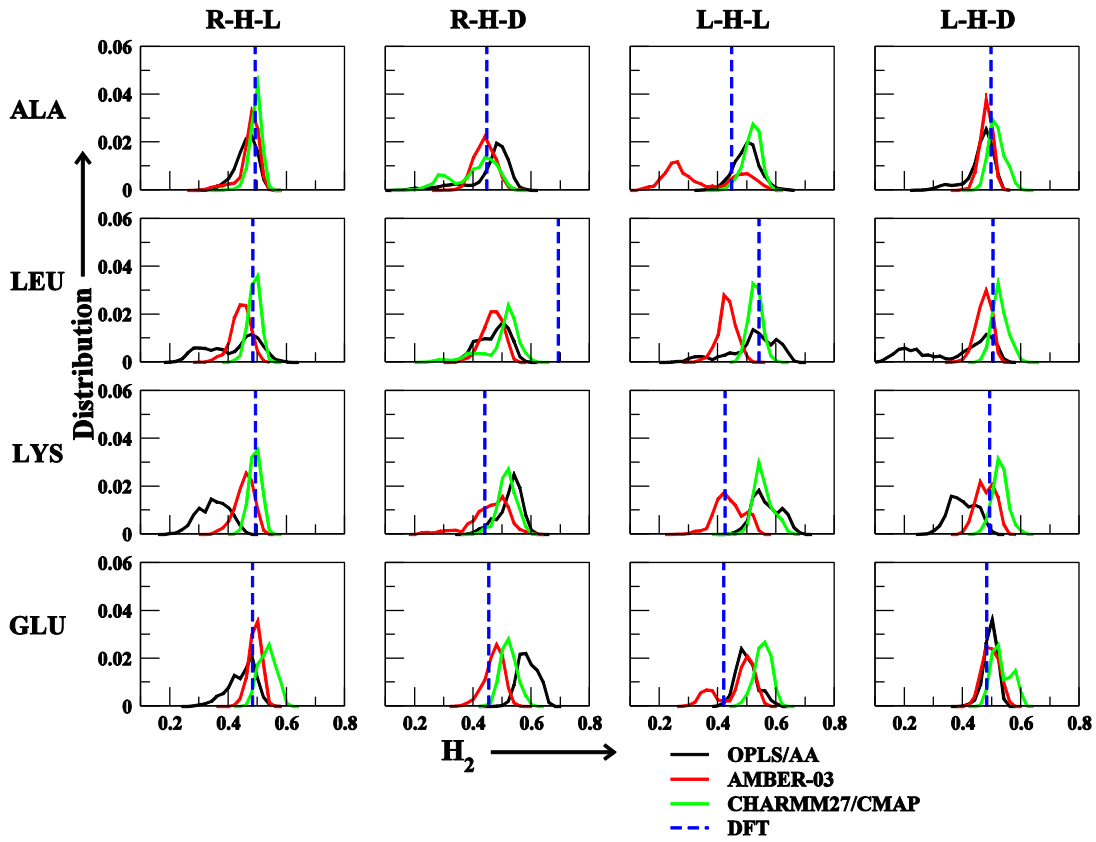


Figure 4.14: Distribution of H_2 helical order parameter of 10-mer of ALA, LEU, LYS and GLU in OPLS/AA, AMBER-03 and CHARMM-27/CMAP force-fields in TIP3P water. The vertical dotted blue line in each plot represents the value of H_2 helical order parameter of QM optimized structure of that conformer.

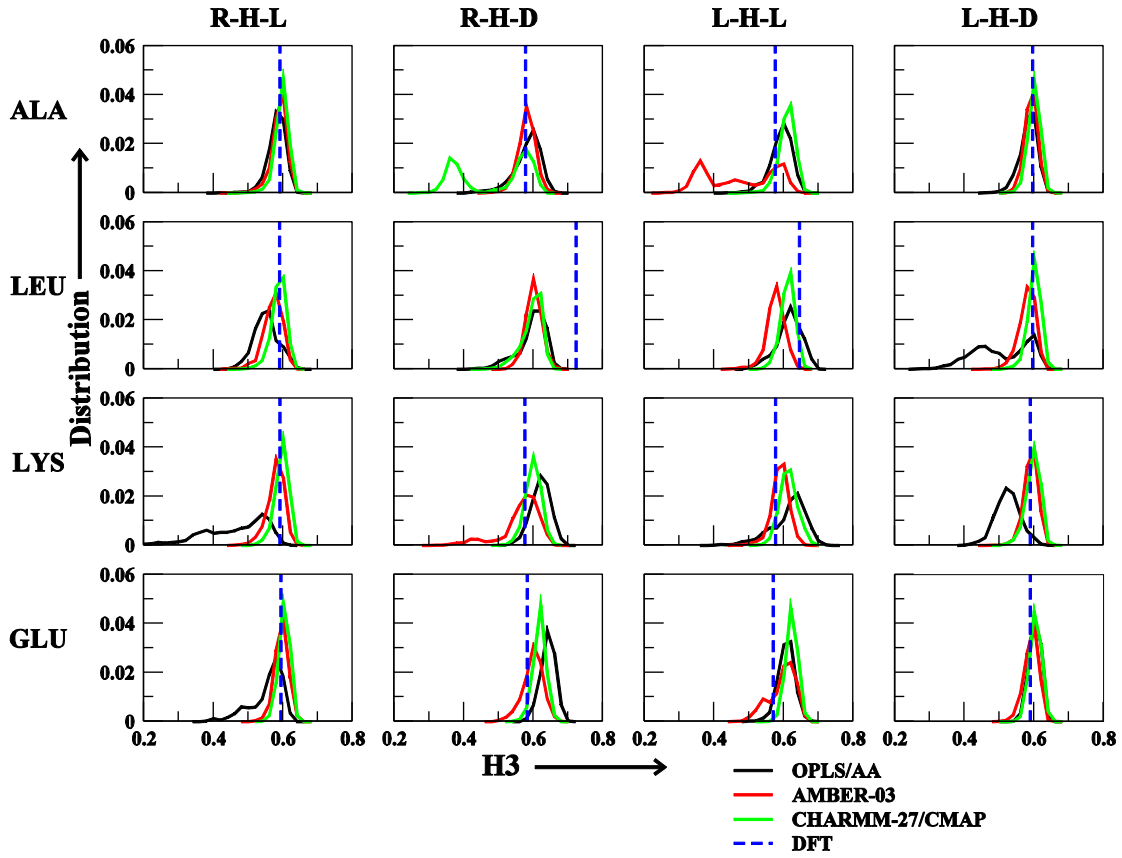


Figure 4.15: Distribution of H_3 helical order parameter of 10-mer of ALA, LEU, LYS and GLU in OPLS/AA, AMBER-03 and CHARMM-27/CMAP force-fields in TIP3P water. The vertical dotted blue line in each plot represents the value of H_3 helical order parameter of QM optimized structure of that conformer.

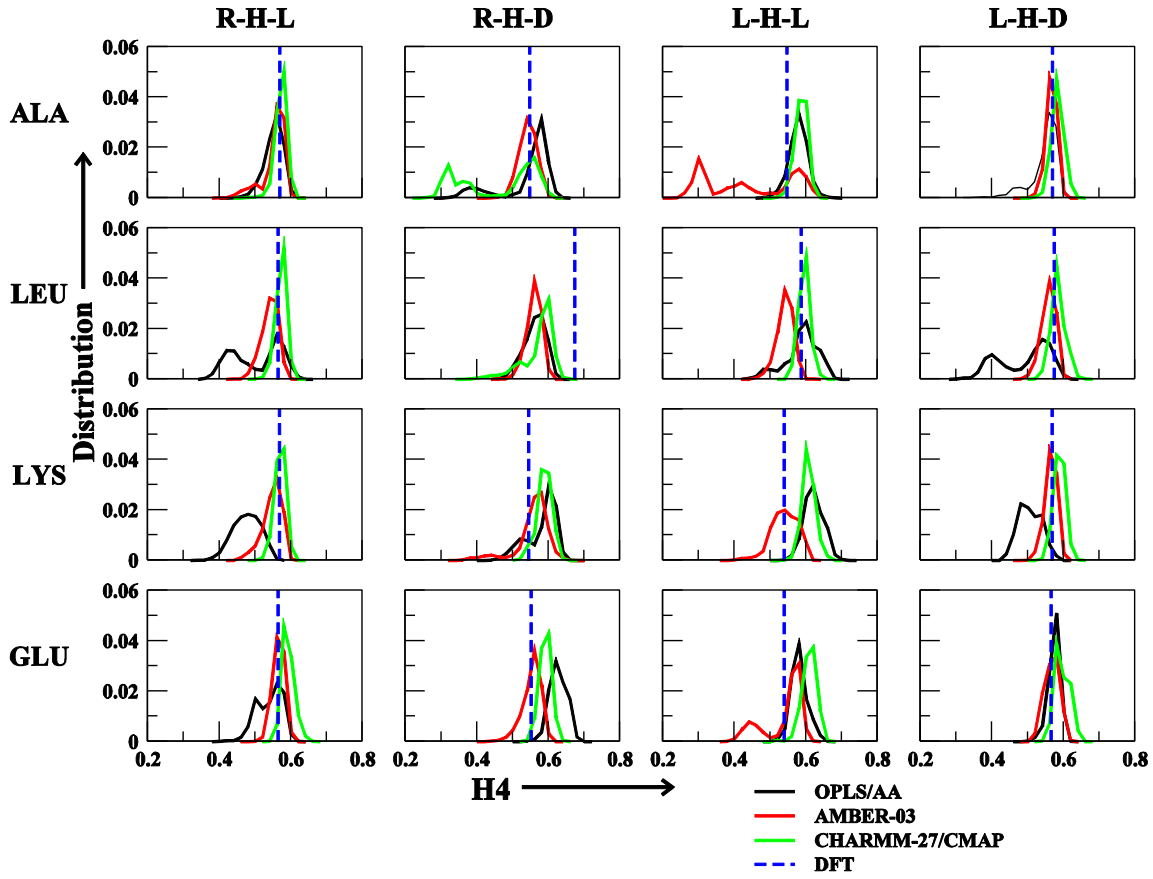


Figure 4.16: Distribution of H_4 helical order parameter of 10-mer of ALA, LEU, LYS and GLU in OPLS/AA, AMBER-03 and CHARMM-27/CMAP force-fields in TIP3P water. The vertical dotted blue line in each plot represents the value of H_4 helical order parameter of QM optimized structure of that conformer.

Broad H_2, H3, H4 distributions in Figure 4.14-4.16, is indicative of fluctuations in the helical structures during MD simulation. In the present study, we have considered the initial structure to be α-helical. The QM optimization and MD simulation of these helices gives structures which are preferably close to α-helices, but in some case they may deviate from α-helical structure depending upon the interactions acting within the polypeptide chain. Thus it is important to compare and contrast helical order parameters obtained from QM and MD simulations more explicitly. Hence percentage deviations of helical order parameters obtained from classical MD from that obtained from QM optimized structures were calculated. The percentage deviations were calculated as,

$$\text{Percentage deviation relative to QM} = 100*(H_{QM}-H_{MD})/H_{QM}$$

Higher deviation indicates that helices in MD are away from that from QM. Percentage deviations H_2, H_3, H_4 parameters are plotted for the different helical chain in Figure 4.17, 4.18, 4.19, respectively. From the percentage deviation plots it appears that the CHARMM-27/CMAP and AMBER-03 force field simulated resulted into better agreement of helical order parameters with QM. This holds good for energetically lower conformers of ALA, LEU, and LYS. But for the GLU it shows larger deviation with CHARMM-27/CMAP force field. AMBER-03 explains the energetically lower conformers of GLU better than CHARMM-27/CMAP. OPLS/AA shows large deviation with in helical parameters with respect to QM optimized structure for almost all of the polypeptides studied. The energetically higher R-H-D and L-H-L conformers of LYS and GLU are the best explained by AMBER-03, but it fails to show good agreement of helical order parameters for energetically higher conformers for the polypeptides consisting of LEU and L-H-L conformers of ALA. OPLS/AA and CHARMM-27/CMAP generally shows large deviation of helical order parameters for energetically higher conformers, but they explain L-H-L conformer of LEU, ALA well. R-H-D conformer of ALA is better explained by OPLS/AA. Hence calculation of percentage deviation in helicity provides a good structural quantification of MD structures compared with QM structures.

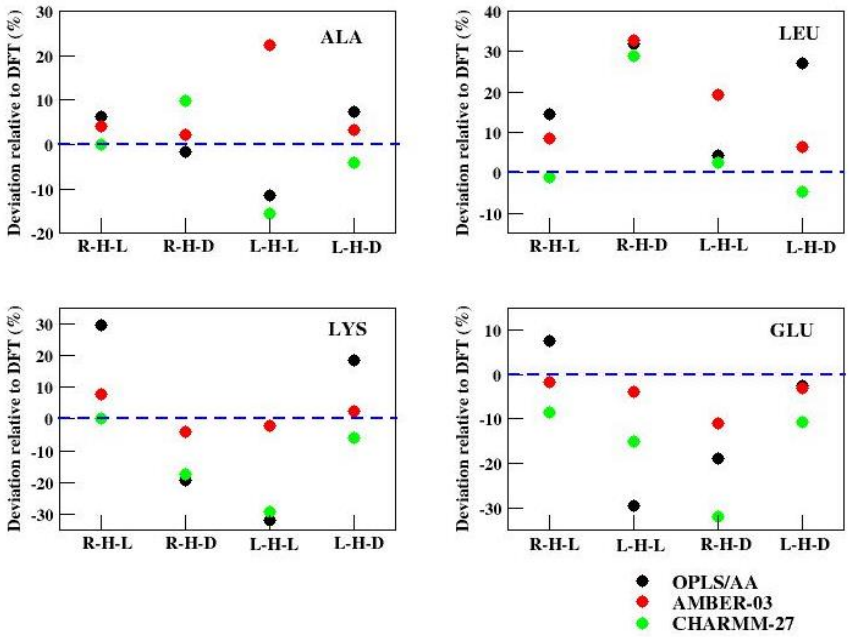


Figure 4.17: Percentage deviation of H_2 helical order parameter obtained from MD simulation relative to QM of 10-mer of ALA, LEU, LYS and GLU in OPLS/AA, AMBER-03 and CHARMM-27/CMAP force-fields in TIP3P water.

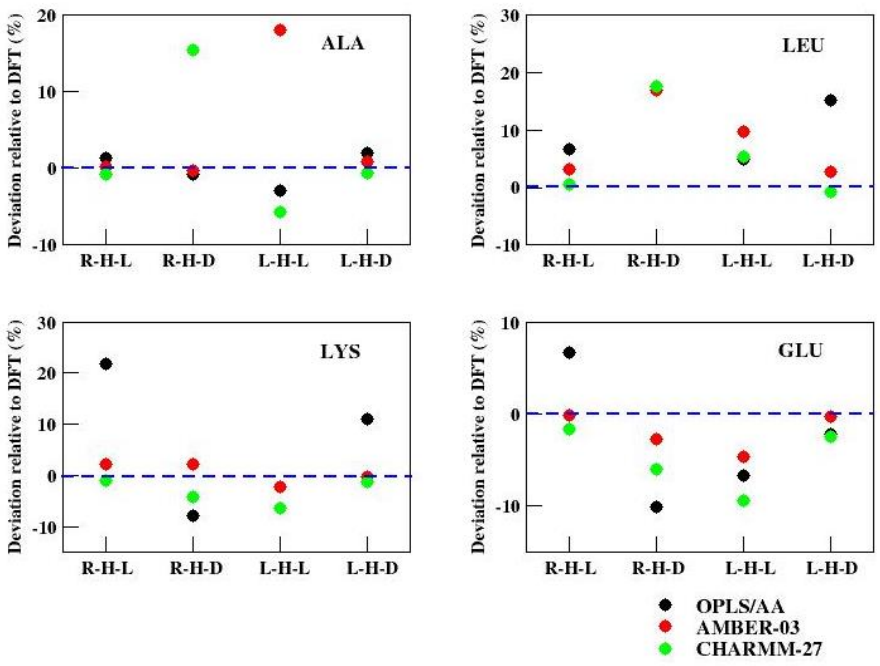


Figure 4.18: Percentage deviation of H_3 helical order parameter obtained from MD simulation relative to QM of 10-mer of ALA, LEU, LYS and GLU in OPLS/AA, AMBER-03 and CHARMM-27/CMAP force-fields in TIP3P water.

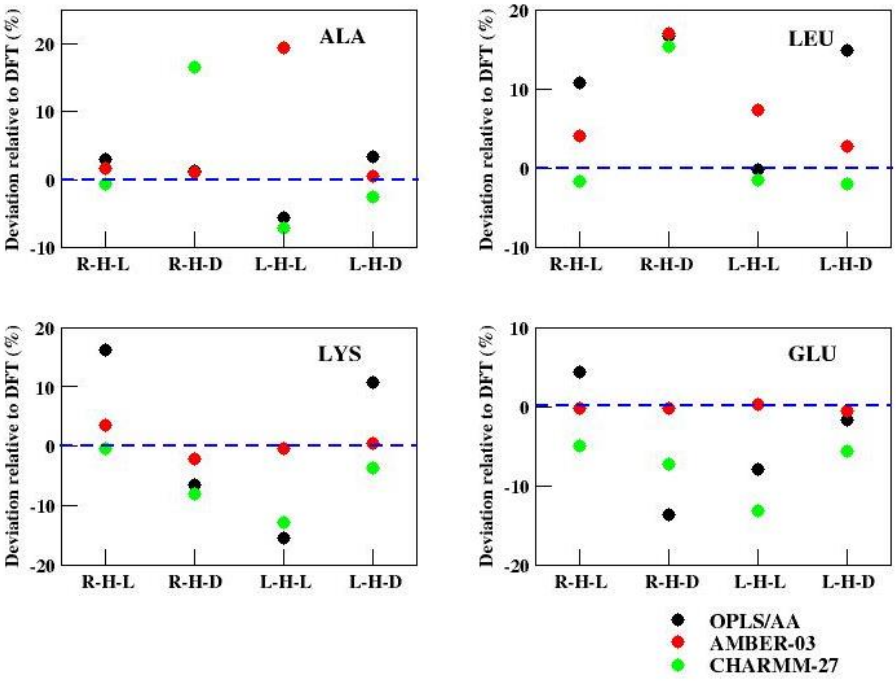


Figure 4.19: Percentage deviation of H_4 helical order parameter obtained from MD simulation relative to QM of 10-mer of ALA, LEU, LYS and GLU in OPLS/AA, AMBER-03 and CHARMM-27/CMAP force-fields in TIP3P water.

4.3.3 Quantification of Interaction between Side Chain and Main Chain

The initial structures for MD simulation are same as the structures taken before QM optimization. From the QM optimization, we obtained the relative energies of different conformers of α-helical polypeptides. The difference between their relative energies of these conformers mainly attributed to the interaction energies between side chain and main chain atoms. It has been verified by the calculation of non-bonded interaction energies between side chain and main chain atoms from MD simulation. The simulation trajectories, obtained from last 500 ps of 1 ns of production run have been analyzed to calculate the interaction energies between side-chain and main-chain atoms. The total of all pair and short range LJ and electrostatic interaction energies as well as long range electrostatic interaction energies between side-chain and main-chain atoms have been calculated from the last 500 ps of production MD run, and converted into histogram and distributions are plotted in Figure 4.20. The histogram plots clearly demonstrate the higher side chain – main chain interaction energies of R-H-D/L-H-L conformers with respect to R-H-L/L-H-D conformers for all the polypeptides and irrespective of any force fields used for simulation. The energy corresponding to the peak of the histogram is considered as interaction energy between side chain and main chain of the conformers. The relative energies of R-H-D, L-H-L, L-H-D conformers with respect to R-H-L conformer have been calculated for all the force field and shown in Table 4.4. The relative energy difference is the highest for CHARMM-27/CMAP force field among all the other force fields. The relative energy of R-H-D/L-H-L of LYS and GLU found to be lower than that of LEU, which is consistent with QM optimization result. Although, the energies obtained from QM optimization cannot be directly compared with energy from MD simulation as no thermal energy is involved during QM optimization but the side chain-main chain interaction energy calculated from MD can well explain the reasons behind the higher relative energies of R-H-D/L-H-L conformers with respect to R-H-L conformation of peptides.

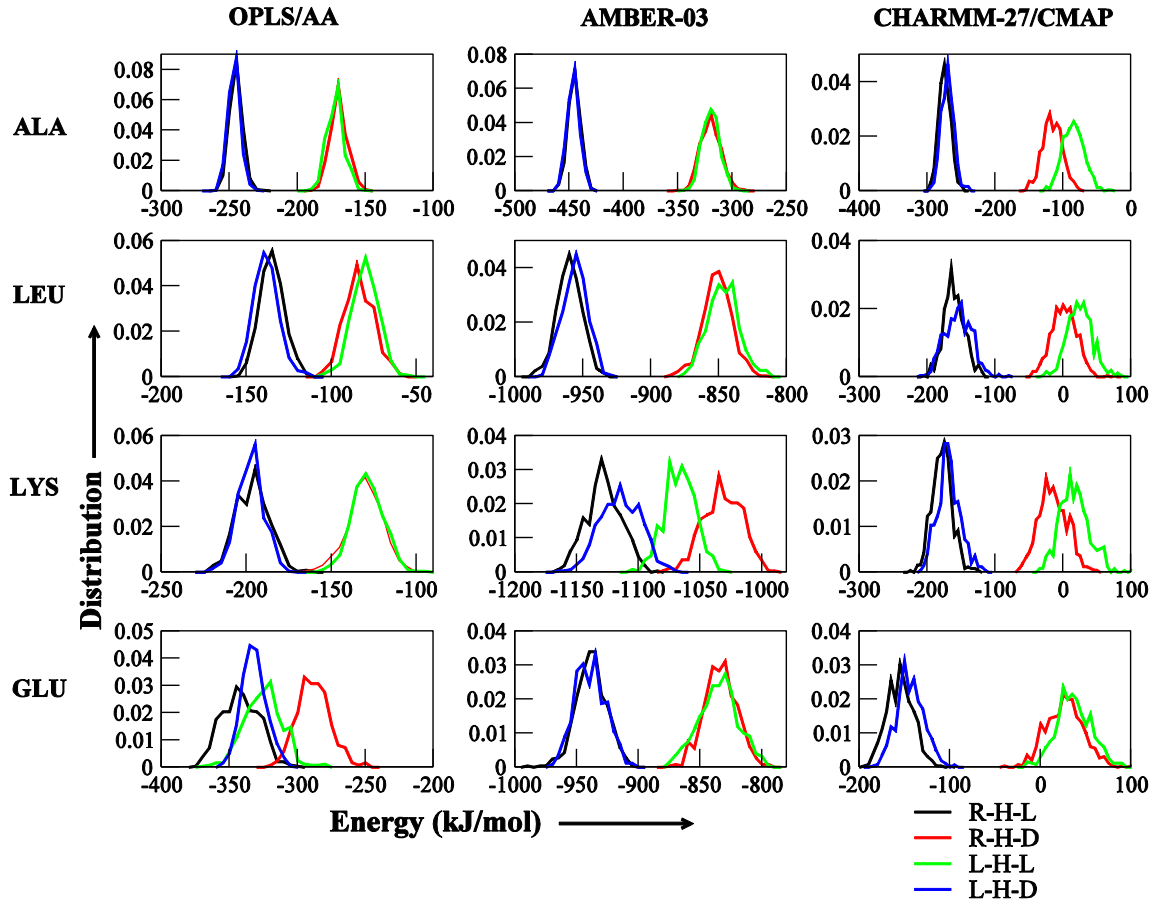


Figure 4.20: Distribution of side chain-main chain interaction energies for four different conformers of 10-mer homopolypeptides consiting of ALA, LEU, LYS and GLU amino acids from MD simulation using OPLS/AA, AMBER-03, CHARMM-27/CMAP force fields. The distribution plots of R-H-L, R-H-D, L-H-L, L-H-D conformers are shown by black, red, green and blue lines respectively.

