Characterization of Hydrophobic Interactions of Polymers with Water and Phospholipid Membranes Using Molecular Dynamics Simulations

Mihaela Drenscko
The Graduate Center, City University of New York

How does access to this work benefit you? Let us know!
Follow this and additional works at: https://academicworks.cuny.edu/gc_etds

Part of the Other Physics Commons

Recommended Citation
https://academicworks.cuny.edu/gc_etds/2289

This Dissertation is brought to you by CUNY Academic Works. It has been accepted for inclusion in All Dissertations, Theses, and Capstone Projects by an authorized administrator of CUNY Academic Works. For more information, please contact deposit@gc.cuny.edu.
CHARACTERIZATION OF HYDROPHOBIC INTERACTIONS OF POLYMERS WITH WATER AND PHOSPHOLIPID MEMBRANES USING MOLECULAR DYNAMICS SIMULATIONS

by

MIHAELA DRENSCKO

A dissertation submitted to the Graduate Faculty in Physics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2017
The Full Title of Your Dissertation Should be Included on this Page:

by

Author Name

This manuscript has been read and accepted for the Graduate Faculty in Physics
in satisfaction of the dissertation requirement for the degree of Doctor of
Philosophy.

08/18/2017
Date

Sharon M. Loverde
Chair of Examining Committee

08/25/2017
Date

Igor L. Kuskovsky
Executive Officer

Supervisory Committee:

Qiao-Sheng Hu
Joel Koplik
Graziano Vernizzi
Shuiqin Zhao

THE CITY UNIVERSITY OF NEW YORK
ABSTRACT

Characterization of Hydrophobic Interactions of Polymers with Water and Phospholipid Membranes Using Molecular Dynamics Simulations

by

Mihaela Drenscko

Advisor: Dr. Sharon M. Loverde

Polymers and lipid membranes are both essential soft materials. The structure and hydrophobicity/hydrophilicity of polymers, as well as the solvent they are embedded in, ultimately determines their size and shape. Understating the variation of shape of the polymer as well as its interactions with model biological membranes can assist in understanding the biocompatibility of the polymer itself. Computer simulations, in particular molecular dynamics, can aid in characterization of the interaction of polymers with solvent, as well as polymers with model membranes. In this thesis, molecular dynamics serve to describe polymer interactions with a solvent (water) and with a lipid membrane.

To begin with, we characterize the hydrophobic collapse of single polystyrene chains in water using molecular dynamics simulations. Specifically, we calculate the potential of mean force for the collapse of a single polystyrene chain in water using metadynamics, comparing the results between all atomistic with coarse-grained molecular simulation. We next explore the scaling behavior of the collapsed globular shape at the minimum energy configuration, characterized by the radius of gyration, as a function of chain length. The exponent is close to one third, consistent with that predicted for a polymer chain in bad solvent. We also explore the
scaling behavior of the Solvent Accessible Surface Area (SASA) as a function of chain length, finding a similar exponent for both all-atomistic and coarse-grained simulations. Furthermore, calculation of the local water density as a function of chain length near the minimum energy configuration suggests that intermediate chain lengths are more likely to form dewetted states, as compared to shorter or longer chain lengths.

Next, in order to investigate the molecular interactions between single hydrophobic polymer chains and lipids in biological membranes and at lipid membrane/solvent interface, we perform a series of molecular dynamics simulations of small membranes using all atomistic and coarse-grained methods. The molecular interaction between common polymer chains used in biomedical applications and the cell membrane is unknown. This interaction may affect the biocompatibility of the polymer chains. Molecular dynamics simulations offer an emerging tool to characterize the interaction between common degradable polymer chains used in biomedical applications, such as polycaprolactone, and model cell membranes. We systematically characterize with long-time all-atomistic molecular dynamics simulations the interaction between single polycaprolactone chains of varying chain lengths with a model phospholipid membrane. We find that the length of polymer chain greatly affects the nature of interaction with the membrane, as well as the membrane properties. Furthermore, we next utilize advanced sampling techniques in molecular dynamics to characterize the two-dimensional free energy surface for the interaction of varying polymer chain lengths (short, intermediate, and long) with model cell membranes. We find that the free energy minimum shifts from the membrane-water interface to the hydrophobic core of the phospholipid membrane as a function of chain length. These results can be used to design polymer chain lengths and chemistries to optimize their interaction with cell membranes at the molecular level.
Acknowledgments

M. D. would like to thank Dr. Sharon Loverde from Department of Chemistry of College of Staten Island, Dr. Alex Tzanov from CUNY High Performance Computing Center, Lei Huang and Virginia Trueheart from the Texas Advanced Computing Center and Daniel Moy, Assistant Program Officer of Physics Program at Graduate Center of CUNY.
# Table of Contents