Table 4.4: Difference in side-chain and main-chain interaction energies (kJ/mol) of different conformers of 10-mer homo-polypeptides of ALA, LEU, LYS, GLU with respect to their R-H-L conformer from MD simulation using OPLS/AA, AMBER-03, CHARMM-27/CMAP force fields.

	OPLS/AA			AMBER-03			CHARMM-27/CMAP		
	R-H-D	L-H-L	L-H-D	R-H-D	L-H-L	L-H-D	R-H-D	L-H-L	L-H-D
ALA	77	73	1	125	125	0.7	158	189	4
LEU	51	54	-3	109	113	3	162	185	6
LYS	66	67	1	96	62	13	163	195	10
GLU	54	16	10	105	103	0.7	182	192	11

4.4 Conclusion

The present study demonstrates how commonly used force fields for bimolecular simulations perform in representing helical conformations of homopolymeric peptide helices constructed from ALA, LEU, LYS and GLU. To achieve this objective, deviations with respect to the QM optimized helical configurations on using different classical MD force fields were quantified. 32 α-helical peptides (4 amino acids × 4 conformations of 5- and 10-mer from ALA, LEU, LYS and GLU) in four conformations (R-H-L, R-H-D, L-H-L and L-H-D) were considered for this study. Relative energies computed from QM calculations (using M06/6-311g (d, p) level of theory) indicated that R-H-L and L-H-D are lower energy conformations compared to R-H-D and L-H-L. Further, the helical conformations obtained after classical MD simulations using CHARMM-27/CMAP, AMBER-03, and OPLS/AA were compared with that obtained from QM calculations. Comparisons were performed depending upon two distinct structural properties of helical configurations. The first approach was from the equal contribution of backbone dihedrals (Φ and Ψ angles), and intra-helical *i-i+4* H-bonding, and the second one was from helical order parameters (H_2, H_3, H_4). From the Φ and Ψ dihedral angles and intra helical H-bonds contribution we observed that R-H-L, R-H-D, L-H-L and L-H-D conformers of all α-helical peptides of ALA, LEU, LYS and GLU are better represented by CHARMM-27/CMAP force-field than AMBER-03 and OPLS/AA force fields. From the helical order parameters (H_2, H_3, H_4) distribution obtained from MD simulation we observed that almost all conformers of all polypeptides remain in the α-helical region by using CHARMM-27/CMAP and AMBER-03 force-field, although AMBER-03 force field shows the better agreement of helical order parameters to QM than CHARMM-27/CMAP for most of the polypeptides.

References

[1] Kabsch, W., and Sander, C., "Dictionary of Protein Secondary Structure - Pattern-Recognition of Hydrogen-Bonded and Geometrical Features," Biopolymers 22, 2577-2637 (1983).

[2] Cintas, P., "Chirality of Living Systems: A Helping Hand from Crystals and Oligopeptides (Vol 41, Pg 1139, 2002)," Angewandte Chemie-International Edition 41, 2227-2227 (2002).

[3] Branden, C., and Tooze, J., "Introduction to Protein Structure," Garland Publishing, New York (1999).

[4] Zawadzke, L. E., and Berg, J. M., "The Structure of a Centrosymmetric Protein Crystal," Proteins: Structure, Function, and Bioinformatics 16, 301-305 (1993).

[5] Novotny, M., and Kleywegt, G. J., "A Survey of Left-Handed Helices in Protein Structures," Journal of Molecular Biology 347, 231-241 (2005).

[6] Wade, D., Boman, A., Wåhlin, B., Drain, C. M., Andreu, D., Boman, H. G., and Merrifield, R. B., "All-D Amino Acid-Containing Channel-Forming Antibiotic Peptides," Proceedings of the National Academy of Sciences 87, 4761-4765 (1990).

[7] Oren, Z., Hong, J., and Shai, Y., "A Repertoire of Novel Antibacterial Diastereomeric Peptides with Selective Cytolytic Activity," Journal of Biological Chemistry 272, 14643-14649 (1997).

[8] Colonna-Cesari, F., Premilat, S., Heitz, F., Spach, G., and Lotz, B., "Helical Structures of Poly(D-L-Peptides). A Conformational Energy Analysis," Macromolecules 10, 1284-1288 (1977).

[9] Nanda, V., and DeGrado, W. F., "Computational Design of Heterochiral Peptides against a Helical Target," Journal of the American Chemical Society 128, 809-816 (2006).

[10] Brown, R. A., Marcelli, T., De Poli, M., Solà, J., and Clayden, J., "Induction of Unexpected Left-Handed Helicity by an N-Terminal L-Amino Acid in an Otherwise Achiral Peptide Chain," Angewandte Chemie 124, 1424-1428 (2012).

[11] Martinek, T. A., and Fulop, F., "Peptidic Foldamers: Ramping up Diversity," Chemical Society Reviews 41, 687-702 (2012).

[12] Raghavachari, K., Trucks, G. W., Pople, J. A., and Head-Gordon, M., "A Fifth-Order Perturbation Comparison of Electron Correlation Theories (Reprinted from Chemical Physics Letters)," Chemical Physics Letters 589, 37-40 (2013).

[13] Stephens, P. J., Devlin, F. J., Chabalowski, C. F., and Frisch, M. J., "Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields," The Journal of Physical Chemistry 98, 11623-11627 (1994).

[14] Perdew, J. P., Burke, K., and Ernzerhof, M., "Generalized Gradient Approximation Made Simple," Physical Review Letters 77, 3865-3868 (1996).

[15] Zhao, Y., and Truhlar, D. G., "The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals," Theoretical Chemistry Accounts 120, 215-241 (2008).

[16] Zhao, Y., and Truhlar, D. G., "Density Functional for Spectroscopy: No Long-Range Self-Interaction Error, Good Performance for Rydberg and Charge-Transfer States, and Better Performance on Average Than B3lyp for Ground States," J Phys Chem A 110, 13126-13130 (2006).

[17] Zhao, Y., and Truhlar, D. G., "Density Functionals with Broad Applicability in Chemistry," Accounts Chem Res 41, 157-167 (2008).

[18] Hohenstein, E. G., Chill, S. T., and Sherrill, C. D., "Assessment of the Performance of the M05-2x and M06-2x Exchange-Correlation Functionals for Noncovalent Interactions in Biomolecules," J Chem Theory Comput 4, 1996-2000 (2008).

[19] Zhang, Y., Ma, N., and Wang, W. Z., "Assessment of the Performance of the M05-Class and M06-Class Functionals for the Structure and Geometry of the Hydrogen-Bonded and Halogen-Bonded Complexes," J Theor Comput Chem 11, 1165-1173 (2012).

[20] Walker, M., Harvey, A. J. A., Sen, A., and Dessent, C. E. H., "Performance of M06, M06-2x, and M06-Hf Density Functionals for Conformationally Flexible Anionic Clusters: M06 Functionals Perform Better Than B3lyp for a Model System with Dispersion and Ionic Hydrogen-Bonding Interactions," J Phys Chem A 117, 12590-12600 (2013).

[21] Jiang, J., Wu, Y., Wang, Z.-X., and Wu, C., "Assessing the Performance of Popular Quantum Mechanics and Molecular Mechanics Methods and Revealing the Sequence-Dependent Energetic Features Using 100 Tetrapeptide Models," J Chem Theory Comput 6, 1199-1209 (2010).

[22] Kaschner, R., and Hohl, D., "Density Functional Theory and Biomolecules: A Study of Glycine, Alanine, and Their Oligopeptides," J Phys Chem A 102, 5111-5116 (1998).

[23] Aleman, C., "A Conformational Study of the Alpha-L-Aspartate-Containing Dipeptide," J Phys Chem A 104, 7612-7616 (2000).

[24] Gould, I. R., Cornell, W. D., and Hillier, I. H., "A Quantum-Mechanical Investigation of the Conformational Energetics of the Alanine and Glycine Dipeptides in the Gas-Phase and in Aqueous-Solution," Journal of the American Chemical Society 116, 9250-9256 (1994).

[25] Jorgensen, W. L., Maxwell, D. S., and TiradoRives, J., "Development and Testing of the Opls All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids," Journal of the American Chemical Society 118, 11225-11236 (1996).

[26] Kaminski, G. A., Friesner, R. A., Tirado-Rives, J., and Jorgensen, W. L., "Evaluation and Reparametrization of the Opls-Aa Force Field for Proteins Via Comparison with Accurate Quantum Chemical Calculations on Peptides†," The Journal of Physical Chemistry B 105, 6474-6487 (2001).

[27] Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A., "A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules (Vol 117, Pg 5179, 1995)," Journal of the American Chemical Society 118, 2309-2309 (1996).

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

[29] Schuler, L. D., Daura, X., and Van Gunsteren, W. F., "An Improved Gromos96 Force Field for Aliphatic Hydrocarbons in the Condensed Phase," Journal of Computational Chemistry 22, 1205-1218 (2001).

[30] Oostenbrink, C., Villa, A., Mark, A. E., and Van Gunsteren, W. F., "A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: The Gromos Force-Field Parameter Sets 53a5 and 53a6," Journal of Computational Chemistry 25, 1656-1676 (2004).

[31] Boehm, H. J., and Brode, S., "Ab Initio Scf Calculations on Low-Energy Conformers of N-Acetyl-N'-Methylalaninamide and N-Acetyl-N'-Methylglycinamide," Journal of the American Chemical Society 113, 7129-7135 (1991).

[32] Gould, I. R., and Kollman, P. A., "Ab Initio Scf and Mp2 Calculations on Four Low-Energy Conformers of N-Acetyl-N'-Methylalaninamide," The Journal of Physical Chemistry 96, 9255-9258 (1992).

[33] García, A. E., and Sanbonmatsu, K. Y., "A-Helical Stabilization by Side Chain Shielding of Backbone Hydrogen Bonds," Proceedings of the National Academy of Sciences 99, 2782-2787 (2002).

[34] Mackerell, A. D., Feig, M., and Brooks, C. L., "Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations," Journal of Computational Chemistry 25, 1400-1415 (2004).

[35] Best, R. B., Buchete, N. V., and Hummer, G., "Are Current Molecular Dynamics Force Fields Too Helical? (Vol 95, Pg L07, 2008)," Biophys J 95, 4494-4494 (2008).

[36] Lange, O. F., van der Spoel, D., and de Groot, B. L., "Scrutinizing Molecular Mechanics Force Fields on the Submicrosecond Timescale with Nmr Data," Biophys J 99, 647-655 (2010).

[37] Cino, E. A., Choy, W. Y., and Karttunen, M., "Comparison of Secondary Structure Formation Using 10 Different Force Fields in Microsecond Molecular Dynamics Simulations," J Chem Theory Comput 8, 2725-2740 (2012).

[38] Best, R. B., and Hummer, G., "Optimized Molecular Dynamics Force Fields Applied to the Helix−Coil Transition of Polypeptides," The Journal of Physical Chemistry B 113, 9004-9015 (2009).

[39] Piana, S., Lindorff-Larsen, K., and Shaw, David E., "How Robust Are Protein Folding Simulations with Respect to Force Field Parameterization?," Biophys J 100, L47-L49.

[40] Lindorff-Larsen, K., Maragakis, P., Piana, S., Eastwood, M. P., Dror, R. O., and Shaw, D. E., "Systematic Validation of Protein Force Fields against Experimental Data," Plos One 7, (2012).

[41] Pandey, P. R., and Roy, S., "Early Stages of Unwinding of Zwitterionic A-Helical Homopolymeric Peptides," Chemical Physics Letters 514, 330-335 (2011).

[42] Pandey, P. R., and Roy, S., "Distinctions in Early Stage Unwinding Mechanisms of Zwitterionic, Capped, and Neutral Forms of Different A-Helical Homopolymeric Peptides," The Journal of Physical Chemistry B 116, 4731-4740 (2012).

[43] Chécot, F., Lecommandoux, S., Klok, H. A., and Gnanou, Y., "From Supramolecular Polymersomes to Stimuli-Responsive Nano-Capsules Based on Poly(Diene-B-Peptide) Diblock Copolymers," Eur. Phys. J. E 10, 25-35 (2003).

[44] Cai, C., Zhang, L., Lin, J., and Wang, L., "Self-Assembly Behavior of Ph- and Thermosensitive Amphiphilic Triblock Copolymers in Solution: Experimental Studies and

Self-Consistent Field Theory Simulations," The Journal of Physical Chemistry B 112, 12666-12673 (2008).

[45] Cai, C., Lin, J., Chen, T., Wang, X.-S., and Lin, S., "Super-Helices Self-Assembled from a Binary System of Amphiphilic Polypeptide Block Copolymers and Polypeptide Homopolymers," Chemical Communications 2709-2711 (2009).

[46] Naik, S. S., Ray, J. G., and Savin, D. A., "Temperature- and Ph-Responsive Self-Assembly of Poly(Propylene Oxide)-B-Poly(Lysine) Block Copolymers in Aqueous Solution," Langmuir 27, 7231-7240 (2011).

[47] Breedveld, V., Nowak, A. P., Sato, J., Deming, T. J., and Pine, D. J., "Rheology of Block Copolypeptide Solutions: Hydrogels with Tunable Properties," Macromolecules 37, 3943-3953 (2004).

[48] Gazit, E., "Self-Assembled Peptide Nanostructures: The Design of Molecular Building Blocks and Their Technological Utilization," Chemical Society Reviews 36, 1263-1269 (2007).

[49] Sarkar, S., Pandey, P. R., and Roy, S., "Propensity of Self-Assembled Leucine-Lysine Diblock Copolymeric A-Helical Peptides to Remain in Parallel and Antiparallel Alignments in Water," The Journal of Physical Chemistry B 119, 9520-9531 (2015).

[50] Pati, D., Shaikh, A. Y., Das, S., Nareddy, P. K., Swamy, M. J., Hotha, S., and Gupta, S. S., "Controlled Synthesis of O-Glycopolypeptide Polymers and Their Molecular Recognition by Lectins," Biomacromolecules 13, 1287-1295 (2012).

[51] Nick Pace, C., and Martin Scholtz, J., "A Helix Propensity Scale Based on Experimental Studies of Peptides and Proteins," Biophys J 75, 422-427 (1998).

[52] Tomasi, J., Mennucci, B., and Cammi, R., "Quantum Mechanical Continuum Solvation Models," Chem Rev 105, 2999-3094 (2005).

[53] Bjelkmar, P., Larsson, P., Cuendet, M. A., Hess, B., and Lindahl, E., "Implementation of the Charmm Force Field in Gromacs: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models," J Chem Theory Comput 6, 459-466 (2010).

[54] Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G. M., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J. M., and Kollman, P., "A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations," Journal of Computational Chemistry 24, 1999-2012 (2003).

[55] Barlow, D. J., and Thornton, J. M., "Helix Geometry in Proteins," Journal of Molecular Biology 201, 601-619 (1988).

[56] Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J., Gaussian, Inc., Wallingford CT (2009).

[57] Tomasi, J., Mennucci, B., and Cammi, R., "Quantum Mechanical Continuum Solvation Models," Chem Rev 105, 2999-3093 (2005).

[58] Boys, S. F., and Bernardi, F., "The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures with Reduced Errors," Molecular Physics 19, 553-566 (1970).

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

[60] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L., "Comparison of Simple Potential Functions for Simulating Liquid Water," The Journal of Chemical Physics 79, 926-935 (1983).

[61] Patra, M., Karttunen, M., Hyvönen, M. T., Falck, E., Lindqvist, P., and Vattulainen, I., "Molecular Dynamics Simulations of Lipid Bilayers: Major Artifacts Due to Truncating Electrostatic Interactions," Biophys J 84, 3636-3645.

[62] Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R., "Molecular Dynamics with Coupling to an External Bath," The Journal of Chemical Physics 81, 3684-3690 (1984).

Chapter 5

Structural integrity of bilayers formed by glycopolypeptide-polypropylene oxide di-block copolymers: Effect of chiral heterogeneity in glycopolypeptide

5.1 Introduction

Glycopeptides are peptides with sugar groups attached to the amino acid residues. There are different cell surface protein receptors like lecitin which specifically interacts with carbohydrate of glycopolymers to enhance its biomolecular recognisation.[1] Carbohydrate moieties play a crucial role in different processes such as tissue regeneration, intercellular recognition, cell adhesion, immunological recognition.[2] The pendent carbohydrate groups attached to polypeptide backbone attributes different biological functionality in glycopeptides. Thus glycopolymers have drawn attention in field of chemistry, biology, and medicine. Its biological activities can be tuned for different biomedical purposes. Glycopolymers is abundant in the bacterial cell wall. The vesicular structures formed by glycopeptides have the similarity with the structure of the viral capsid made of glycoproteins.[3] Glycopolymers are also favourite for its usage as antibiotics that inhibit the formation of bacterial cell wall and folds into different types of secondary structures like α-helix and β-sheet.

Glycoproteins show different biological activities because of its occurrence in different macromolecules like enzyme, immunoglobulin, cell adhesion molecules, etc. It has encouraged researchers to synthesize the conjugates where carbohydrate is attached to the hydrophobic chains.[4, 5] These diblocks chains can self-assemble into different structures like vesicles[3], polymersomes[6, 7], nanorods[8]. Glycopeptides itself can also self-assemble into well defined structure like hydrogels.[9] This is used in different medicinal applications like eye surgery because of its toxic nature, and ability of hindering postoperative fibrosis.[10-12] Glycopeptide having a lactose unit attached to the C-terminal of hydrophobic helical peptide forms vesicle in water.[13] Now-a-days, glycopolymers are the material of very high importance because of their novel applications in drug delivery, tissue engineering, and scaffolds.[14-16] Glycosylation helps in the process of self-assembly of some peptides like diphenylalanine which is less soluble and difficult to study for the self-assembly. Glyscosylation of this enhances its solubility and improves its self-assembling properties.[17] Antitumor vaccine consisting of glycopeptides conjugates enhances specificity towards tumor cells. The vaccine is designed by multilayer self-assembly of positive T-helper cell peptide epitope at core, negatively charged polyglutamic acid in inner shell, and positively charged MUC1 glycopeptides bearing antigen at outer shell.[18] Some glycopeptides enhance the bioactivity of carbohydrate binding proteins. Galectin specifically interacts with the cell surface and extra-cellular matrix glycoprotein controlling cell

signalling in different cellular process, whose bioactivity helps in different medical applications, tissue regeneration, cancer therapeutics, etc. Restuccia et al [19] have mimicked the natural glycoprotein-galectin interaction by attaching a N-acetylglucosamine at the N-terminating asparagines residue of a peptide.

The knowledge of self-assembly and interaction of hydrophobic-hydrophilic segments can be applied for designing the systems of highly specific applications. The micelles consisting of hydrophobic core can enclose hydrophobic drugs where the hydrophilic shell on the exterior makes it water-soluble. Polymersome is a hollow sphere which contains an aqueous solution in the core enclosed by bi-layer membrane. They are potential candidate for carrying drug as they are stable in throughout the blood circulation time.[20] People have synthesized polymersomes of different thickness and permeability by varying their molecular weight and compositions.[21, 22] The majority of the studies have been carried out upon it for using it as a drug carrier in cancer therapy. The maximum number of drugs fails to cure cancer due to its poor water solubility, bioavailability at different receptor sites, and severe side effects. The encapsulation of drugs within nanocarriers can increase its bioavailability by acquiring poorly water-soluble drugs in water compatible carriers. Furthermore, it can protect the drug from the releasing into other organs, thus it minimizes the side effects. For nanocarriers to deliver drug selectively to a tumour, they need to have a long enough blood circulation half-life for reaching to tumour cells. The polymersome are prepared to be stimuli-responsive during its synthesis so that it can be used more smartly in response of pH, temperature, redox potential etc. It can also be modified to use for targeted drug delivery by attaching a biological reorganisation unit on the surface. Polymersome consisting of hydroxyproline rich glycoprotein block co-polymers form scaffolds by aggregation of hydrophobic peptide and hydrophilic glycopeptides.[23] Polypeptide-based copolymers show considerable promise as building blocks for polymersomes. Holowka et al have found that the helical conformation of a hydrophobic peptide segment like poly (γ-benzyl-L-glutamate) (PBLG) gives rise to a stable vesicle.[24]

Generally, diblock copolymers of different hydrophobicity form compact bilayer structure. Several experimental techniques and molecular simulations have been successfully carried out to investigate the structures and properties of vesicles and polymersomes. The stability and conformation of a bilayer depends upon the secondary structure of the constituting chain. The secondary structure of glycopeptides can be α-helical where peptide

sequence is an important factor maintaining helicity.[25, 26] Das et al[27] have studied the self-assembly of glycopeptides-b-poly-(propylene oxide) in aqueous solution to form polymersomes. Polypropylene is a FDA- approved hydrophobic component, and becomes biocompatible when it's attached to hydrophilic glycopeptides. It can encapsulate both hydrophobic and hydrophilic drugs simultaneously. Their study revealed that the packing of the ordered helical glycopolypeptide segment is required for the self-assembly into spherical morphology.

Huang et al[28] have varied the chain length of hydrophilic segment containing galactose in the b-poly (γ-benzyl-L-glutamate) –b (galactosylated propargylglycine) di-block chain and observed formation of various type self-assembled structures. There is no macroscopic assembly occurs at lower weight ratio (upto ~38%) of hydrophilic segment, but at higher weight ratio (~ 55 % - 68 %) mixture of spherical and worm-like structure are found. The constituent of the backbone of glycopeptide and its secondary structure is crucial in self-assembly process. It's assumed that the α-helicity content of di-block chains plays a role in stabilising bilayer. In nature, protein mainly consists of l-amino acids. But, the existence of d-amino acids is also possible.[29, 30] There are many recent studies on the self-assembly of d-amino acids. The chirality of amino acid of helical glycopeptides blocks effect its percentage α-helicity in the self-assembled bilayer. Therefore, our aim is to investigate whether the α-helicity content of hydrophilic glycopeptides block have any influence on the stability of the bilayer.

Thus, we have performed MD simulation of two bilayers constituting of two types of di-block of b-glycopolypetide-b-polypropylene oxide (b-GP-b-PPO) chains. In one chain, glycopolypeptide block consists of purely l-amino acids, another is comprised of alternate d, l-amino acids. The chains are packed into hexagonal closed packed arrangement to form each monolayer of the bilayer as it gives the best packing efficiency. The bilayer where glycopeptides block consist of purely l-amino acid is found to be substantially stable and the chains remained integrated throughout the simulation time. We intended to understand the key interaction which controls the stability of bilayers.

5.2 Computational Details

The all-atomistic molecular dynamics simulations of bilayers consisting of diblocks chains of capped GP_{12}-PPO_{44} chain were performed. The N terminal of hydrophilic GP block was

capped by acetyl group (ACE), whereas hydrophobic PPO block was capped by butane (BUT). GP blocks of two kinds of bilayer consist of all l-amino acid, and alternate l, and d amino acids. All the MD simulations were carried out using Gromacs 4.6.3. pacakage.[31] The force field parameters of glycopeptides part were taken from the work of Das et al.[27] The atomic charges of the atoms of PPO and terminal butane groups were calculated by multi-conformational fitting of them using RESP method.[32] The O-C-C-H dihedral angle of PPO and C-C-C-C dihedral of BUT varied form 0° to 180° to generate 7 different initial configurations which were used for fitting of atomic charges. The charges were calculated using Hartree-Fock method and 6-31g (d) basis set. All the fitting were made using R.E.D. server.[33, 34] All the force field parameters were consistent with AMBER99SB force field. The force field for PPO were validated by calculating the radius of gyration of PPO_7 polymer chains in water and density of the melt system. The details about the simulation of single PPO_7 polymer chains in water and for melt polymer system are given in the following sections. The density of the melt system at 298 K was found to be 1005 kg/m^3, very close to experimental value[35] (Figure 5.3).

5.2.1 Simulation of single PPO_7 chain in water

A single chain of PPO_7 chain was solvated into water. It was energy minimized using steepest descent algorithm, and then the 20 ns of simulation run was carried out. The system was simulated at temperature of 300 K and in pressure of 1 atm. The temperature of the system was maintained using V-rescale thermostat with coupling constant of 0.1 ps. The pressure of the system was kept at 1 atm using Berendsen Barostat with 0.1 ps coupling constant. The Vander Waal interaction was calculated from LJ 12-6 interaction upto cut-off value of 0.9 nm and direct electrostatic interaction was determined using coulombs law within 1.4 nm of atom to atom distance. Long range electrostatic interaction was taken care of using PME electrostatics. R_g of polymer chain was calculated by averaging over last 10 ns of simulation time and converted into distribution (Figure 5.1). The average radius of gyration (0.5 nm) of single chain in water was in very well agreement with the work of Hezaveh et al.[36]

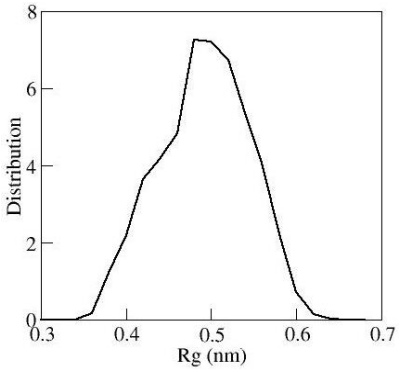


Figure 5.1: Distribution of radius of gyration of polypropylene oxide chain over last 10 ns of simulation time.

5.2.2 Simulations of PPO_7 melt system

The PPO_7 melt system was constructed by randomly packing 121 polymer chains into a box of volume (5.7 × 5.89 × 6.19) nm^3. The system was energy minimized and then NPT simulation was performed at 300 K. All the simulation parameters were kept same as previously mentioned except the compressibility value of 3.0×10^{-5} which is used for the melt system.[36]. The density of the system was converged within 5 ns of simulation time. The simulation was performed for 17 ns and density with time plot (figure 5.2) shows that density of melt system is 1005 kg/m^3 at 300 K. The result matches with experimental observation.[35]

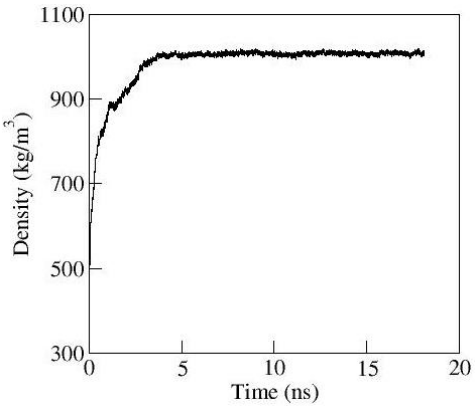
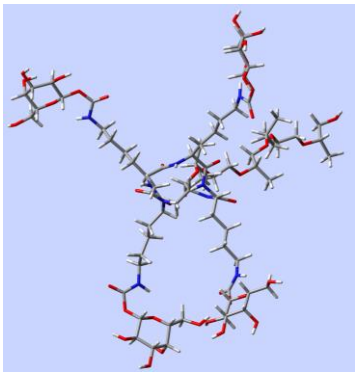


Figure 5.2: Density of polypropylene oxide polymer melt with simulation time.

5.2.3 Preparation of the systems

The initial configuration of the ACE-LYG_{12}-PPO_{44}-BUT is prepared using geometrical parameters of quantum mechanical (QM) optimized structure using HF/6-31g (d) method of smaller chain length ACE-LYG_4-PPO_7-BUT. The optimized structure is given in Figure 4. Total 4 residues of glycopeptide were taken for giving one complete turn for α-helix. The QM optimization produces the appropriate conformation prior to the lowest energy of the side chains of LYG for starting MD simulation. The bond angle and dihedral of connection of PPO and LYG residues obtained from QM optimization have been used to build the initial conformation of long length ACE-LYG_{12}-PPO_{44}-BUT chain.



5.3: QM optimized struture of ACE-LYG4-PPO7-BUT using HF/6-311 g (d,p) basis set

5.2.3.a Glycopeptides block consisting of l-amino acids

A single chain of ACE-LYG_{12}-PPO_{44}-BUT consisting of l-amino acids was solvated in water. NPT simulation run was performed in water to attain a globular conformation of hydrophobic PPO chain beneath the hydrophilic glycopeptide block. The radius of gyration converged after 20 ns of simulation (figure 5.4). The structure obtained after 20 ns of simulation was used as constituent for preparing bilayer. The chains were packed into hexagonal closed packed arrangement in each monolayer for obtaining the highest packing efficiency (74%). The bilayer system is first energy minimized using steepest descent method. Then, 1 ns of semi-isotropic NPT simulation was performed so that the polymer chains packed more closely through the interaction among them. The conformation obtained after 1 ns of simulation was taken, and then water was added to both sides of hydrophilic glycopeptides part so that it faces water and hydrophobic part remain buried inside bilayer. The system was energy minimized and then semi-isotropic NPT simulation was performed. Semi-isotropic

pressure coupling was applied by individually coupling to the xy plane and z-direction (normal to bilayer-water interface).

5.2.3.b Glycopeptides block consisting of alternate d, l-amino acid

The same polymer chain was constructed with the glycopeptides block consisting of alternate d and l-amino acids residues. The initial structure of d-amino acid obtained from the conformation obtained from QM optimization of ACE-LYG_4-PPO_7-BUT chain by changing sign of the dihedral angles associated with the C1 atom of the attached side chain and hydrogen atom with Cα of the was converted into d-amino acids. The chain was solvated into water, then energy minimized. Then 400 ps of NPT run was carried out by position restraining PPO atoms as side chains of glycopeptides can relax. Then removing positional restraint of PPO atoms, NPT simulation run was carried out so that long PPO chains get coiled to form a hydrophobic core below the glycopeptide. After 6.5 ns of simulation, the radius of gyration (Figure 5.4) of PPO chain becomes almost similar to the PPO of the chain where GP block consisting of l-amino acid. The conformation which obtained after 6.5 ns simulation was used to make the bilayer-water system in same way as discussed previously.

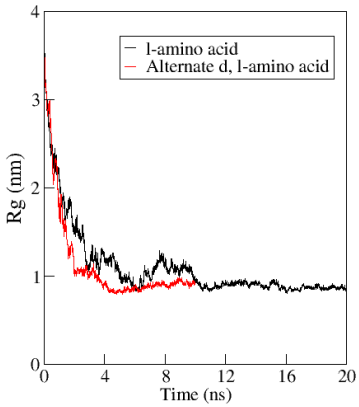


Figure 5.4: Radius of gyration of PPO chain of di-block glycopeptide-PPO copolymer chain in water. The diblock in which glycopeptide consist of pure l-amino acid, and alternate d, l – amino acids are shown by black, and red colours, respectively.

5.3 Results and Discussion

The bilayer having glycopeptide block consisting of l-amino acids remain integrated after long time simulation (980 ns). Water molecules enter into the exterior of bilayer consisting of

hydrophilic glycopeptide region but they don't penetrate the hydrophobic polypropylene oxide (PPO) core region of the bilayer. The bilayer consisting of alternate d, l-amino acids is simulated for 600 ns and water molecule is found to invade the bilayer from the starting of simulation. The diblock chains of the middle portion of bilayer move aways from each other and water molecules enters into the core region of bilayer. The snap shots after 980 ns of simulation of bilayer consisting of l-amino acid and 600 ns simulation of bilayer consisting alternate d, l-amino acids are given in the Figure 5.5.

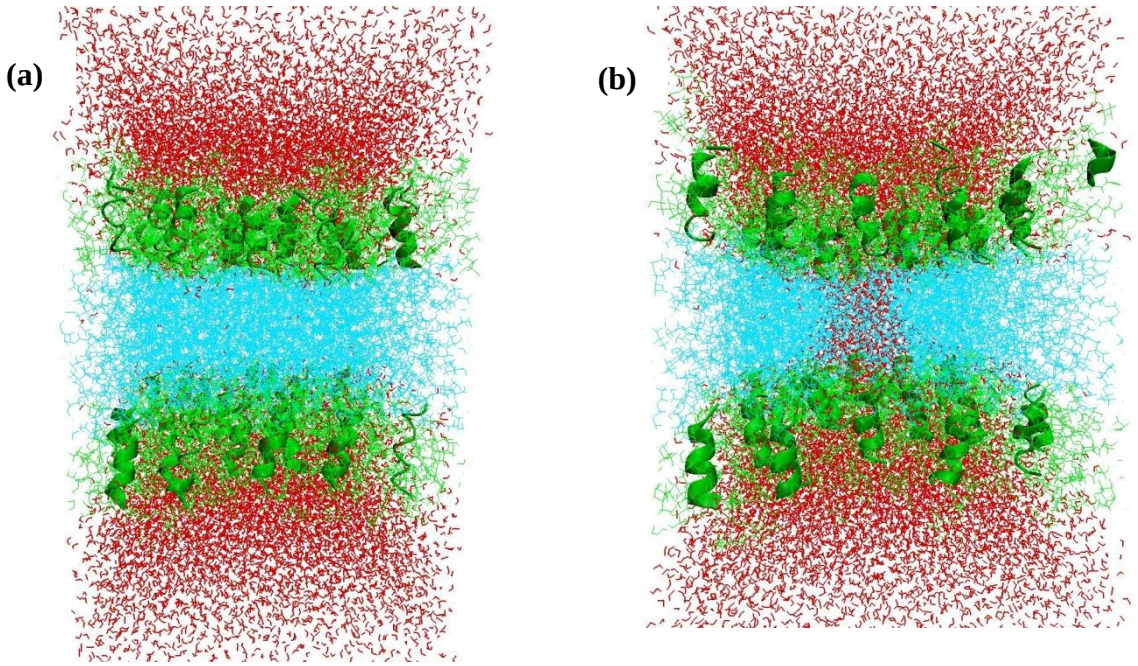


Figure 5.5: The Snapshot of bilayers consisting of diblocks chains of glycopeptide-polypropylene oxide (*b*-GP-*b*-PPO) in water. (a) The structure of the bilayer after 980 ns where the GP block consisting of l-amino acids. (b) The snapshot of bilayer after 600 ns simulation run where GP block consisting of alternate d, l-amino acids. GP, PPO, and water are shown by green, cyan, and red colours respectively.

The penetration of water molecules in the bilayer is well understood from the calculation of partial densities of hydrophobic PPO, hydrophilic glycopeptides and water molecules along the normal of bilayer-water interface. The average partial densities of different components are calculated over different time frames for 2 ns and shown in figure 5.6. From the figure 5.6a, it's evident that water does not enter into core hydrophobic region of PPO of bilayer constituting of l-amino acids, whereas for bilayer consisting of d, l-amino acids water starts

entering into the hydrophobic region from the beginning of simulation. The huge amount of water penetrated in the PPO region disintegrates the bilyer and the chains are getting apart from each other with the progress of simulation time.

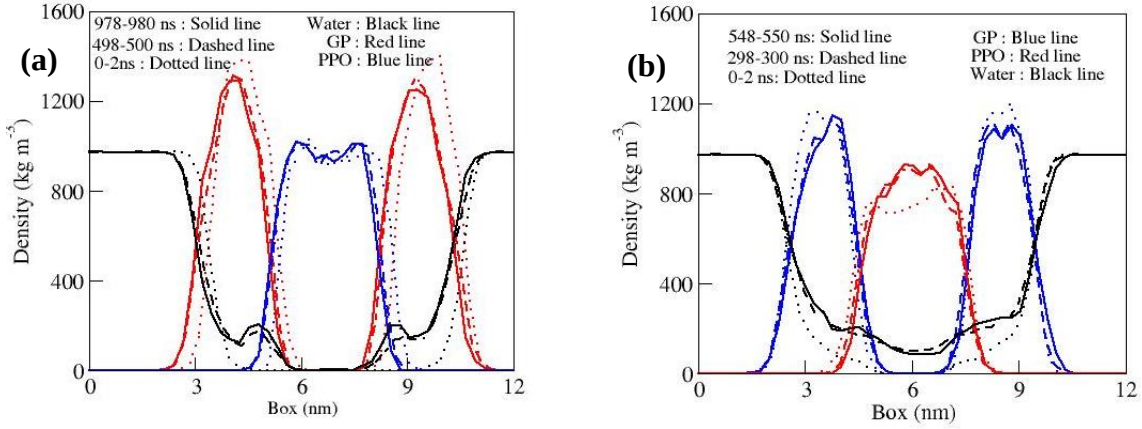


Figure 5.6: Partial densities of GP, PPO and water along the normal of bilayer-water interface at different time intervals of bilayer consisting of (a) l-amino acid. (b) alternate d, l-amino acids.

The average area per chain shows the integrity of the chains in the bilayer. The area per chain is calculated by dividing the area of simulation box which is bilayer-water interface by the number of polymer chains. The average area per chain with function of time is shown in figure 5.7. The area per chain found to be less for the bilayer consisting of only l-amino acid which clearly indicates that they remain in more compact form and occupy less surface area (~1.38 nm^2/chain) whereas the chains of alternate d, l-amino acid are far apart (~1.57 nm^2/chain) from each other.

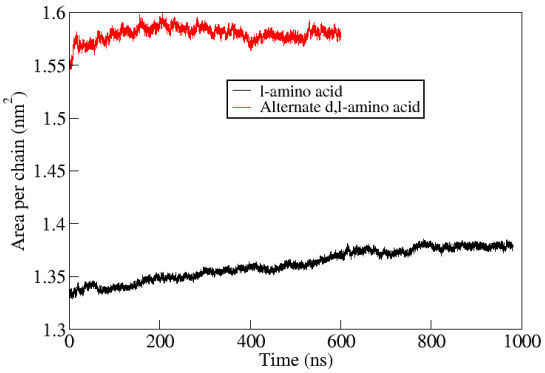
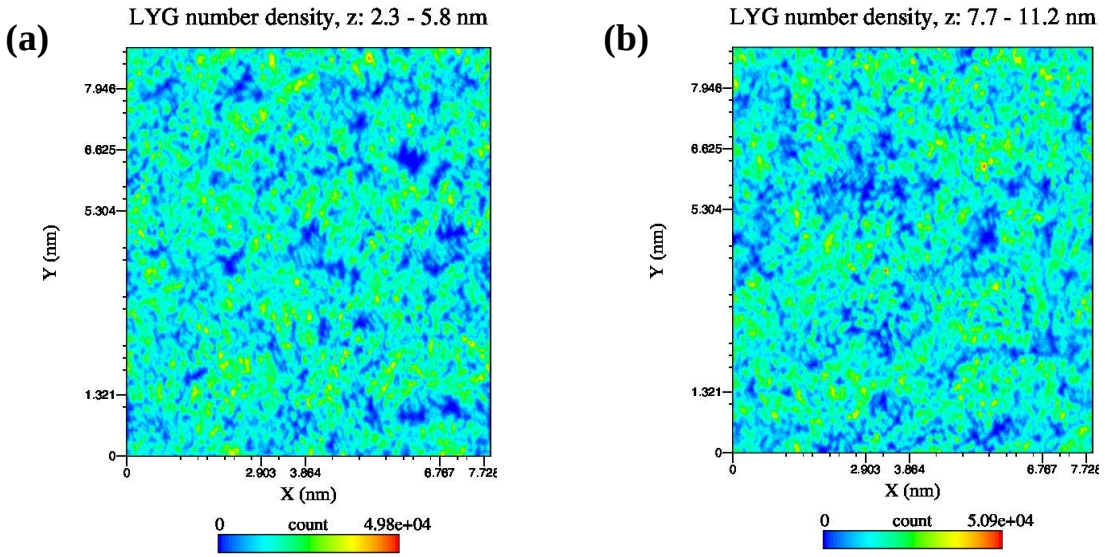


Figure 5.7: Area per polymer chain with simulation time for bilayers consisting of l amino acid (black) and alternate d, l-amino acids (red).

From the area per chain, it's understood that the packing of di-block chains is more compact in bilayer where glycopeptide block consist of purely l-amino acids. Therefore, to know the positional arrangement of the chains, we calculated the local density of all the chains in the plane parallel to water-bilayer interface. The chains are not homogeneously distributed throughout the bilayer. The penetration of water molecules in the bilayer alters the positional distribution of chains because of its interaction with glycopeptides. The side chains of glycopeptides have different arrangements in two bilayers because of different chirality of amino acid residues. It results into the distinct interactions among the glycopeptides chains in two bilayers. Thus, the penetration of water molecules affects differently the glycopeptides which leads into different distribution of them. The average densities of the glycopeptides block of both the monolayers have been calculated separately. This density is represented by colour coding; the low density to high density is represented by varying colour from blue to red (Figure 5.8). The homogeneous distribution of green and blue colours in lower (Figure 5.8a) and upper monolayer (Figure 5.8b) of bilayer indicates that the glycopeptides is uniformly distributed in both the monolayers of glycopeptides consisting of l-amino acids. The blue patches present in the middle of the both lower (Figure 5.8c) and upper monolayer (Figure 5.8d) of the bilayer consisting of alternate d, l-amino acid indicate that there are many void spaces occupied by water.



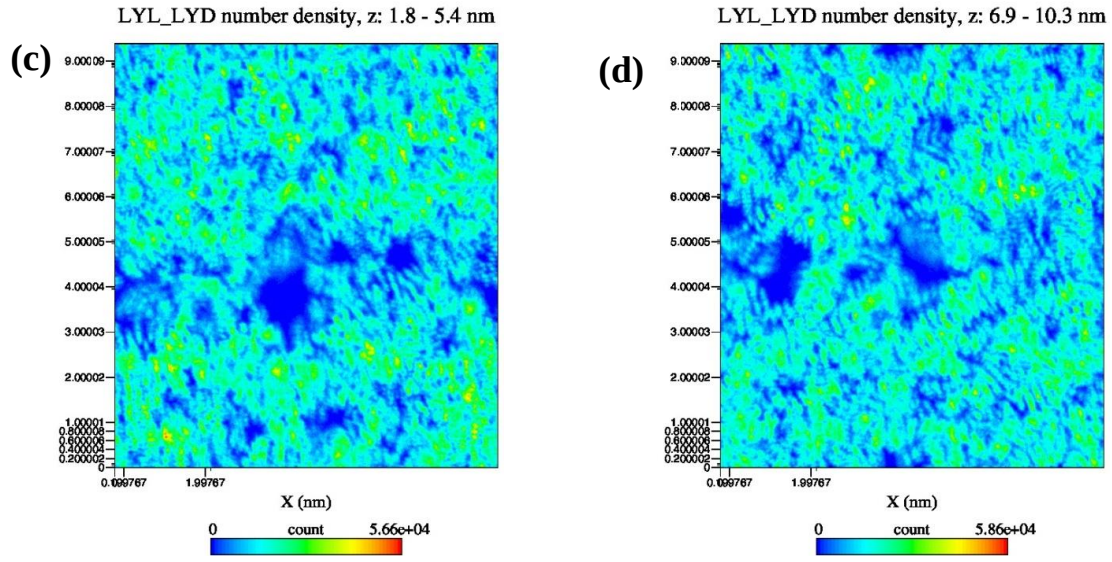


Figure 5.8: 2D density map of glycopeptide in xy plane consisting of (a) block of l-amino acids of lower monolayer (b) block of l-amino acids of upper monolayer (c) block of alternating l and d amino acids of lower monolayer. (d) block of alternating l and d amino acids of upper monolayer.

As the simulation progresses the α-helical glycopeptides chains started to unwind. The primary difference between these two bilayers is the α-helical glycopeptide block. Thus, whether there is any role of α-helicity behind the dissembling of bilayer containing alternate d, l-amino acid chains is required to be investigated. Therefore, the number of glycopeptide residues in α-helical form is calculated. If the hydrogen bonding energy between CO of i-th residue NH of (i+4) th residue is less than -0.5 kcal/mol then it's considered to be in alpha-helical form.[37] The percentage of such number residues is calculated and plotted as a function of simulation time (figure 5.9b, 5.9c), and also converted into their distribution (figure 5.9a). It's found that the polymer chains of alternate d, l-amino acids contain more percentage α-helicity (~46 %) than that of l-amino acids (~32%). The helicity of a single di-block chain in water has also been calculated. It's found to be slightly higher for the glycopeptide segment containing alternating d, l amino acid (Figure 5.10).

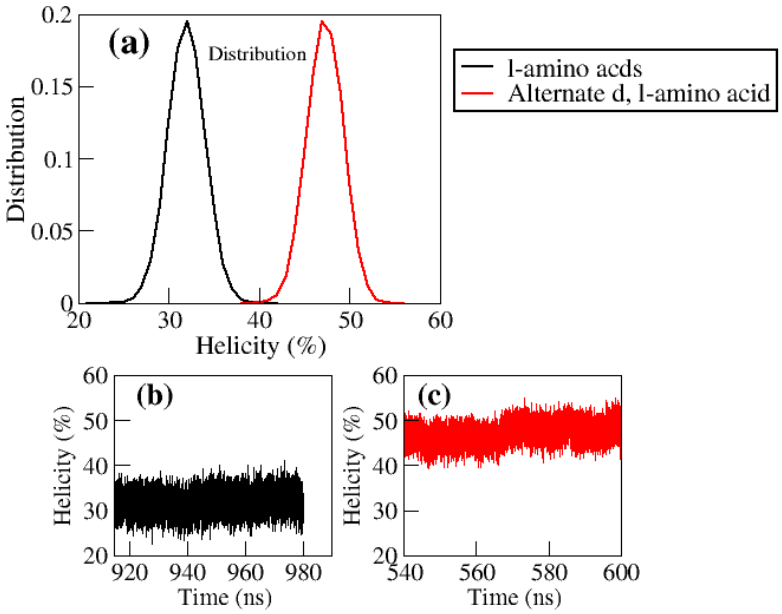


Figure 5.9: (a) Distributions of percentage α-helicities of bilayers consisting of l amino acid (black) over last 60 ns of 980 ns mdrun and alternate d, l-amino acids (red) over last 60 ns of 600 ns mdrun (b) The percentage α- helicities with time of bilayer consisting of l-amino acid, (c) The percentage α-helicities with time of bilayer consisting of alternate, d, l-amino acid.

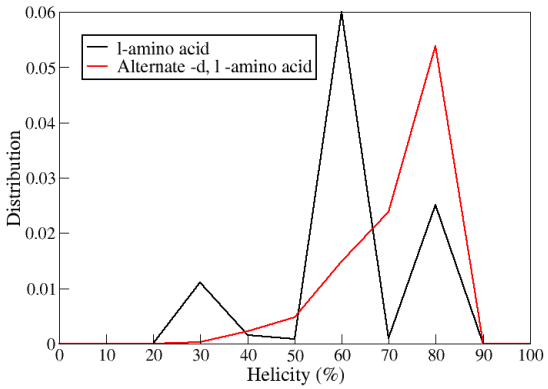


Figure 5.10: Distributions of percentage of all types of helicities of single chains in water consisting of l-amino acid (black) and alternate d, l-amino acids (red) over 450 ns of simulation time.

It has been found that the chains of bilayers remained integrated and uniformly distributed in the assembled structure. The reason behind difference of integrity of the two bilayers is

primary motive of our study. There are many hydrogen bond forming sites in glycopeptides chain. The glucose attached to the side chain of lysine has five OH groups which can take part in hydrogen bond formation. Moreover, there are CO and NH groups in between of glucose and side chain of lysine. These sites can form side chain – side chain H-bonds. The main chain of peptide also contains CO and NH groups which are potential site for forming side chain – main chain hydrogen bonds.

The hydrogen bonds have been calculated from geometrical aspects. The distance distribution between hydrogen and hydrogen bond acceptor for possible sites of interaction have been calculated and converted into distributions (Figure 5.11). The first minimum in the distribution plots around 0.25 nm indicates that the H-bond is possible in between side chain OH – side chain OH, side chain OH and side chain CO, side chain OH and side chain NH. Therefore, the distance criteria of hydrogen – hydrogen bond acceptor of the hydrogen bond is considered to be less than 0.25 nm. The angle formed by hydrogen bond donor – hydrogen – hydrogen bond acceptor was taken to be greater than 130° as the criteria for H-bond formation.

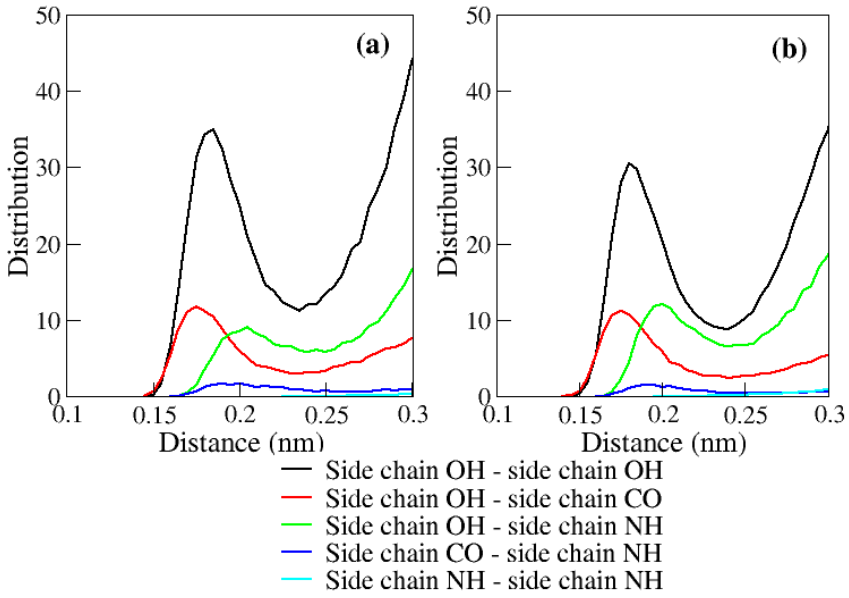


Figure 5.11: Distance distribution between H and H-bond acceptor for different possible hydrogen bonding sites (OH, CO, and NH groups) of side chain for bilayer of which glycopeptides part consisting of (a) l-amino acids, and (b) alternate l, d amino acids.

In each chain, the glucose OH group, CO and NH group in the side chain of different glycopeptides residues can form hydrogen bonds among themselves. This results into formation of intra side chain – side chain hydrogen bond. To explore whether is there is any relationship in between α-helicity of glycopeptides block and intra side chain hydrogen bonds, we calculated the intra side chain – side chain hydrogen bond. We have calculated the number of side chain - side chain hydrogen bonds with simulation time (Figure 5.12a, 5.12b) and converted into their distribution (Figure 5.12c). Although the number is found to be slightly higher for bilayer consisting of alternate d, l- amino acids, but, their difference is insignificant with respect to the difference in percentage α-helicities of glycopeptide blocks of these two bilayers. Therefore, it can be concluded that side chain- side chain hydrogen bonds does not affect percentage α-helicity of glycopeptides blocks of bilayers.

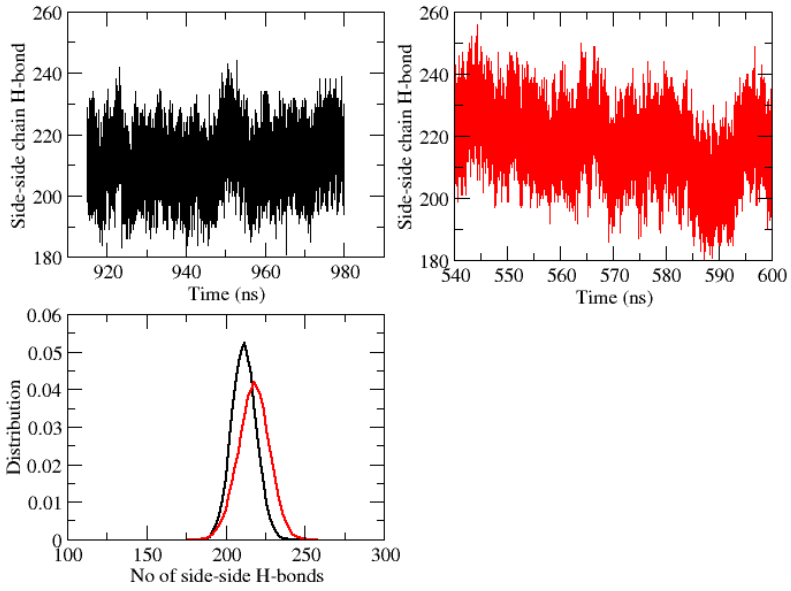


Figure 5.12: The number of side chain – side chain hydrogen bonds with time for (a) bilayer consisting of l-amino acids. (b) Bilayer consisting of alternate d, l-amino acids. (c) Distribution of number of side chain – side chain hydrogen bonds for bilayers consisting of l-amino acid (black line) and alternate d, l amino acid (red line)

The numbers of different types of inter chain hydrogen bonds with simulation time are calculated and shown in Figure 5.13. From the Figure 5.13, it's observed that the bilayer in which glycopeptides block is made of l-amino acid has the higher number of side chain – side chain and side chain - main chain hydrogen bonds. The major contribution comes from the H-

bonding in between side chain NH and main chain CO group for l-amino acid (number of H-bonds ~670), whereas comparatively the number of hydrogen bonds in between side chain OH and main chain CO is found to be much less (~40). The number of hydrogen bonds between glucose hydroxyl groups is found to be highest (~300) among all the possible side chain – side chain hydrogen bonds. In alternate d, l-amino acid bilayer, side chain – side chain hydrogen bonds in between glucose OH groups contributes most towards the total number of H-bonds. But, it's found to be lesser (~260) than what is found in bilayer consisting of l-amino acid. In both bilayers, the number of side chain OH - side chain CO found to be similar (~100). The main difference is observed in the number of side chain NH - main chain CO hydrogen bonds which found to be very high in the bilayer containing l-amino acid compare to alternate d, l-amino acid.

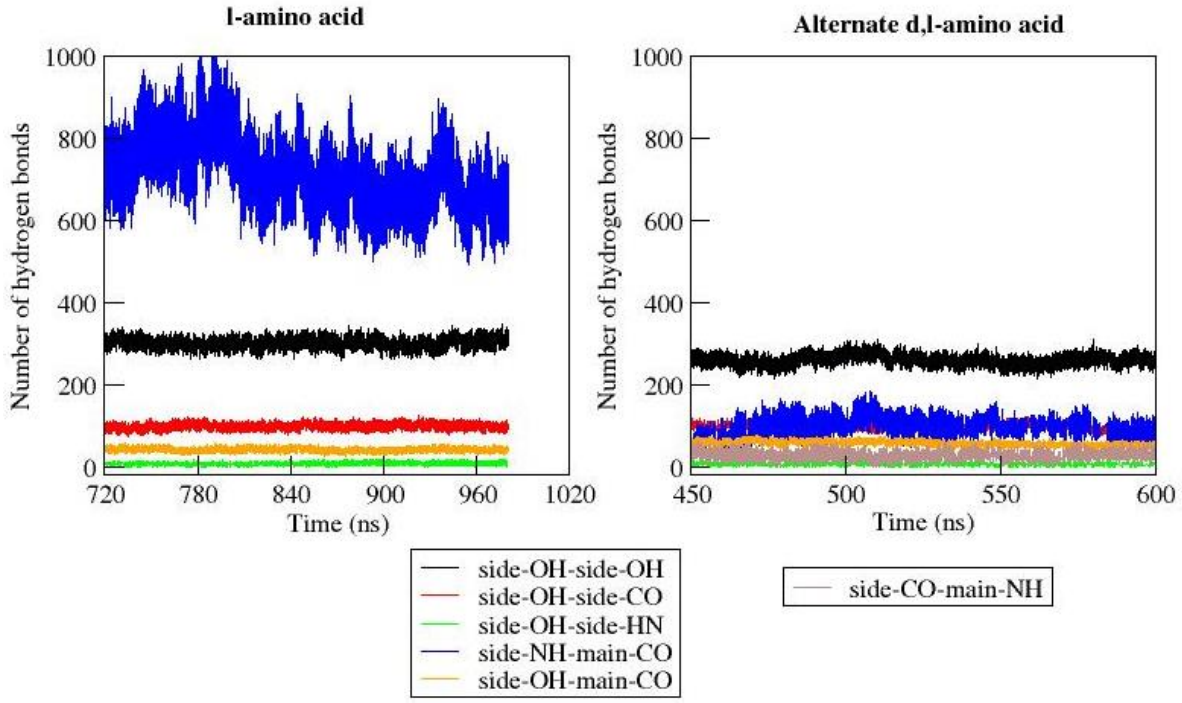


Figure 5.13: The number of different types of inter side chain - side chain and inter side chain - main hydrogen bonds with simulation time for bilayers consisting of l-amino acid (left side) and alternate l and d-amino acid (right side).

5.4 Conclusion

We have done comparative study on the integrity of two bilayers, consisting of diblocks polyglycopeptide-polypropylene oxide chains. The built of hydrophilic glycopeptides blocks

are different, in one bilayer, glycopeptides block comprise of purely l-amino acid, and in another it comprises of alternate l, and d-amino acids. The first bilayer remained integrated after long time simulation. Water molecules penetrate the bilayer consisting of alternate d, l-amino acids residues in large amount which results the separation of diblocks chains from each other. The average area per chain shows that the di-block chains of bilayer consisting of purely l-amino acid residue is more efficiently packed. The di-block chains consisting of l-amino acids are found to be more homogeneously distributed in the bilayer. The percentage α-helicity of the chains consisting of alternate d, l amino is found to be higher than the chain consisting of l-amino acid. We have found that the higher number of side chain – main chain and side chain – side chain hydrogen bonds in the bilayer where polyglycopeptide block consists of l-amino acids keep integrity of the bilayer.

References

[1] Ting, S. R. S., Chen, G., and Stenzel, M. H., "Synthesis of Glycopolymers and Their Multivalent Recognitions with Lectins," Polymer Chemistry 1, 1392-1412 (2010).

[2] Lindhorst, T. K., in *Host-Guest Chemistry: Mimetic Approaches to Study Carbohydrate Recognition*, edited by Penadés, S. (Springer Berlin Heidelberg, Berlin, Heidelberg, 2002), pp. 201-235.

[3] Schatz, C., Louguet, S., Le Meins, J.-F., and Lecommandoux, S., "Polysaccharide-Block-Polypeptide Copolymer Vesicles: Towards Synthetic Viral Capsids," Angewandte Chemie International Edition 48, 2572-2575 (2009).

[4] Aoi, K., Tsutsumiuchi, K., Aoki, E., and Okada, M., "First Synthesis of Glycopeptide Macromonomers and Graft-Type Sugar-Containing Polymers with Glycopeptide Side Chains," Macromolecules 29, 4456-4458 (1996).

[5] Tsutsumiuchi, K., Aoi, K., and Okada, M., "Synthesis of Polyoxazoline−(Glyco)Peptide Block Copolymers by Ring-Opening Polymerization of (Sugar-Substituted) A-Amino Acid N-Carboxyanhydrides with Polyoxazoline Macroinitiators," Macromolecules 30, 4013-4017 (1997).

[6] Upadhyay, K. K., Meins, J. F. L., Misra, A., Voisin, P., Bouchaud, V., Ibarboure, E., Schatz, C., and Lecommandoux, S., "Biomimetic Doxorubicin Loaded Polymersomes from Hyaluronan-Block-Poly(Γ-Benzyl Glutamate) Copolymers," Biomacromolecules 10, 2802-2808 (2009).

[7] Upadhyay, K. K., Bhatt, A. N., Mishra, A. K., Dwarakanath, B. S., Jain, S., Schatz, C., Le Meins, J.-F., Farooque, A., Chandraiah, G., Jain, A. K., Misra, A., and Lecommandoux, S., "The Intracellular Drug Delivery and Anti Tumor Activity of Doxorubicin Loaded Poly(Γ-Benzyl L-Glutamate)-B-Hyaluronan Polymersomes," Biomaterials 31, 2882-2892 (2010).

[8] Pati, D., Kalva, N., Das, S., Kumaraswamy, G., and Sengupta, S., "Multiple Topologies from Glycopolypeptide–Dendron Conjugate Self-Assembly: Nanorods, Micelles, and Organogels," Journal of American Chemical Society 134, 7796-7802 (2012).

[9] Xu, X.-D., Liang, L., Cheng, H., Wang, X.-H., Jiang, F.-G., Zhuo, R.-X., and Zhang, X.-Z., "Construction of Therapeutic Glycopeptide Hydrogel as a New Substitute for Antiproliferative Drugs to Inhibit Postoperative Scarring Formation," Journal of Materials Chemistry 22, 18164-18171 (2012).

[10] Kretsinger, J. K., Haines, L. A., Ozbas, B., Pochan, D. J., and Schneider, J. P., "Cytocompatibility of Self-Assembled B-Hairpin Peptide Hydrogel Surfaces," Biomaterials 26, 5177-5186 (2005).

[11] Lee, K. Y., and Mooney, D. J., "Hyddrogels for Tissue Engineering," Chemical Reviews 101, 1869-1879 (2001).

[12] Mahler, A., Reches, M., Rechter, M., Cohen, S., and Gazit, E., "Rigid, Self-Assembled Hydrogel Composed of a Modified Aromatic Dipeptide," Advanced Materials 18, 1365-1370 (2006).

[13] Kimura, S., Muraji, Y., Sugiyama, J., Fujita, K., and Imanishi, Y., "Spontaneous Vesicle Formation by Helical Glycopeptides in Water," Journal of Colloid and Interface Science 222, 265-267 (2000).

[14] Manning, D. D., Hu, X., Beck, P., and Kiessling, L. L., "Synthesis of Sulfated Neoglycopolymers: Selective P-Selectin Inhibitors," Journal of the American Chemical Society 119, 3161-3162 (1997).

[15] Wang, Y., and Kiick, K. L., "Monodisperse Protein-Based Glycopolymers Via a Combined Biosynthetic and Chemical Approach," Journal of the American Chemical Society 127, 16392-16393 (2005).

[16] Prescher, J. A., and Bertozzi, C. R., "Chemistry in Living Systems," Nat Chem Biol 1, 13-21 (2005).

[17] Roytman, R., Adler-Abramovich, L., Kumar, K. S. A., Kuan, T.-C., Lin, C.-C., Gazit, E., and Brik, A., "Exploring the Self-Assembly of Glycopeptides Using a Diphenylalanine Scaffold," Organic & Biomolecular Chemistry 9, 5755-5761 (2011).

[18] Liu, Y.-F., Sun, Z.-Y., Chen, P.-G., Huang, Z.-H., Gao, Y., Shi, L., Zhao, Y.-F., Chen, Y.-X., and Li, Y.-M., "Glycopeptide Nanoconjugates Based on Multilayer Self-Assembly as an Antitumor Vaccine," Bioconjugate Chemistry 26, 1439-1442 (2015).

[19] Restuccia, A., Tian, Y. F., Collier, J. H., and Hudalla, G. A., "Self-Assembled Glycopeptide Nanofibers as Modulators of Galectin-1 Bioactivity," Cellular and Molecular Bioengineering 8, 471-487 (2015).

[20] Discher, B. M., Won, Y.-Y., Ege, D. S., Lee, J. C.-M., Bates, F. S., Discher, D. E., and Hammer, D. A., "Polymersomes: Tough Vesicles Made from Diblock Copolymers," Science 284, 1143-1146 (1999).

[21] Awasthi, V. D., Garcia, D., Klipper, R., Goins, B. A., and Phillips, W. T., "Neutral and Anionic Liposome-Encapsulated Hemoglobin: Effect of Postinserted Poly(Ethylene

Glycol)-Distearoylphosphatidylethanolamine on Distribution and Circulation Kinetics," Journal of Pharmacology and Experimental Therapeutics 309, 241-248 (2004).

[22] Bermudez, H., Brannan, A. K., Hammer, D. A., Bates, F. S., and Discher, D. E., "Molecular Weight Dependence of Polymersome Membrane Structure, Elasticity, and Stability," Macromolecules 35, 8203-8208 (2002).

[23] Wegenhart, B., Tan, L., Held, M., Kieliszewski, M., and Chen, L., "Aggregate Structure of Hydroxyproline-Rich Glycoprotein (Hrgp) and Hrgp Assisted Dispersion of Carbon Nanotubes," Nanoscale Research Letters 1, 154-159 (2006).

[24] Holowka, E. P., Pochan, D. J., and Deming, T. J., "Charged Polypeptide Vesicles with Controllable Diameter," Journal of American Chemical Society 127, 12423-12428 (2005).

[25] Palian, M. M., Jacobsen, N. E., and Polt, R., "O-Linked Glycopeptides Retain Helicity in Water," The Journal of Peptide Research 58, 180-189 (2001).

[26] Pati, D., Shaikh, A. Y., Das, S., Nareddy, P. K., Swamy, M. J., Hotha, S., and Gupta, S. S., "Controlled Synthesis of O-Glycopolypeptide Polymers and Their Molecular Recognition by Lectins," Biomacromolecules 13, 1287-1295 (2012).

[27] Das, S., Sharma, D. K., Chakrabarty, S., Chowdhury, A., and Sen Gupta, S., "Bioactive Polymersomes Self-Assembled from Amphiphilic Ppo-Glycopolypeptides: Synthesis, Characterization, and Dual-Dye Encapsulation," Langmuir 31, 3402-3412 (2015).

[28] Huang, J., Bonduelle, C., Thévenot, J., Lecommandoux, S., and Heise, A., "Biologically Active Polymersomes from Amphiphilic Glycopeptides," Journal of the American Chemical Society 134, 119-122 (2012).

[29] Zawadzke, L. E., and Berg, J. M., "The Structure of a Centrosymmetric Protein Crystal," Proteins: Structure, Function, and Bioinformatics 16, 301-305 (1993).

[30] Wade, D., Boman, A., Wåhlin, B., Drain, C. M., Andreu, D., Boman, H. G., and Merrifield, R. B., "All-D Amino Acid-Containing Channel-Forming Antibiotic Peptides," Proceedings of the National Academy of Sciences 87, 4761-4765 (1990).

[31] Hess, B., Kutzner, K., Spoel, D. v. d., and Lindahl, E., "Gromacs 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation," Journal of Chemical Theory and Computation 4, 435-447 (2008).

[32] Basma, M., Sundara, S., Çalgan, D., Vernali, T., and Woods, R. J., "Solvated Ensemble Averaging in the Calculation of Partial Atomic Charges," Journal of Computational Chemistry 22, 1125-1137 (2001).

[33] Vanquelef, E., Simon, S., Marquant, G., Garcia, E., Klimerak, G., Delepine, J. C., Cieplak, P., and Dupradeau, F.-Y., "R.E.D. Server: A Web Service for Deriving Resp and

Esp Charges and Building Force Field Libraries for New Molecules and Molecular Fragments," Nucleic Acids Research 39, W511-W517 (2011).

[34] Wang, J., Cieplak, P., Kollman, P. A., and "How Well Does a Restrained Electrostatic Potential (Resp) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules?," Journal of Computational Chemistry 21, 1049-1074 (2000).

[35] Sadeghi, R., and Ziamajidi, F., "Thermodynamic Properties of Aqueous Polypropylene Oxide 400 Solutions from Isopiestic Measurements over a Range of Temperatures," Fluid Phase Equilibria 249, 165-172 (2006).

[36] Hezaveh, S., Samanta, S., Milano, G., and Roccatano, D., "Molecular Dynamics Simulation Study of Solvent Effects on Conformation and Dynamics of Polyethylene Oxide and Polypropylene Oxide Chains in Water and in Common Organic Solvents," The Journal of Chemical Physics 136, 124901 (2012).

[37] Kabsch, W., and Sander, C., "Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features," Biopolymers 22, 2577-2637 (1983).

Chapter 6

Phase Diagram of Self-assembled Sophorolipid Morphologies from Mesoscale Simulations

6.1 Introduction

Three-dimensional microstructures formed by the self-assembly of amphiphilic molecules in aqueous medium have been subjected to scientific investigations by experimentalists as well as theoreticians for a long time. The spontaneous assembly of amphiphilic molecules leads to the formations of many interesting microstructures like vesicles or micelles or membranes depending on molecular structures, interactions between different moieties, and most importantly, the concentration.[1] The wide scope of practical applications of these microstructures in different fields that encompasses from effective drug delivery[2, 3] to modern electronics[4] has enhanced scientific investigations recently in this field. Although a significant amount of study has been done but still many questions remain unanswered like how the mechanism of formation of different microstructures in phase diagram is influenced by various parameters like relative strengths of hydrophobic-hydrophilic interactions, concentration, relative size and shapes of hydrophilic and hydrophobic blocks or moieties in the molecule etc. The detailed description about structure, relative arrangements of amphiphilic chains inside microstructure, the dynamics in confined space etc. is necessary. Experimental techniques have acute limitations to bring such insights out. In such scenario, multiscale molecular dynamics can play an important role to investigate the formation of different microstructures by amphiphilic molecules under different concentrations of water and extract information which can help us to get deeper level of understanding of the phenomena.

Bolaamphiphiles are one such important class of amphiphilic molecules which have two hydrophilic head groups. These two head groups may be same or different in size as well as in hydrophilicity. They are linked together through a hydrophobic alkyl tail. Bolaamphiphile's structure is similar to the Archeobacterias membrane liquid.[5, 6] They exhibit high thermal resistance and are found in the cell membrane of thermophilic bacteria. They can be of two types; the one in which same polar groups at the ends is called symmetric bolaamphiphile, and anti-symmetric bolaamphiphile is one which has different head groups. Also it can be categorised into different classes based on the type of head groups: bolapeptides, bolalipids, and carbohydrate containing bolaamphiphile. Bolaamphiphiles can self-assemble into different structures like vesicle, micelle, hydrogels due to their amphiphilic nature.[7, 8] Their properties of self-assembly can be exploited for various applications in drug delivery, gene delivery[9, 10], bio-imaging[11]. There are some bolaamphiphile prodrug

like PDDZ, ZPPD which have very stable self-assembled structure but they also become active in presence of plasma. They are efficiently taken up by the mononuclear phagocyte system (mps) and effectively reach to the target organs.[12, 13]

The formation of different microstructures of bolaamphiphiles in different aqueous concentrations is essentially a macroscopic phenomenon in terms of length and timescale. The exploration of different morphologies by tuning concentration can provide us an entire spectrum of phase-field description. Then, by extracting the information about the structural properties of the molecules at different regions of microstructure formed at different positions of phase diagram can be correlated in an organised manner to get a coherent and robust perspective. All atomistic molecular dynamics simulations are useful in looking at very local level of interactions in details but to achieve a macroscopic morphology like discussed before requires a time and length scale that demands huge computational power beyond general capacity. However, this limitation can be overcome by applying coarse-grained simulation technique which is capable to sample both length and time scale to such an extent that the microstructures can be achieved at the cost of affordable computational efficiency. At the same time, it is important to maintain a reasonable connection to atomistic details. For our study presented here, it is important to connect these two levels of resolutions by systematic mapping so that perfect coarse-grained interaction parameters and structures can be achieved. We have adopted Dissipative Particle Dynamics (DPD) as the coarse-graining technique to form microstructures of bolaamphiphile at different aqueous concentrations along with all-atomistic molecular dynamics simulation which was used to derive necessary input DPD interaction parameters. This mapping and parameterization are essential initial parts of this study. DPD simulations will provide us necessary insights which, we think, can contribute to molecular level understanding of microstructures of bolaamphiphiles in a qualitative way.

The applications of bola-amphiphiles in different fields have drawn researcher's attention for last few years. Recently, Marnet et al[14] have characterized different micelles formed by glycolipid bolaamphiphile with small angle X-ray scattering, small angle neutron scattering and molecular dynamics simulation. Anionic bolaamphiphile with succinate head group forms unilamellar and multilamellar vesicle structure in water. They are highly stable with respect to the changes in temperature and ionic strength.[15] Lu et al[16] studied the aggregation of different types of surfactant, like gemini surfactant by using dynamic light scattering and transmission electron microscopy. They found that the formation of spherical vesicles facilitated by the hydrocarbon part of polar head group. The controlled release of

proteins in brain by bolaamphiphile cationic vesicles carrier is studied by Dakwar et al[17]. Baccile et al[18] have studied glycolipid based bola containing sophorose as larger head group and –COOH as smaller head group separated by $–C_{15}$ alky tail. They have observed that the morphology of self-assembled structure changes with degree of ionisation of –COOH group. The mixture of carbohydrate based bola-amphiphile and sodium dodecyl surfactant forms micelles in aqueous medium.[19]

There are some simulation studies on self-assembly of polypeptide, lipid, and polymer like materials by atomistic as well as by coarse grain methods. But, the self-assembly of bola-amphiphile by particle based simulation are rare. Mikami et al[20] have first studied sugar-based solid bolaamphiphiles by all atomistic molecular dynamics simulation. They have found that the intermolecular interaction energy between neighbouring molecules prevailed over layer-wise interaction. Crane et al[21] shown rigidity, flexibility, relative molar volume and phase behaviour of model bolaamphiphile crystals from coarse graining simulation. Wang et al[22] obtained the self-assembly of amphiphilic dendritic multiarm copolymers with hydrophobic core and hydrophilic linear arm copolymers into multimolecular spherical micelles by performing DPD simulation. This method is widely used in simulation of polymers[23-28], complex fluids[29-31], liquid with interfaces[32, 33], colloids[34-36]. We have not found any study which has explored the concentration dependent microstructure profile of bolaamphiphiles so far available in literature.

The most important study which motivated our work was done by Dhasaiyan et al.[37] Dhasaiyan et al have experimentally studied the self-assembly of lionlenic acid (LNSL) in water. Their study revealed that the LNSL molecules can have two types of arrangements in self-assembled structure. They have also simulated the bolaamphiphile in different concentration in water and found that it forms spherical aggregate at 1:3 weight ratio of water.[38] They have mainly obtained a spherical structure because of smaller length and time scale of atomistic simulation and they have missed other types of self-assembled morphologies as well depending on the amount of water in the system. But their study lacks the detailed insight about the structural and dynamical properties of constituent chains that govern the inherent microstructures obtained at different concentrations. It is thus important to address structural properties in details to understand the underlying phase diagram. In our work we have used DPD to investigate the morphologies of these self-assembled structures at different concentration of bolaamphiphiles. For performing DPD simulation we have mapped different functional groups of bolaamphiphile into beads by mapping from atomistic

structural details. The interaction parameters between DPD beads were calculated from all atomistic MD simulations, experimental data and finally using Flory Huggins solubility theory. The details of DPD parameters and how we obtained those provided in this work. We have presented the detailed analysis of the arrangement of the bolaamphiphile chains in the self-assembled structure and their dynamics inside confined shell as well. DPD trajectories are used to quantify the flipping of the bolaamphiphile chains in self-assembled structures.

6.2 Computational details

6.2.1 Theoretical Background of DPD

DPD is a mesoscopic simulation method which is first presented by Hoogerbrugge and Koleman[39] to predict complex hydrodynamic behaviour. Later on Espanol and Warren slightly modified it to ensure the correct thermodynamic equilibrium state.[40] Their modification leads to the control of temperature by fluctuation-dissipation theorem. Two bead particles interact through three types of forces; each of them is pair wise additive.

$$f_i = \sum_{j \neq i} (\vec{F}_{ij}^{C} + \vec{F}_{ij}^{D} + \vec{F}_{ij}^{R}) \tag{1}$$

$\vec{F}_{ij}^{C}$ is a conservative force exerted on particle i by particle j, $\vec{F}_{ij}^{D}$ is a dissipative fore, and $\vec{F}_{ij}^{R}$ is random force. The conservative force is repulsive in nature, acting up to cut-off radius which is 1.

$$\vec{F}_{ij}^{C} = \begin{cases} a_{ij}\left(1 - \dfrac{r_{ij}}{r_c}\right)\hat{r}_{ij} & (r_{ij} < 1) \\ 0 & (r_{ij} \geq 1) \end{cases} \tag{2}$$

Where a_{ij} is maximum repulsion between particles i and j; $\vec{r}_{ij} = \vec{r}_i - \vec{r}_j$ and $\hat{r}_{ij} = \vec{r}_{ij} \big/ \left|\vec{r}_{ij}\right|$

The drag force and random force work as heat sink and heat source, respectively. Together they maintain constant temperature of the system, and they are given by,

$$\vec{F}_{ij}^{D} = -\gamma \omega_d(r_{ij})\left(\hat{r}_{ij} \cdot \vec{v}_{ij}\right)\ \hat{r}_{ij} \qquad \vec{F}_{ij}^{R} = \sigma \omega_r(r_{ij})\theta_{ij}\hat{r}_{ij} \tag{3}$$

where, $\vec{v}_{ij} = \vec{v}_i - \vec{v}_j$ is relative velocity and γ is friction coefficient, ω_d and ω_r are r-dependent weight-function and vanishes when r> r_c, θ_{ij} is a random variable with zero mean and variance

of one. The noise amplitude σ is fixed to 3 following the work of Groot and Warren. According to the fluctuation-dissipation theorem for the DPD two weight functions are related to each other by $\omega_D = \omega_R^2$ relation, and two amplitudes are related by $\sigma^2 = 2k_B T.\gamma$

6.2.2 Calculation of interaction parameters between DPD beads

We have performed DPD simulation of bolaapmphiphile-water mixture systems in different molar ratios. For our study it is necessary to adopt systematic mapping scheme to connect atomistic details with coarse-grain details. At first, we mapped structural details based on relative molar volumes of different components, and then, the necessary interaction parameters are calculated by using Flory-Huggins theory. In both cases the atomistic details about structure and atomistic level interactions are important input information for mapping. So the all-atomistic MD simulations have been performed before DPD simulations. Bolaamphiphile has two hydrophilic head groups and one hydrophobic tail group. Sophorose disaccharide head group is in larger size compare to another head group propanoic acid. Tail is C_{15} alkyl chain with three *cis* double bonds in its intermediate position. In DPD simulation the chemical identity of individual atoms is not taken into account. The bolaamphiphile molecule is actually represented by the collection of three different types of beads. Sophorose group is represented by four beads (H), the hydrophobic tail portion consists of five beads (T), and propanoic acid head group is represented by one bead (D). The bead-wise representation of bolaamphiphile is depicted in figure 6.1. Three atomistic water molecules collectively represent one DPD bead (W) for water in our calculations. The interactions among these beads are different depending on their scale of hydrophobicity and hydrophilicity.

The interaction between different beads can be calculated from Flory-Huggins χ parameter and given by the equation[41],

$$\chi^{AB} = (0.231 \pm 0.001)\, \Delta a_{AB} \qquad (4)$$

where $\Delta a_{AB} = a_{AB} - a_{AA.}$ Interaction parameters are same between similar beads and its value is 78.

The χ_{AB} parameter between two species can be obtained from their solubility mismatch, and related to Hildrebrand solubility parameter by

$$\chi^{AB} = V\,(\delta_A - \delta_B)^2/kT \qquad (5)$$

where V is average molar volume of the beads A and B. The solubility parameter is calculated from the square root of the cohesive energy density.

$$\delta = \sqrt{\frac{(\Delta H_v - RT)}{Vm}} \quad (6)$$

The enthalpy of vaporisation of propanoic acid and water is obtained from literature.[42] But, as the enthalpy of vaporization for sophorose and 3, 6, 9-tripentadecene is not available in the literature. We have calculated this using all-atomistic molecular dynamics simulation (MD). We have performed all atomistic MD of sophorose disachharide using charmm36 force field parameters[43] of glucose. The charges of connecting atoms of two glucose units are slightly adjusted to make the molecule neutral. The initial structure for starting all atomistic MD simulation of sophorose is obtained from quantum chemical (QM) optimization with Gaussian 09 software[44] using B3LYP functional and 6-311g (d, p) basis set. For MD simulation 213 sophorose molecules are randomly packed into box size of 6.2×5.7×5.5 nm^3 cubic box using Gromacs 4.6.3 version.[45] Energy minimization was done using steepest descent algorithm. Pure sophorose is a solid compound (melting point 470 K). Therefore, sophorose simulated for 30 ns using NPT-ensemble at three different higher temperatures near to its melting point from 420 K to 440 K with an interval of 10 K. The values of enthalpy and density obtained from the equilibrated systems in these temperatures were interpolated to obtain the enthalpy and density at 298 K (Table 6.1). For NPT ensemble we have used V-rescale thermostat with coupling constant 0.1 ps. Berendsen barostat used to maintain pressure at 1 atm applying 1 ps coupling constant. Electrostatic interaction beyond 1 nm of cut-off was taken care using Particle Mesh Ewald (PME) method. The last 10 ns of the trajectories were analyzed for calculating density and enthalpy.

The initial structure of 3,6,9-tripentadecene was obtained from QM optimization using b3lyp functional and 6-311g (d, p) basis set. Total 314 3,6,9-tripentadecene chains were randomly packed into box of 6.8×7.0×6.84 nm^3 volume. The bulk system was periodically heated and cooled in the temperature range of 298-498 K with an interval of 20 K using simulated annealing (SA) method. After 4 cycles of SA the structure was taken for NPT run for 10 ns. The density is converged to 787 kg/m^3, similar to the alkane of same chain length. The inter-molecular non-bonded interaction energy is calculated by adding columbic and Lennard-Jones interaction, and averaging over last 7 ns of production simulation run. The enthalpy of vaporisation is related to non-bonded energy by thermodynamic law $\Delta H = - E_{nb} + RT$. The

solubility parameters of different components are calculated by Eq. 6 and given in Table 6.2. The interactions parameters between different beads have been calculated Eq. 4 and 5 and given in Table 6.3.

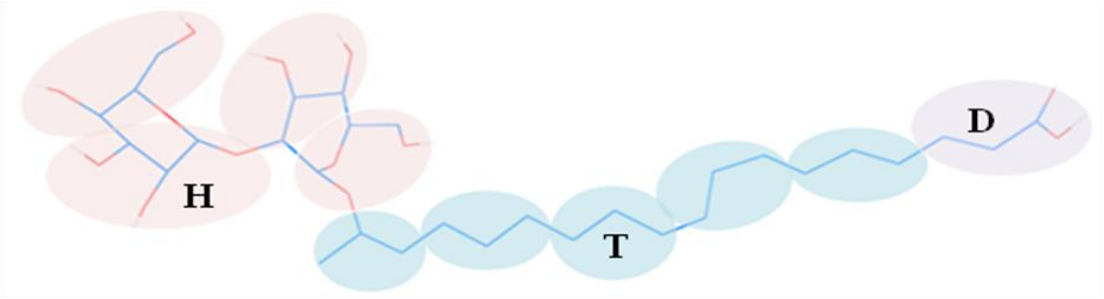


Figure 6.1: Bead-wise representation of bola-amphiphile. Sophorose head group represented by four H beads (pink), alkyl tail by five T beads (blue), and propanoic acid head group by one D bead (violet).

Table 6.1: Enthalpy and densities of sophorose at different temperatures and their interpolated value at 298 K.

Temperature (K)	450	440	430	298
Enthalpy (kJ/mol)	194.97	198.1	201	240.8
Molar volume (cm^3/mol)	250	248	246.43	215.5

Table 6.2: Enthalpy, molar volume, and Hildebrand solubility parameters (δ) of different beads at 298 K.

Beads	H	T	D	W
Enthalpy (kJ/mol)	240.8	76.8	54.4	40.6
Molar volume (cm^3/mol)	215.5	275.6	75	18
δ (J/cm^3)	33.4	16.7	26.93	47.53

Table 6.3: Interaction parameter (a) between different beads:

Interactions	W-H	W-D	W-T	T-H	T-D	H-D
A	97	126	169	105	90	83

All the DPD simulations have been carried out using NVT ensemble and using DL_meso 2.5 package[46]. Temperature parameter is considered as $k_BT = 1$. The mass of each bead is considered to be 1 dpd unit of mass. Timestep for DPD simulation is chosen as 0.06 τ. The DPD units corresponding to physical observables have been converted into their real units. The time in DPD unit is converted into real time by using the equation of $\tau \sim 25.7\ N_m^{5/3}$, where N_m is number of water molecule corresponding to one bead. Since, three water molecules represent one bead; τ becomes 160.4 ps and time step 0.06 τ found to be 9.62 ps. Therefore each time frame in DPD unit is equal to 9.62 ps. The cut-off radius is given by, $r_c = (\rho V_b)^{1/3}$, where density (ρ) of the system is 3 and volume of an water bead is $(3\times30 Å^3) = 90\ Å^3$. Therefore, the value of r_c in physical unit becomes 6.46 Å.

The total number of the DPD beads taken for each simulation is 5×10^4. They are enclosed by the cubic box of length 25.54 DPD unit (or 25.54*0.646 nm =16.5 nm). The systems are simulated for 5×10^5 steps which is equivalent to ~5 μs simulation. Initially, the trajectories are saved after each 1000 steps. After that few simulations have been extended for 5×10^4 steps further for analysis saving the frames at each step. The DPD simulations of bolaamphiphile in water are carried out with five different concentrations. The ratios of bolaamphiphile to water concentrations have been varied from 5 % (v/v) to 50 % (v/v) with 5 % (v/v) interval. The number of taken bolaamphiphile chains, water beads and the formed morphologies are given in Table 6.4. The DPD simulations for each concentration have been repeated three times. We have obtained same morphologies in all three set of simulations.

Table 6.4: Table showing the number of bolaamphiphile molecules, water beads taken for different concentrations of bolaamphiphile in water and their self-assembled morphologies.

Bolaamphiphile concentration (%v/v)	Number of bolaamphiphile molecules	Number of water beads	Morphology
5	250	47500	Spherical
10	500	45000	Spherical
15	750	42500	Spherical
20	1000	40000	Cylindrical
25	1250	37500	Bilayer
30	1500	35000	Distorted cylinder
35	1750	32500	Distorted cylinder
40	2000	30000	Curved bilayer
45	2250	27250	Intermediate Structure
50	2500	25000	Double bilayer

6.2.3 Structure factor Calculation

The type of morphology formed by 30% (v/v) bolaamphiphile in water is verified from structure factors. The structure factors have been evaluated in two ways. In one method, it's calculated considering two different beads, one is sophorose head groups (H) and another is alkyl tail groups (T). The other calculation involves only T beads. S (k) is given by

$$S(k) = \rho_M(k)\rho_M(-k)/N_M$$

Where $\rho_M(k)$ is particle density for M type of bead in the reciprocal space and N_M is particle of number of M beads. ρ_M (k) is given by,

$$\rho_M(k) = \sum_{i=1}^{N} \exp(ikr)$$

$k= (2\pi/L) (k_x, k_y, k_z)$, L is box dimension, k_x, k_y, k_z are the integers.

S (k) is evaluated by averaging the last 50 time frames of equilibration trajectory. The plot of structure factor, S (k) with k^2 is shown in Figure 6.2. The peaks appear at the k^2 =1, 4, 6, 9. Thus the morphology shows peaks at k=1, √4, √6, √9. The peak value best matches with the cylindrical morphology which shows peak at 1, √3, √4, √7, √9.[47, 48] The peak at √3 is missing and another peak appears at √6 instead of √7. Thus, the structure is actually a kind of distorted cylinder, which formed by the bisection of two cylinders in perpendicular direction.

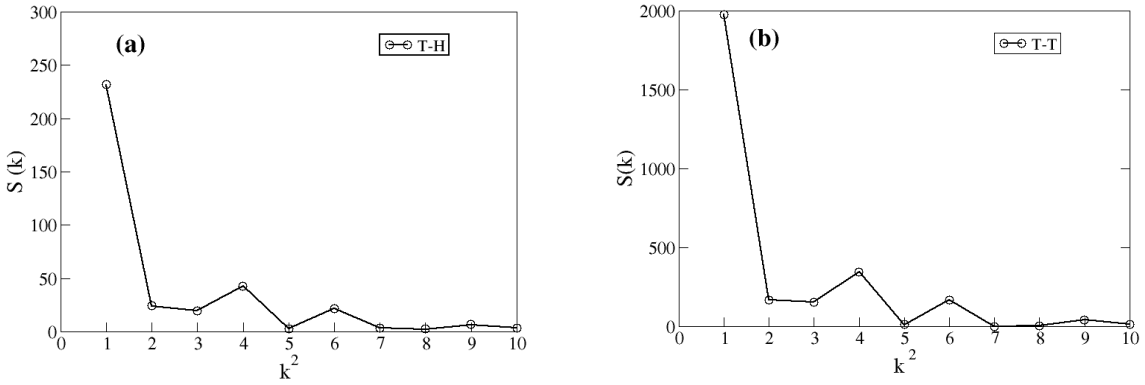


Figure 2: (a) Structure factor plot of the aggregated morphology formed by system with 30 % (v/v) bolaamphiphile in water and calculated (a) in between DPD beads of alkyl tail (T) and sophorose head groups (H), (b) among DPD beads of alkyl tail.

6.3 Results and Discussion

In this study, DPD simulations are performed on randomly dispersed bolaamphiphile chains with different concentration in water. These chains self-assembled into different morphologies depending upon their concentration in water. These morphologies have been examined in details. The arrangements of two different hydrophilic groups and hydrophobic moieties have been explored. The chains in aggregate can take different orientations and they can flip flop from one orientation into another. The residence time of these chains in different orientations and their frequency of flipping have been studied in details.

6.3.1 Morphologies of self-assembled structure

The bolaamphiphile aggregates into different morphologies according to their concentration in water. The concentration dependent self-assembled morphologies of bolaamphiphile molecules in depicted in figure 6.3. Bolaamphiphiles self-assemble into spherical morphology upto its concentration of 15% (v/v) in water. The sophorose head groups are found be on the surface and in the core of the spherical aggregate. The alkyl chains are organized in between surface and core of the sphere. When the concentration of bolaamphiphiles in water is 20 % (v/v), formation of cylindrical morphology has been observed. The sophorose head group gets aggregated at the surface as well as the in core similar to spherical morphology. Bilayer is formed at 25 % (v/v) concentration of bolaamphphiles in water. Sophorose head groups are found at the exterior of bilayer, and alkyl tail and propanoic acid head group remain at the interior of bilayer. Bolaamphiphile with 30% (v/v) and 40% (v/v) concentration in water aggregates into distorted cylindrical structure and curved bilayer structures, respectively. The distorted cylindrical structure has been identified from structure factor. The details of the structure factor calculation and the results are presented in computational details section. The system with 50% (v/v) chain concentration form double bilayer after self-assembly. Sophorose is found be at the exterior of bilayer facing water molecules as well as at the interior of double bilayer. The alkyl tail remains inside of each bilayer. Dhasaiyan et al[37] have observed also similar phenomenon from the atomistic simulation study of randomly dispersed LNSL in water that it forms double bilayer type of morphology and sophorose-sophorose interaction is more predominant than the others. Baccile et al[49] found different morphologies of the self-assembled of sophorolipid in water. The spherical micelle is formed at low concentration of sophorolipid, and it becomes cylindrical upon increasing concentration. Larson et al[1] derived the phase diagram of a tri-component system constituting of water, oil and surfactant through Monte Carlo simulation method. They have shown a change from spherical to cylindrical morphology and finally the lamellar structure by increasing concentration of the surfactant in a surfactant-water mixture. We have also obtained the similar concentration dependent profile of the aggregated morphologies by performing DPD simulation of bolaamphiphiles. All the self-assembled structures formed by bolaamphiphile at different concentration are shown in Figure 6.4. The structures like spherical, cylindrical, bilayer, and double bilayer are further analyzed in the next sections.

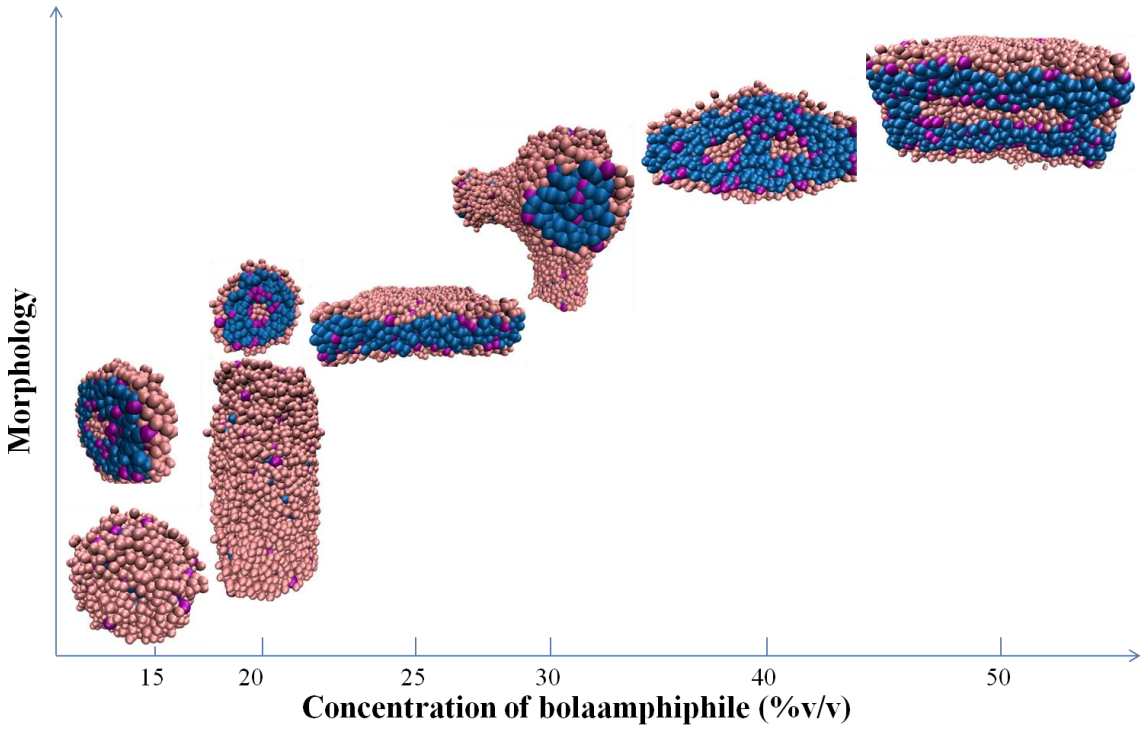
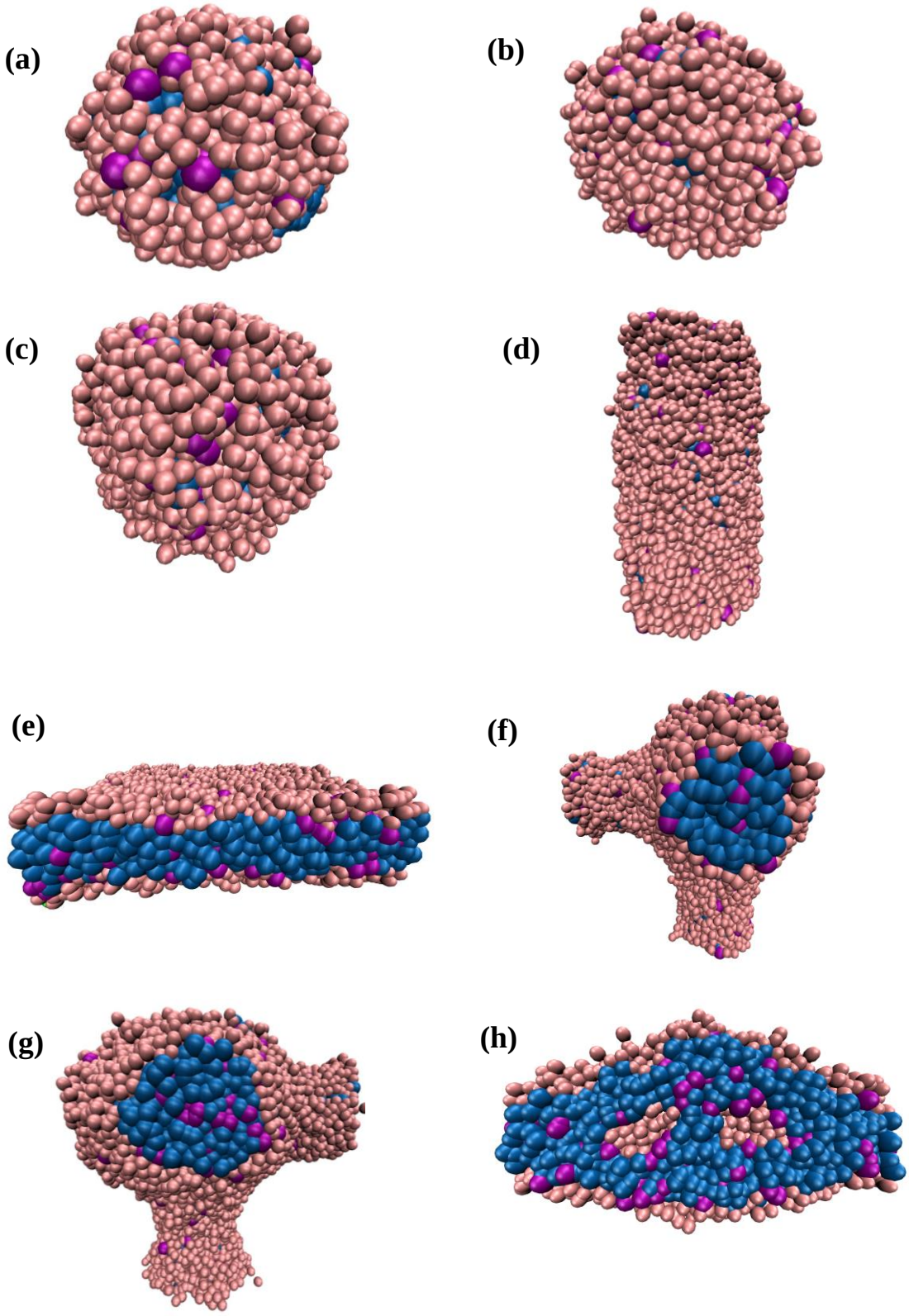


Figure 6.3: (a) The concentration dependent self-assembled morphologies formed by bolaamphiphile molecules in water. Spherical aggregate formed by 10 % (v/v) bolaamphiphile in water. The lower and upper pictures represent periphery and cross-section of spherical morphology, respectively. H, T, and D beads are shown by pink, blue, and purple coloured sphere, respectively. The system with 20 % (v/v) bolaamphiphile concentration in water forms cylindrical aggregate. The lower and upper side pictures represent side view, and top view of cylindrical morphology, respectively. Bolaamphiphile with 25 % (v/v) concentration in water forms bilayer. Distorted cylindrical structure formed by 30 % (v/v) bolaamphiphile in water. The bolaamphiphile with 40 % (v/v) concentration in water aggregated into curved bilayer structure. The double bilayer structure formed by 50 % (v/v) bolaamphiphile in water.


(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)

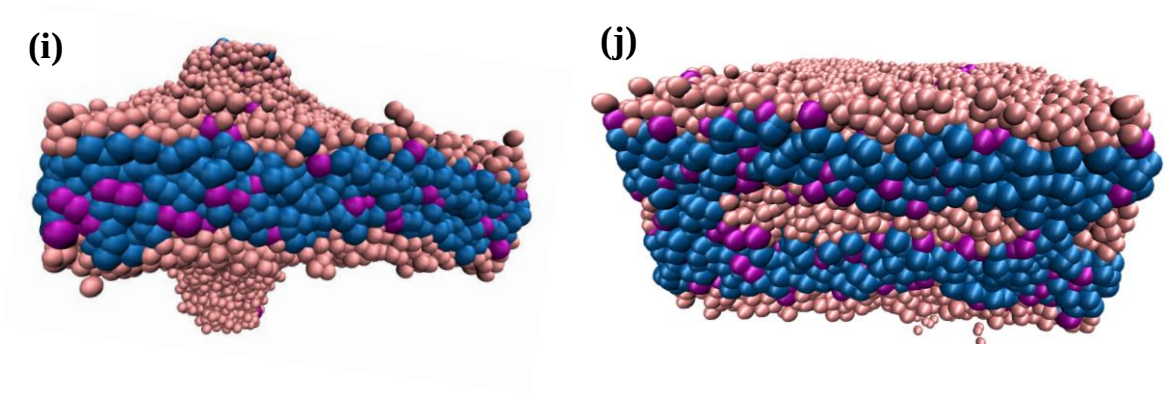


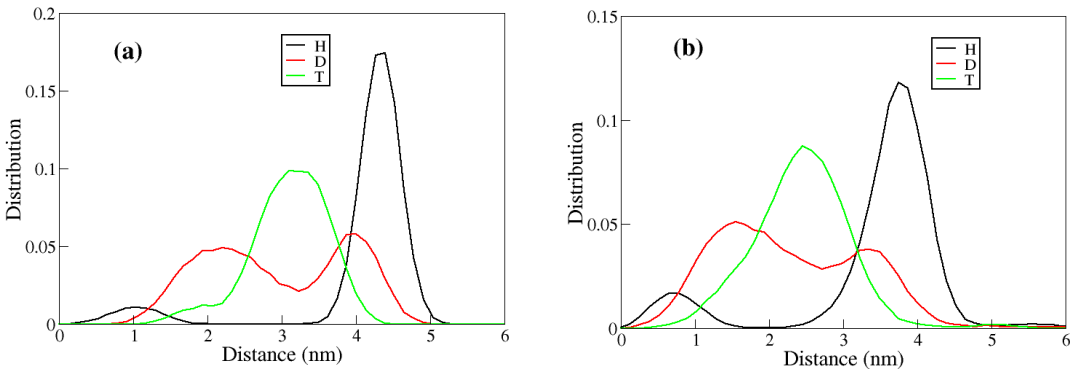
Figure 6.4: The self-assembled morphologies formed by bolaamphiphiles with different concentrations in water. H, T, and D beads are shown by pink, blue, and purple coloured sphere, respectively. (a-c) Spherical aggregates formed by bolaamphiphile with 5% (v/v) to 15 % (v/v) concentrations in water. (d) Bolaamphiphiles with 20 % (v/v) concentration in water form cylindrical morphology. (e) Bilayer formed by 25 % (v/v) bolaamphiphile in water (f-g) Distorted cylindrical structures formed by 35 % (v/v) and 40 % (v/v) bolaamphiphile concentration in water. (h) Bolaamphiphile with 40 % (v/v) concentration in water forms curved bilayer. (i) Intermediate structure in between curved bilayer and double bilayer is formed by 45 % (v/v) bolaamphiphile in water. (j) Bolaamphiphile with 50 % (v/v) concentration in water forms double bilayer.

6.3.2 Arrangement of hydrophobic and hydrophilic groups of sophorolipids

Now, it's very important to envisage how quantitatively these different groups of sophorolipid chains are distributed throughout the self-assembled structure. To calculate the distributions of hydrophobic and two hydrophilic head groups in different morphologies we have deployed different methods described below. For the spherical aggregates formed by 10 % bola present in water, we have first calculated the centre of mass (COM) of whole aggregate consisting of bolaamphiphile chains. Then the COM of 4 beads constituting sophorose head group and COM of five beads constituting C_{15} tail groups have been calculated. The distances between these COMs of different groups from the COM of whole aggregate have been calculated and converted into a histogram (figure 6.5a). The distribution plot clearly shows that the sophorose remains on the periphery of spherical aggregate facing water molecules and also self-assemble at the core. Two peaks appear for the smaller head groups, one in periphery adjacent to the peak position of sophorose and another inside, towards the hydrophobic tails. The tail group remains in between core and periphery. From figure 6.5a it is also evident that the population of smaller head groups are higher near the

core region of the spherical aggregate. Now, inside the aggregate sophorolipid chains can exist either in coiled or elongated form. In coiled form, the smaller head group comes near the larger head group, thus its peak just appear after the peak of larger head group in the distribution plot (figure 6.5a). The elongated chains have its smaller and larger group are apart from one another.

The distribution of the different groups in the cylindrical aggregate formed by the 20% (v/v) bolaamphiphile is calculated in a different way. The axis of cylindrical aggregate is parallel to the x-axis. Thus, the cylindrical structure is divided into different slabs perpendicular the x-axis. The COM of the bolaamphiphile beads residing in each slab is calculated first. Then the x-component of COM of each bola-amphiphile chains is calculated to determine the grid in which COM falls. If it falls into certain slab then the COM of different groups of that chain is calculated. Then the distances between COM of that particular slab in which the chain belongs to and the COM of the different groups of the same chain are calculated. These distances have been converted into histograms (figure 6.5b). The distributions (i.e. the arrangement in the cylindrical aggregate) of the hydrophobic and hydrophilic groups are almost similar to that of the 10% (v/v) bolaamphiphile.



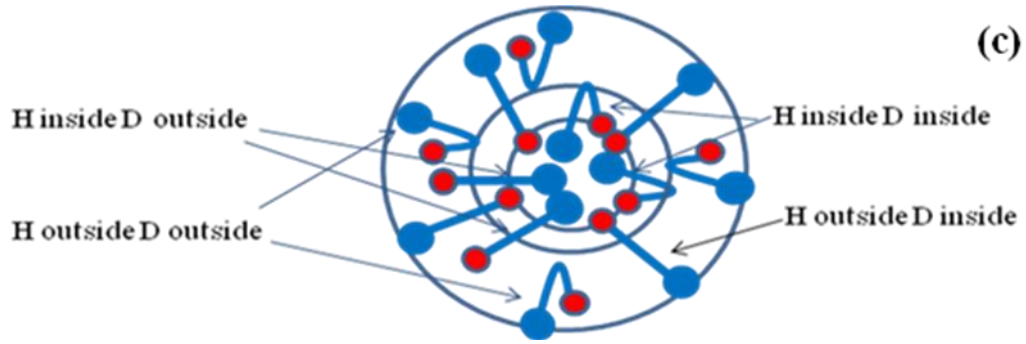


Figure 6.5: (a) Distribution of distance between centres of mass of different beads of chains from COM of spherical aggregate, (b) Distribution of distance between centres of mass of different beads of chains from the axis of cylindrical morphology. (c) Schematic representation of orientation of bolaamphiphile chains in spherical aggregate. Smaller and larger head group are represented by blue and red spheres, respectively. Tail is shown by blue stick. The innermost circle (radius ~2.6 nm) corresponds to the region where no H bead is found. The middle one (radius ~3.2 nm) indicates the peak position of T beads in distribution plots.

The system containing 25% (v/v) bolaamphiphile forms bilayer, whereas the system with 50 % (v/v) bolaamphiphile in water forms double bilayer. Dhaisyan et al also obtained such bilayer like arrangement of bolaamphiphile chains in vesicle.[37] The bilayer-water interface is perpendicular to x-axis. Thus how the density of different groups changes in the bilayer and double bilayer structures is determined by calculation of number densities of beads along x-axis.To calculate the number of beads along the x-axis, the structure is subdivided into 100 slabs parallel to yz-plane. The number of different types of bead is calculated in each slab, and then it's divided by the volume of the slab to get the number density. The number densities of different beads plotted with the position of slab along x-axis are shown in figure 6.5a for bilayer and in figure 6.5b for double bilayer. These plots show that the sophorose remain in exterior of the structure and it faces water for both the structures. The alkyl chain and small head group remained on the inner side. In double bilayer the number of H bead decreases from the exterior to inner side, reaches to minimum and then it again increases in the middle of the structure (figure 6.5b). The whole distribution appears to be divided into two parts along x-axis where the number of each type of beads in the one part is a mirror image of another. Each part of the whole structure is actually resembles a bilayer. The one side (hydrophilic heads) of each bilayer faces water on the exterior and in the interior there is a head-on approach between hydrophilic head groups. The alkyl chain is found to be buried inside each of bilayer. The number density of beads of alky tail and propanoic acid head group along x-axis indicates that the each monolayer constituting bilayer can interact through both of them. Thus, the chains either can be coiled or stretched. The chains are in coiled form when the propanoic acid head group remains near to sophorose and they are extended when propanoic acid head group remains in the interface of two monolayers. Dhaisyan et al[37] have done an experiment and atomistic simulation study of LNSL sophorolipid in water. They found the double bilayer –like of arrangements which we also found from our DPD

simulation with 50 % (v/v) bolaamphiphile in water. They found different arrangements of bolaamphiphile chains among which sophrose-sophrose interaction was more predominant. Pandey et al[38] have performed atomistic simulation of different sophorolipids like linolenic acid, oleic acid in their different concentrations in water. They have shown that the number of hydrogen bonds between sophorose-sophorose head group is more than sophorsoe-propanoic acids and propanoic acid-propanoic acid. We also found from partial density plot and snapshots that bilayers interact with each other mainly through sophorse head groups. They have also obtained a higher number of hydrogen bonds between water and sophorose head group than the rest of bolaapmphile chains. This is also observed in our study as sophorose moiety is found to occupy, predominantly, the exterior part of bilayer facing the water.

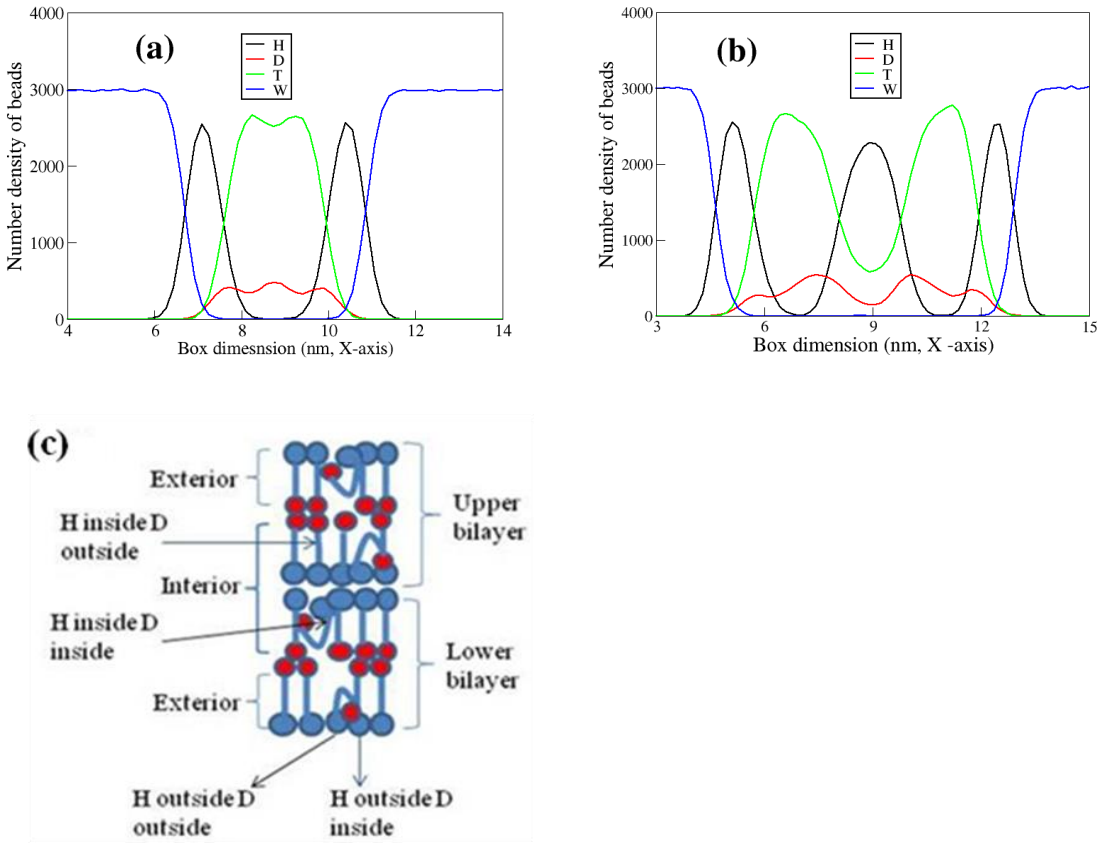


Figure 6.6: (a) Number density of different beads in the bilayer along the normal to bilayer water interface. (b) Number density of different bead along the normal to double-bilayer water interface. (c) Schematic representation of arrangement of chains in double bilayer. Larger and smaller head groups are represented by blue and red spheres, respectively. Tail is shown by blue stick.

6.3.3 Orientation of chains and their flips

In spherical aggregate, the head groups exist in core as well as in periphery and there is certain region where there is no sophorose head group is found. This region exists from 2 to 3.2 nm from the COM of spherical aggregate. The middle of this region is at 2.6 nm which is used to classify the orientation of bola-amphiphile chains based on the position of their sophorose head group. The chains whose sophorose head group are positioned within 2.6 nm from the COM of spherical aggregate is considered to exist in core and if it exists beyond that distance is called then the chains considered to be in exterior region. The peak heights of distribution plot of smaller head group are found to be at 2.3 nm and 4 nm and a peaks minimum exist at 3.2 nm in between them. The chains can again be classified in two orientations depending upon the position of its smaller head groups exists inside, i.e., within peak minima or outside which is beyond peak minima (3.2 nm). Thus the total 4 orientations are possible, the chain which have larger head group (H) inside can have its smaller head group (D) inside or outside. The same is possible for the chains which have H outside. For the cylindrical structure, the head groups are found to be aggregated on periphery and in the core as well. The average number of chains experiencing different orientation is calculated and shown in Table 6.5. The majority of the bolaamphiphile chains are in H outside D outside (~43%) and H outside D inside (~49%) orientations, whereas much less number of chains is found to exist in H inside D outside orientation (~0.7%). The cylindrical aggregate is not completely uniform in the shape and size. Therefore, the distribution plots have been separately calculated for each of the slab as discussed previously. The orientation of the chains in different slabs is schematically represented by figure 6.5c. The minima in between the two peaks of each of the histogram plot for large (H) as well as small (D) head groups are determined. These minima (represented by innermost circle in figure 6.5c) corresponding to histograms of H group is of each grid is used to distinguish the chains in periphery and in the core. The chains in the periphery as well as in core have two possible conformations. The D group is considered to be outside if it is beyond the minimum (represented by middle circle in figure 6.5c) of distribution of D beads from the axis of cylindrical aggregate and inside if it within the value. Thus four orientations of chains are possible in cylindrical aggregates. The chains more prefer to remain in H outside D outside (~39 %), and H outside D inside (~50 %) orientations than the H inside D inside (~9.7 %) and H inside D outside orientations (~1.4%) (Table 6.5).

Two peaks of number density plots of H bead (figure 6.6a) in the bilayer suggest that sophorose faces water on the exterior part of bilayer. The smaller propanoic acid head group and the alky tail groups remain inside the bilayer. The bolaamphiphile chains belong which of the two constituent monolayers are determined from the position of its COM along x-axis with respect to the mid-point (8.7 nm) of the partial density plot (Figure 6.6a) where no sophorose is found. Now, the position of D bead along x-axis differentiates the two orientations of chains. If COM's and D bead's x-component (xyz-coordinate) positioned in same monolayer then the chain is in coiled form, otherwise it is in elongated conformation. The majority (~67 %) of the chains is found to be coiled form. In the double bilayer structure, three peaks are found in the number density plot of H beads (Figure 6.6b). It shows that the sophorose group occupies the exterior regions (2.8 – 7 nm and 11 - 15 nm) as well as interior regions (7 – 11 nm). Therefore, if the first H bead of any chain is found in the exterior, then it is considered to be in the exterior region of double bilayer, otherwise it exists in interior region. The chains of upper bilayer and lower bilayer are distinguished by using a cut-off value of 9 nm passing through middle of the interior part of double bilayer. If the x-component of coordinate of COM of a chain whose H beads lie in this interior region are greater than this cut-off value then it is assumed that the chain belongs to the upper bilayer, otherwise the chain is in lower bilayer. The chains in the exterior and interior regions can have two types of arrangements. In each bilayer, there is a region where the probability of finding H bead is very less and that of T bead very high. The position of D bead within or beyond this point (for upper and lower layer 11 nm, 7 nm respectively) gives rise to two orientations. Therefore, the chains on the exterior as well as interior have two possible orientations which lead into total four orientations. In the double bilayer structure the number of bolaamphiphile chains existing in all three orientations are similar (near 30%) except H inside D outside orientation (~8%). The orientations of these chains in double bilayer are schematically represented in figure 6.6c.

Table 6.5: Table showing percentage of bolaamphiphile chains existing in different orientations for different morphologies.

Morphology	H outside D outside (%)	H outside D inside (%)	H inside D inside (%)	H inside D outside (%)	Coiled (%)	Elongated (%)
Spherical	43	49.3	7	0.7		
Cylindrical	39.2	49.7	9.7	1.4		
Bilayer					67	33
Double bilayer	29.4	28.7	33.6	8.3		

Bolalipids or their mixture with simple monopolar lipids can form membranes which are mechanically stable, show lower permeability while maintaining membrane fluidity. Bulacu et al[50] have atomistically studied flip flop motion of cyclic di-DPPC from coiled to stretched configuration in a self-assembled bilayer. They have found that the flip flop takes place on a 10 ns time scale. Moss et al[51] have studied symmetric bolaamphiphile lipid chain in a bilayer. They have found the flip-flop motion of the chains from coiled to bridging conformation in lipid bilayer. We have found that the chains flip flop in different orientations. We have calculated the number of flips of each chain undergoes from the particular orientation, and the number of frames it remains in the same orientation throughout the last 50000 time steps of equilibrated condensed trajectory. The number of flip is divided by the number of frame in which it remained in the same orientation to obtain the flipping frequency in terms of $step^{-1}$. We have divided the obtained flipping frequency by time step (9.62×10^{-3} ns) to obtain flipping frequency in terms of ns^{-1}. The probability distribution of flipping frequency plots (Figure 6.7) have been normalized by the number of chains experiencing that orientation. The most probable flip frequency values of bolaamphiphile chains in different orientations for different morphologies are shown in Table 6.6. The flip frequencies are given in real time scale (nanosecond) for comparing different morphologies, however these are qualitative numbers, as DPD interaction are parameterized are based on thermodynamic quantities and systems are highly coarse-grained.

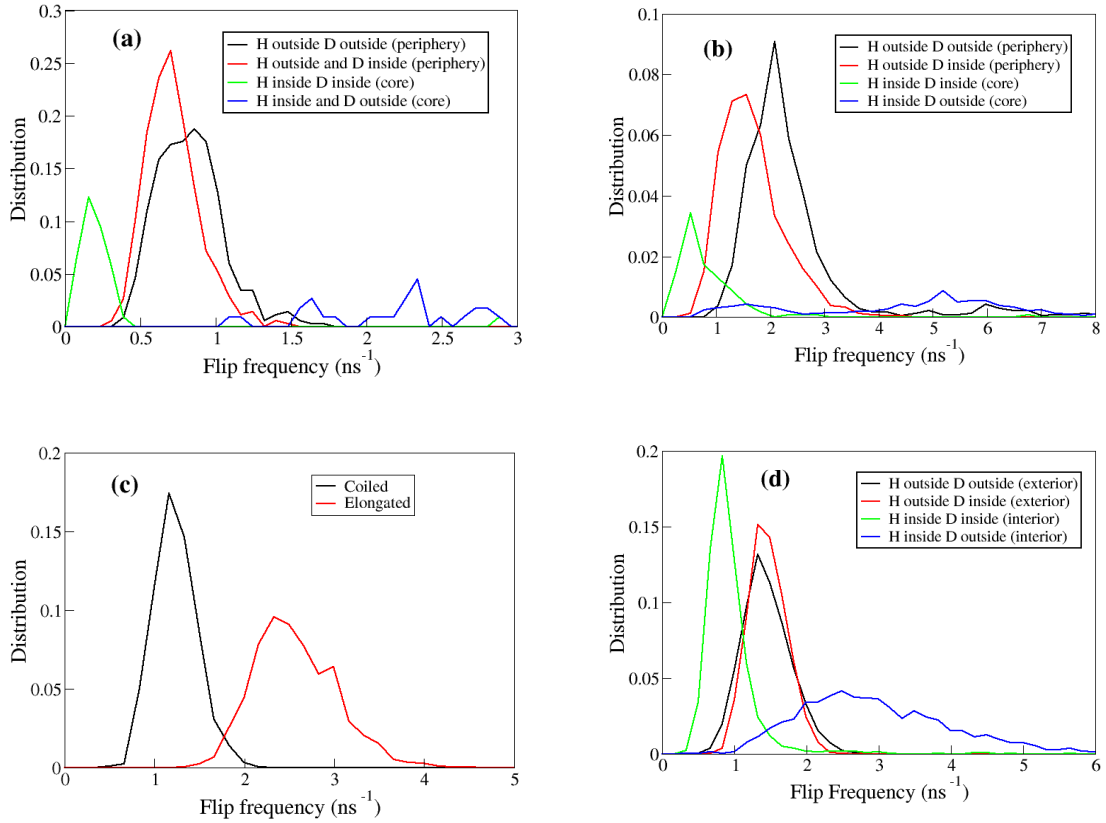


Figure 6.7: Flip frequency distributions of chains for different orientations in (a) spherical aggregate (b) cylindrical aggregate, (c) bilayer structure (d) double bilayer structure.

Table 6.6: Table showing maximum probability of flip frequency (ns^{-1}) of bolaamphiphile chains in different orientations for the different morphologies.

Morphology	H outside D outside	H outside D inside	H inside D inside	H inside D outside	Coiled	Elongated
Spherical	0.8	0.7	0.16	1.6		
Cylindrical	2	1.5	0.5	5		
Bilayer					1.16	2.32
Double bilayer	1.3	1.3	0.85	2.5		

The chains residing in different positions of self-assembled structure have different flip frequency. In spherical aggregate, the chains which exist in core have wide variation of flip

frequency for its two orientations. We found it to be the lowest (peak maxima at 0.16 ns^{-1}) when it's small head group (D) remains inside and the highest (higher than 1 ns^{-1}) when it is outside (Figure 6.7a). The chains which exist in periphery have similar flip frequency for its two orientations. The highest probability of peak distribution found to be at 0.85 ns^{-1}, and 0.7 ns^{-1} for D outside and D inside, respectively. The higher flip frequency of the chains with H inside and D outside orientation indicates that they are less favourable in this orientation and more likely to switch over to the other orientations. The results are similar for the cylindrical aggregate but the flipping frequency of the chains are higher than the same kind of orientation for the chains of spherical aggregate. The maximum probability of the flip frequency found to be at 0.5 ns^{-1}, and 5 ns^{-1} for the chains in the core with D inside and D outside orientations, respectively (Figure 6.7b). The chains in the periphery have flipping frequency in between them, 1.5 ns^{-1} when D inside and 2 ns^{-1} when D outside. It indicates that chains in the cylindrical aggregate are more likely to flip from one state to another than the chains of spherical aggregate. The chains of bilayer shows higher flip frequency (Figure 6.7c) in the elongated form (2.3 ns^{-1}) than the coiled form (1.16 ns^{-1}). In the double bilayer structure, the chains in the interior with H inside and D outside orientation have the lowest flip frequency (0.85 ns^{-1}) where the chains with H inside and D outside have higher flip frequency (2.5 ns^{-1}) (Figure 6.7d). The flip frequencies of chains in periphery are found to be 1.3 ns^{-1} for both the orientations.

The flipping frequency differentiates among chains and they have the tendency to flip-flop from one state to another state residing in different regions. But it can't be discerned that when a chain remains in periphery or in core which state is more preferable between its available two states. The chains in core or inside of the aggregate may prefer to remain in one state more than the other. Therefore, it is important to calculate residence time of the each chain to remain in a particular state. Residence time of chains in particular orientation shows their survival time in that orientation.

The time spent by each chain in a particular orientation is termed here as residence time. We have calculated residence time distribution by converting the number of frames each chain continuously resides in a particular orientation into a histogram. Total 50000 steps of condensed equilibrated trajectory were saved after each time step. These trajectories were used for calculating residence time distribution. The residence time distributions plots for the different orientations are normalized by two factors. The probability of distribution plot increases with the total number of chains undergoing into a particular orientation. Therefore,

the probabilities of the distribution plots of different orientation are divided by the total number of chains which undergoes that orientation. Secondly, if the chains residing in a particular orientation have higher flipping frequency, then probability of residence time distribution will add up. Therefore, they are divided by the highest probable flip frequency number experienced by the bolaamphiphile chains in the corresponding orientations over the 50000 steps of the trajectories. The residence time distributions plots are fitted by single exponential functions and the fitted plots are shown in Figure 6.8.

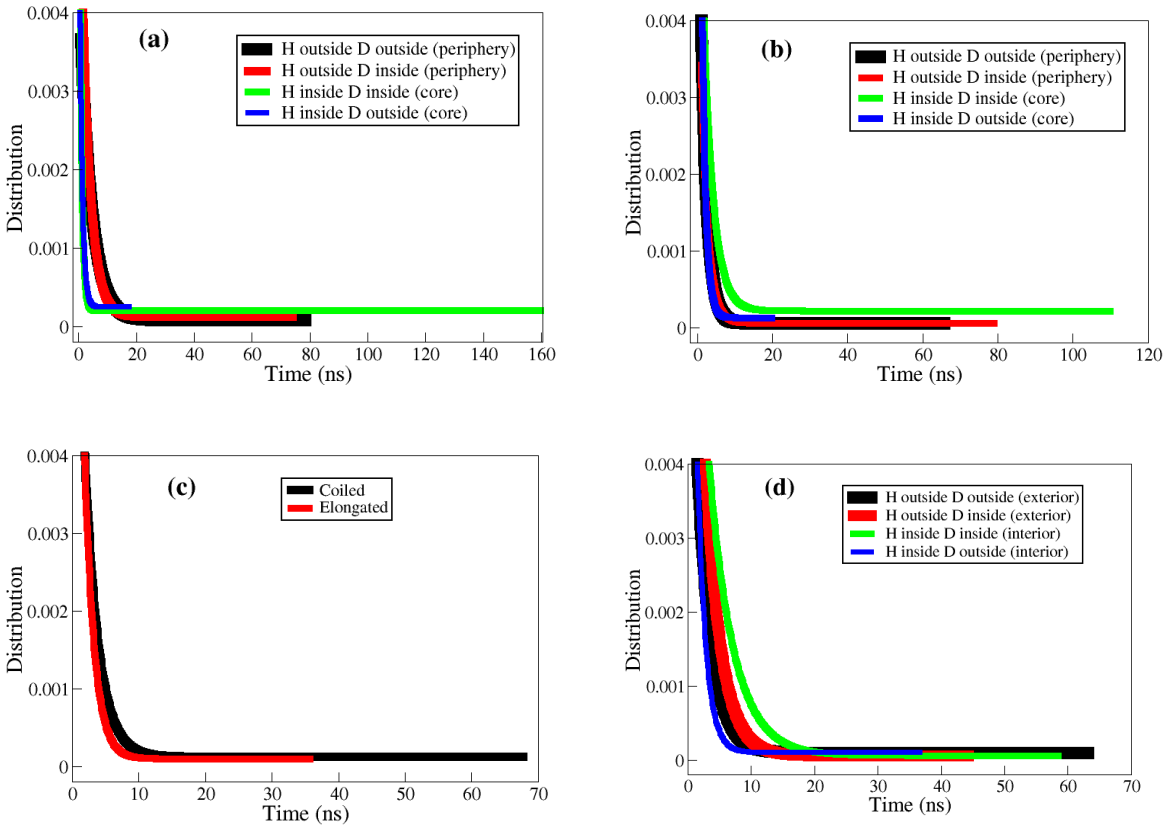


Figure 6.8: Residence time distributions of chains for different orientations in (a) spherical aggregate, (b) cylindrical aggregate, (c) bilayer structure and (d) double bilayer structure.

The residence time distribution in figure 6.8 is mainly shown from its lower probability amplitude region to highlight the tail part or higher residence time region. The maximum survival times of the chains in each possible orientation for different morphologies have been represented in Table 6.7. From figure 6.8a it appears that the residence time distributions of both the orientations of the bolaamphiphile chains existing at the periphery are similar and

chains reside in that orientation up to approximately 80 ns. Residence time of H inside and D inside orientation is higher than the H inside D outside orientation for the bolaamphiphile chains which exist in the core. The survival time of H inside and D inside orientation is found to be 160 ns, whereas the chains with H inside and D outside orientation can survive at most 18 ns. Therefore in the spherical aggregate, the chains which exist in periphery, both the orientations are equally favourable. However, orientations with both H and D inside are more favourable than the other for the chains existing in the core. In the cylindrical aggregate the chains at the periphery with H outside and D inside orientation have slightly higher residence time (79 ns) than the H outside and D inside orientation (65 ns). But, the residence time of the chains in a cylindrical aggregate having H inside and D inside orientation is lower than the chains of spherical aggregate for the same orientation (figure 6.8b). Similar to spherical aggregate, the chains in the core with H inside and D inside orientation have higher residence time (up to ~110 ns) than the other orientation (~20 ns). From figure 6.8c, it's found that the chains in bilayer morphology have two orientations in which coiled conformation tend to show higher residence time (~68 ns) than the elongated chains (~36 ns). In the double bilayer, the number of chains with H inside D outside orientation survives longer time (upto ~35 ns) than the same orientations of other morphologies (figure 6.8d). In all the morphologies, the chains with H inside D outside have lowest residence times among all the orientations. The chains which exist in the exterior facing water it is found that in the spherical aggregate it can survive for higher time (~80 ns) than the others self-assembled morphologies. This is also comparable to the observed lowest flip frequency (~0.8 ns^{-1}) in spherical aggregate among others. The chains which exist in interior with H inside and D inside orientation show higher residence time than all the other orientation for spherical and cylindrical morphologies. But, it is found that it can survive the longer time (~ upto 160 ns) in the spherical morphology than in cylindrical morphology (~110 ns) and double bilayer (~59 ns). The same observation is found form the comparison of flip frequencies of this orientation in all the morphologies. It's found to be the lowest in spherical morphology (0.16 ns^{-1}) compared to the cylindrical (0.5 ns^{-1}) and double bilayer (0.85 ns^{-1}) morphologies.

Table 6.7: Table showing maximum survival times (ns) of bolaamphiphile chains in different orientations for different morphologies.

Morphology	H outside D outside	H outside D inside	H inside D inside	H inside D outside	Coiled	Elongated
Spherical	79	75	160	18		
Cylindrical	65	79	110	20		
Bilayer					68	36
Double bilayer	63	44	59	37		

6.4 Conclusions

The self-assembly of linolenic acid bolaamphiphile is investigated in their different concentrations in water by DPD simulation. We have found a complete phase diagram of different self-assembled morphologies for their different concentration in water. The arrangements of different hydrophilic and hydrophobic portion of bolaamphiphile chains in different kinds of aggregates have been studied. The hydrophilic sophorose self-assembles into the exterior facing water and they are also found to be assembled in core for different of morphologies. The hydrophobic tail buried inside in between periphery and core. We have observed from the distribution plots of the beads that the chains remain in different orientations and position of self-assembled aggregate. They can flip from one orientation to another. The flipping frequencies of these bolaamphiphile chains in different orientations and various self-assembled morphologies have been compared. The flipping frequency is found be the lowest for H inside-D-inside orientation, and highest for H inside–D outside orientation among the four possible orientations for all types of morphologies. The chains existing in H outside - D outside and H outside D inside orientations show similar flip frequencies. The flipping frequencies for all the orientations are found to be less in spherical morphology in comparison to others. The residence time has been calculated for the chains experiencing different orientation for all the morphologies. The results obtained from residence time distribution plots are in agreement with frequency distribution. These detailed

information about chain conformations and their dynamics extracted from the microstructures impart crucial understanding about the morphologies at different concentrations. The molecular level insight gained is expected to contribute significantly in future investigations in this field of study.

References

[1] Larson, R. G., "Self-Assembly of Surfactant Liquid Crystalline Phases by Monte Carlo Simulation," The Journal of Chemical Physics 91, 2479-2488 (1989).

[2] Grinberg, S., Kolot, V., Linder, C., Shaubi, E., Kas'yanov, V., Deckelbaum, R. J., and Heldman, E., "Synthesis of Novel Cationic Bolaamphiphiles from Vernonia Oil and Their Aggregated Structures," Chemistry and Physics of Lipids 153, 85-97 (2008).

[3] Popov, M., Linder, C., Deckelbaum, R. J., Grinberg, S., Hansen, I. H., Shaubi, E., Waner, T., and Heldman, E., "Cationic Vesicles from Novel Bolaamphiphilic Compounds," Journal of Liposome Research 20, 147-159 (2010).

[4] Györvary, E., Schroedter, A., Talapin, D. V., Weller, H., Pum, D., and Sleytr, U. B., "Formation of Nanoparticle Arrays on S-Layer Protein Lattices," Journal of Nanoscience and Nanotechnology 4, 115-120 (2004).

[5] Fuhrhop, J.-H., and Mathieu, J., "Routes to Functional Vesicle Membranes without Proteins," Angewandte Chemie International Edition in English 23, 100-113 (1984).

[6] Arakawa, K., Eguchi, T., and Kakinuma, K., "36-Membered Macrocyclic Diether Lipid Is Advantageous for Archaea to Thrive under the Extreme Thermal Environments," Bulletin of the Chemical Society of Japan 74, 347-356 (2001).

[7] da Silva, E. R., Alves, W. A., Castelletto, V., Reza, M., Ruokolainen, J., Hussain, R., and Hamley, I. W., "Self-Assembly Pathway of Peptide Nanotubes Formed by a Glutamatic Acid-Based Bolaamphiphile," Chemical Communications 51, 11634-11637 (2015).

[8] Meister, A., Drescher, S., Garamus, V. M., Karlsson, G., Graf, G., Dobner, B., and Blume, A., "Temperature-Dependent Self-Assembly and Mixing Behavior of Symmetrical Single-Chain Bolaamphiphiles," Langmuir 24, 6238-6246 (2008).

[9] Jain, N., Arntz, Y., Goldschmidt, V., Duportail, G., Mély, Y., and Klymchenko, A. S., "New Unsymmetrical Bolaamphiphiles: Synthesis, Assembly with DNA, and Application for Gene Delivery," Bioconjugate Chemistry 21, 2110-2118 (2010).

[10] Jain, N., Goldschmidt, V., Oncul, S., Arntz, Y., Duportail, G., Mély, Y., and Klymchenko, A. S., "Lactose-Ornithine Bolaamphiphiles for Efficient Gene Delivery in Vitro," International Journal of Pharmaceutics 423, 392-400 (2012).

[11] Shi, E., Gao, Z., Yuan, M., Wang, X., and Wang, F., "Self-Assembly of Benzothiadiazole-Functionalized Dinuclear Platinum Acetylide Bolaamphiphiles for Bio-Imaging Application," Polymer Chemistry 6, 5575-5579 (2015).

[12] Jin, Y., Qi, N., Tong, L., and Chen, D., "Self-Assembled Drug Delivery Systems. Part 5: Self-Assemblies of a Bolaamphiphilic Prodrug Containing Dual Zidovudine," International Journal of Pharmaceutics 386, 268-274 (2010).

[13] Jin, Y., Xin, R., Tong, L., Du, L., and Li, M., "Combination Anti-Hiv Therapy with the Self-Assemblies of an Asymmetric Bolaamphiphilic Zidovudine/Didanosine Prodrug," Molecular Pharmaceutics 8, 867-876 (2011).

[14] Manet, S., Cuvier, A.-S., Valotteau, C., Fadda, G. C., Perez, J., Karakas, E., Abel, S., and Baccile, N., "Structure of Bolaamphiphile Sophorolipid Micelles Characterized with Saxs, Sans, and Md Simulations," The Journal of Physical Chemistry B 119, 13113-13133 (2015).

[15] Lo Nostro, P., and Gabrielli, G., "Osmotic and Structural Studies on Bolaamphiphile Vesicles," Colloids and Surfaces 44, 119-137 (1990).

[16] Lu, T., Han, F., Mao, G., Lin, G., Huang, J., Huang, X., Wang, Y., and Fu, H., "Effect of Hydrocarbon Parts of the Polar Headgroup on Surfactant Aggregates in Gemini and Bola Surfactant Solutions," Langmuir 23, 2932-2936 (2007).

[17] Dakwar, G. R., Hammad, I. A., Popov, M., Linder, C., Grinberg, S., Heldman, E., and Stepensky, D., "Delivery of Proteins to the Brain by Bolaamphiphilic Nano-Sized Vesicles," Journal of Controlled Release 160, 315-321 (2012).

[18] Baccile, N., Babonneau, F., Jestin, J., Pehau-Arnaudet, G., and Van Bogaert, I., "Unusual, Ph-Induced, Self-Assembly of Sophorolipid Biosurfactants," ACS Nano 6, 4763-4776 (2012).

[19] Gerber, S., Garamus, V. M., Milkereit, G., and Vill, V., "Mixed Micelles Formed by Sds and a Bolaamphiphile with Carbohydrate Headgroups," Langmuir 21, 6707-6711 (2005).

[20] Mikami, M., Matsuzaki, T., Masuda, M., Shimizu, T., and Tanabe, K., "Molecular Dynamics Simulation for the Crystal Structure of Synthetic Sugar-Based Bolaamphiphiles," Computational Materials Science 14, 267-276 (1999).

[21] Crane, A. J., Martinez-Veracoechea, F. J., Escobedo, F. A., and Muller, E. A., "Molecular Dynamics Simulation of the Mesophase Behaviour of a Model Bolaamphiphilic Liquid Crystal with a Lateral Flexible Chain," Soft Matter 4, 1820-1829 (2008).

[22] Wang, Y., Li, B., Zhou, Y., Lu, Z., and Yan, D., "Dissipative Particle Dynamics Simulation Study on the Mechanisms of Self-Assembly of Large Multimolecular Micelles from Amphiphilic Dendritic Multiarm Copolymers," Soft Matter 9, 3293-3304 (2013).

[23] Groot, R. D., and Madden, T. J., "Dynamic Simulation of Diblock Copolymer Microphase Separation," The Journal of Chemical Physics 108, 8713-8724 (1998).

[24] Kong, Y., Manke, C. W., Madden, W. G., and Schlijper, A. G., "Effect of Solvent Quality on the Conformation and Relaxation of Polymers Via Dissipative Particle Dynamics," The Journal of Chemical Physics 107, 592-602 (1997).

[25] Zhang, K., and Manke, C. W., "Simulation of Polymer Solutions by Dissipative Particle Dynamics," Molecular Simulation 25, 157-166 (2000).

[26] Guerrault, X., Rousseau, B., and Farago, J., "Dissipative Particle Dynamics Simulations of Polymer Melts. I. Building Potential of Mean Force for Polyethylene and Cis-Polybutadiene," The Journal of Chemical Physics 121, 6538-6546 (2004).

[27] Chakraborty, S., and Roy, S., "Structure of Nanorod Assembly in the Gyroid Phase of Diblock Copolymer," The Journal of Physical Chemistry B 119, 6803-6812 (2015).

[28] Roy, S., Markova, D., Kumar, A., Klapper, M., and Müller-Plathe, F., "Morphology of Phosphonic Acid-Functionalized Block Copolymers Studied by Dissipative Particle Dynamics," Macromolecules 42, 841-848 (2009).

[29] Xu, Z., and Meakin, P., "A Phase-Field Approach to No-Slip Boundary Conditions in Dissipative Particle Dynamics and Other Particle Models for Fluid Flow in Geometrically Complex Confined Systems," The Journal of Chemical Physics 130, 234103 (2009).

[30] Yamaguchi, T., Ishikawa, T., Imai, Y., Matsuki, N., Xenos, M., Deng, Y., and Bluestein, D., "Particle-Based Methods for Multiscale Modeling of Blood Flow in the Circulation and in Devices: Challenges and Future Directions," Annals of Biomedical Engineering 38, 1225-1235 (2010).

[31] Keaveny, E. E., Pivkin, I. V., Maxey, M., and Em Karniadakis, G., "A Comparative Study between Dissipative Particle Dynamics and Molecular Dynamics for Simple- and Complex-Geometry Flows," The Journal of Chemical Physics 123, 104107 (2005).

[32] Dong, F.-L., Li, Y., and Zhang, P., "Mesoscopic Simulation Study on the Orientation of Surfactants Adsorbed at the Liquid/Liquid Interface," Chemical Physics Letters 399, 215-219 (2004).

[33] Pan, D., Phan-Thien, N., and Khoo, B. C., "Studies on Liquid–Liquid Interfacial Tension with Standard Dissipative Particle Dynamics Method," Molecular Simulation 41, 1166-1176 (2015).

[34] Whittle, M., and Dickinson, E., "On Simulating Colloids by Dissipative Particle Dynamics: Issues and Complications," Journal of Colloid and Interface Science 242, 106-109 (2001).

[35] Dzwinel, W., Yuen, A. D., and Boryczko, K., "Mesoscopic Dynamics of Colloids Simulated with Dissipative Particle Dynamics and Fluid Particle Model," Molecular modeling annual 8, 33-43 (2002).

[36] Whittle, M., and Travis, K. P., "Dynamic Simulations of Colloids by Core-Modified Dissipative Particle Dynamics," The Journal of Chemical Physics 132, 124906 (2010).

[37] Dhasaiyan, P., Pandey, P. R., Visaveliya, N., Roy, S., and Prasad, B. L. V., "Vesicle Structures from Bolaamphiphilic Biosurfactants: Experimental and Molecular Dynamics Simulation Studies on the Effect of Unsaturation on Sophorolipid Self-Assemblies," Chemistry – A European Journal 20, 6246-6250 (2014).

[38] Pandey Prithvi, R., Dhasaiyan, P., Prasad, B. L. V., and Roy, S., "Structural Insight into Self Assembly of Sophorolipids: A Molecular Dynamics Simulation Study," Zeitschrift für Physikalische Chemie 230, 819-836 (2016).

[39] Hoogerbrugge, P. J., and Koelman, J. M. V. A., "Simulating Microscopic Hydrodynamic Phenomena with Dissipative Particle Dynamics," EPL (Europhysics Letters) 19, 155-160 (1992).

[40] Español, P., and Warren, P., "Statistical Mechanics of Dissipative Particle Dynamics," EPL (Europhysics Letters) 30, 191-196 (1995).

[41] Groot, R. D., and Rabone, K. L., "Mesoscopic Simulation of Cell Membrane Damage, Morphology Change and Rupture by Nonionic Surfactants," Biophysical Journal 81, 725-736 (2001).

[42] Verevkin, S. P., "Measurement and Prediction of the Monocarboxylic Acids Thermochemical Properties," Journal of Chemical & Engineering Data 45, 953-960 (2000).

[43] Guvench, O., Greene, S. N., Kamath, G., Brady, J. W., Venable, R. M., Pastor, R. W., and Mackerell, A. D., "Additive Empirical Force Field for Hexopyranose Monosaccharides," Journal of Computational Chemistry 29, 2543-2564 (2008).

[44] Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.;

Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J., Gaussian, Inc., Wallingford CT (2009).

[45] Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E., "Gromacs 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation," Journal of Chemical Theory and Computation 4, 435-447 (2008).

[46] Seaton, M. A., Anderson, R. L., Metz, S., and Smith, W., "Dl_Meso: Highly Scalable Mesoscale Simulations," Molecular Simulation 39, 796-821 (2013).

[47] Hajduk, D. A., Harper, P. E., Gruner, S. M., Honeker, C. C., Kim, G., Thomas, E. L., and Fetters, L. J., "The Gyroid: A New Equilibrium Morphology in Weakly Segregated Diblock Copolymers," Macromolecules 27, 4063-4075 (1994).

[48] Lodge, T. P., Pudil, B., and Hanley, K. J., "The Full Phase Behavior for Block Copolymers in Solvents of Varying Selectivity," Macromolecules 35, 4707-4717 (2002).

[49] Baccile, N., Nassif, N., Malfatti, L., Van Bogaert, I. N. A., Soetaert, W., Pehau-Arnaudet, G., and Babonneau, F., "Sophorolipids: A Yeast-Derived Glycolipid as Greener Structure Directing Agents for Self-Assembled Nanomaterials," Green Chemistry 12, 1564-1567 (2010).

[50] Bulacu, M., Périole, X., and Marrink, S. J., "In Silico Design of Robust Bolalipid Membranes," Biomacromolecules 13, 196-205 (2012).

[51] Moss, R. A., Fujita, T., and Okumura, Y., "Dynamics of a Bolaamphiphilic Lipid in a Bilayer Liposome," Langmuir 7, 2415-2418 (1991).

Chapter 7

Conclusions

In this thesis, we have envisaged important factors controlling the self-assembly of polypeptide based molecules and bolaamphiphile from atomistic and mesoscale simulation, respectively. The self-assembly of monolayer consisting of α-helical diblocks co-polypeptide is mainly understood by performing all-atomistic MD simulation. Bilayer consisting of diblock chains of hydrophilic α-helical poly-glycopeptides and hydrophobic poly-propylene oxide have also been studied in details with the help of MD simulation. In this study our aim was to understand the effect of α-helicity of poly-glycopeptide segment and effective packing of the chains on the integrity of bilayer. Thus, glycopeptide segment is constructed by two types of constituent amino acids conformers; one segment is made of l-amino acid, and another is built by alternate arrangement of d, l amino acid residues. It requires search of a suitable force field among the available ones. The self-assembly of monolayer and bilayer systems have one common feature which is both of them are comprise of di-block chains of different hydrophobilcity. Another common important aspect is that lamellar phase has been considered where the water concentration in the system is approximately 50 % by weight. In the self-assembly simulation of bola-amphiphile, we have further varied our building block from bi-component to tri-component material. We have performed dissipative particle dynamics simulation to study the self-assembly of bola-amphiphile in water at different concentrations.

Self-assembly of block copolymers is a subject which has been well studied. Block copolymers consisting of hydrophobic and hydrophilic segment self-assemble to form different structures like lamella, rods, micelle, etc. But, the structural insight of these aggregates is lacking and also difficult to obtain from experiments. Thus to obtain structural insight and quantification of interaction, MD simulation has been performed with a monolayer consisting of α-helical di-block of capped leucine-lysine. The choice of these particular amino acid comes from the fact that the length of their side chain are similar but presence of an amine group in lysine side chain makes it more hydrophilic than leucine. The preassembled monolayer is considered as initial structure for doing MD simulation as there is no specific interaction site between hydrophobic and hydrophilic sites as in lipid which makes the self-assembly process slower. Moreover, the long time MD simulation does not ensure the formation of lamellar phase as the system may stuck into local energy minima. The lamellar phase, formed by di-block is main interest of our study and thus the concentration of water is maintained approximately 50 % by weight. We have tried to investigate what happens during self-assembly of α-helical di-block copolypetide, whether

hydrophobic and hydrophilic blocks comes near to themselves or there is possibility of hydrophobic block to come within aggregated region of hydrophilic block or vice versa. This question is answered by performing MD simulation with two types of preassembled monolayer systems. One is called parallel monolayer where leucine and lysine blocks are near themselves, and another is anti-parallel where leucine block comes near to lysine block. Both the monolayers are found to be intact and retain their lamellar structures after three steps simulation which includes two NPT simulations with an intermediate simulated annealing method. Water penetrated both the monolayers but it enters more from lysine side region of parallel monolayer. We have envisaged hydrogen bonded interaction to find the reason of the stability of these self-assembled structures. Water forms hydrogen bonds with side chain and main chain of amino acids residues. The formation of water mediated inter-chain hydrogen bonds in large number may play a role stabilising both the monolayers. In both these monolayers, the side chain $–NH_2$ groups on the lysine blocks are capable of forming inter helical H-bonds among themselves, and with the main chain CO groups due to interpenetration of the side chains. In parallel alignment the side-chain–side-chain H-bonds is found more than side-chain–main H-bonds. In case of anti-parallel alignment the number of side-chain – main-chain H-bonds as well as side-chain – side-chain H-bonds are less than that of parallel alignment. The directional preference is observed in H-bond distribution in monolayer. In parallel alignment, side chain - side chain and side chain - main chain hydrogen bonds are formed more in parallel and perpendicular directions than diagonal, whereas in anti-parallel alignment side chain - side chain hydrogen bonds in diagonal direction is prevailing. The interaction among peptide chains of parallel and anti-parallel alignment have been compared by calculation of free energy of binding. The free energies of binding at different angle of orientations between two diblocks chains show that both parallel and anti-parallel alignment are more preferable than other orientations. But parallel orientation found to be slightly more favourable than anti-parallel alignment. Therefore, both parallel and anti-parallel alignment is probable during self-assembly, but parallel alignment is somewhat more preferable. Thus all-atomistic MD simulation gives detailed insight of self-assembled structure consisting of di-block co-polypeptides. Although from the atomistic simulation it’s difficult to obtain self-assembled lamellar structure, but there is scope of obtaining this from coarse-grain simulations. This study can be further extended for bilayer assembly consisting of di-block copolypeptide chains.

The self assembly of right and left handed α-helices made up of L- or D-isomers of amino acids form wide range of structures like nanorods, micelles, etc. Structural arrangements of different helices are not well understood in the self assembled structures. Molecular dynamics (MD) becomes very useful tool in this regard. But before performing MD simulations it is necessary to understand that which Force-field is robust enough to give good results for right and left handed helices made up of L-or D- isomer. QM calculations have been used as benchmark to test the ability of different force fields to reproduce the structural features of different helical conformations constructed of enantiomeric amino acids. All the α-helical conformations are energy optimized by quantum mechanical (QM) method at M06/6-311g (d, p) level of theory in vacuum and water (dielectric constant, ε=78) by polarizable continuum model (PCM). The QM optimization helps to understand the distinction in conformational energies of different conformations formed from a particular amino acid. It has been found that right handed helix from L amino acid, and left handed helix from D amino acid are energetically similar and their energies are lower than right handed helix from D amino acid and left handed helix from L-amino acid. Further, MD simulation of all these conformations have been carried out using OPLS/AA, CHARMM27/CMAP and AMBER03 force fields in vacuum, and in water. The structural comparison between QM and MD is done on the basis of two distinct structural properties of helical configurations. One is from average of number of phi-psi dihedral angles falling in the alpha-helical region and the number of intra helical, i.e., i-i+4 h-bond. This average value from QM is 7 and it's same for all the conformers. MD simulation using CHARMM27 force field produces the closest number to QM, then AMBER 03, and then OPLS/AA. The force fields are validated from another method by calculating helical order parameter. The CHARMM27/CMAP and AMBER03 force field shows better agreement of helical order parameters with QM, particularly for energetically lower conformers of all amino acids except glutamic acid, but AMBER03 explains energetically higher conformers of lysine and glutamic acid better than CHARMM27/CMAP. This comparative study of force fields shows that both AMBER03 and CHARMM27/CMAP keeps peptide in α-helical conformation but AMBER-03 shows better agreement of helical order parameter to QM.

Next, we tried to understand the integrity of bilayers consisting of diblocks glycopolypeptide-polypropylene oxide chains. Now, whether there is any effect of α-helicity content of hydrophilic glycopeptides block on the stability of the bilayer have been studied. It has been elucidated by doing MD simulation of two bilayers where the built of α-helical glycopeptides

blocks are different. They comprise of purely l-amino acid in one bilayer, and alternate l and d-amino acids in another. The detailed insight about interaction among these diblocks chains in bilayer is obtained from performing all-atomistic MD simulation. The bilayer of which glycopeptide block consists of pure l-amino acid residue, found to be integrated after long time simulation. The diblocks chains remain more efficiently packed in this bilayer. Water penetrates in large amount and reaches into core of the bilayer where glycopeptide block consists of alternate d, l-amino acid. The percentage α-helicity of the chains consisting of alternate d, l amino is found to be higher than the chain consisting of l-amino acid which indicates that α-helicity does not play any role in the integrity of bilayer. There are different hydrogen bonds forming sites in the hydrophilic segment of diblocks chains. The higher number of side chain – main chain and side chain – side chain hydrogen bonds in the bilayer consisting of l-amino acids keep integrity of the bilayer.

Next we have changed our self-assembling material from di-block chains to tri-block component. The self-assembly of bolaamphiphile which has two hydrophilic group of different hydrophilicities connected through long hydrophobic tail, has been studied. The different concentrations of bolaamphiphile in water are taken for doing dissipative particle dynamics simulation. The different aggregated morphologies like spherical, cylindrical, lamellar, gyroid, etc have been found upon increasing its concentration in water. Earlier, we have limited our study only in the self-assembled lamellar phase with preassembled structure. All-atomistic MD simulation gives detailed insight about self-assembled structure but it limited our study in a particular phase instead of other phases also. Dissipative particle dynamics simulation facilitates simulation of bolaamphiphile at various phases starting from random arrangement of chains at different concentration in water. In all these self-assembled morphologies, two head groups of bolaamphiphile chains are found on the exterior facing water and they are also found to be assembled in core. Bolaamphiphiles chains are arranged in different orientations and positions in these self-assembled aggregates. They can flip from one orientation to another. The flipping frequencies of these bolaamphiphile chains have been compared for different orientations in various self-assembled morphologies. The flipping frequency is found to be lowest where two hydrophilic head groups are in the core of aggregates and it is highest where larger head group is in the core and smaller head group in the periphery.