Abstract .................................................................................................................................................. iv

Acknowledgments .................................................................................................................................... vi

List of Tables ........................................................................................................................................... xi

List of Figures ........................................................................................................................................... xii

Chapter 1. Introduction .......................................................................................................................... 1

1.1 Research premises .............................................................................................................................. 1

1.2 Thesis organization ............................................................................................................................ 4

Chapter 2. Comprehensive background .............................................................................................. 7

2.1 Polymer science .................................................................................................................................. 7

2.1.1 Flory theory for a single polymer chain ....................................................................................... 8

2.1.2 Polymer in a solvent: Edwards model ......................................................................................... 12

2.1.3 Flory theory predictions ............................................................................................................... 13

2.2 Phospholipid bilayers ....................................................................................................................... 14

2.2.1 Biomembranes ............................................................................................................................... 14

2.2.1.1 Molecular constituents of biomembranes ............................................................................... 14

2.2.1.2 Lipid composition of biomembrane ....................................................................................... 15

2.2.1.3 Protein composition of biomembrane ....................................................................................... 15

2.2.1.4 Multimolecular structure of biomembranes ............................................................................ 15

2.2.1.5 Fluidity of the lipid bilayer ...................................................................................................... 16

2.2.2 Phospholipids ................................................................................................................................. 17

2.2.2.1 Phospholipid bilayer sheets ..................................................................................................... 19

2.2.2.2 Phospholipid bilayer fluidity ..................................................................................................... 20

2.3. Polymer interactions with lipid membranes .................................................................................... 21

2.3.1 Size and shape ............................................................................................................................... 21

2.3.2 Polymers ......................................................................................................................................... 21
2.4 Simulation techniques ...........................................................................................................22
2.4.1 Introduction: theoretical underpinnings of molecular dynamics ........................................22
2.4.2 Molecular dynamics .........................................................................................................22
2.4.3 Molecular interactions ......................................................................................................23
2.4.3.1 Non-bonded interactions ............................................................................................23
2.4.3.2 Bonded interactions ....................................................................................................25
2.4.3.3 Force calculations ........................................................................................................26
2.4.4 MD Algorithm ..................................................................................................................27
2.4.4.1 Verlet algorithm ..........................................................................................................28
2.4.4.2 Constraints ...................................................................................................................28
2.4.4.3 Periodic boundary conditions ......................................................................................30
2.4.4.4 Neighbor lists ..............................................................................................................31
2.4.5 Time dependence .............................................................................................................33
2.4.5.1 Liouville theorem ........................................................................................................33
2.4.5.2 Propagators and the Verlet algorithm .........................................................................36
2.4.6 Rigid molecule rotation ...................................................................................................37
2.4.7 Molecular dynamics in different ensembles .......................................................................38
2.4.8 Steps in molecular dynamics simulations .........................................................................43
2.4.9 Free energy methods ........................................................................................................45
2.4.9.1 Metadynamics ..............................................................................................................46
2.4.10 Coarse grained simulations ............................................................................................48
2.4.10.1 Multiscale simulations ...............................................................................................48
2.4.10.2 Coarse-grained (CG) models .....................................................................................49
2.4.10.3 Pseudoatoms ..............................................................................................................50
2.4.10.4 Finding optimum energy function $U_{CG}$ ..................................................................51
2.4.10.5 Iterative Boltzmann inversion ....................................................................................52
2.4.10.6 Force matching ..........................................................................................................53
2.4.10.7 CG models for polymers ...........................................................................................54
## Chapter 4. AA Molecular dynamics simulations of phospholipid bilayers with PCL

4.1 Lipid bilayer interactions with polymers .................................................................84

4.2 Introduction ...........................................................................................................84

4.3 Simulation Systems ...............................................................................................87

4.4 Free energy calculations .......................................................................................90

4.5 Chain collapse .....................................................................................................90

4.6 Radius of gyration and end-to-end distance .........................................................93

4.6.1 Initial simulations: PCL in water ......................................................................93

4.6.1.1 Initial simulations: PCL in water - metadynamics ....................................96

4.6.2 PCL in POPC membrane ................................................................................100

4.7 Membrane thickness ..........................................................................................103

4.8 Area per lipid .....................................................................................................106

4.9 Density profiles ..................................................................................................107

4.10 Tail order parameters .......................................................................................108
4.11 Solvent accessible surface area .............................................................. 110
4.12 Metadynamics and free energy calculations ........................................... 112
4.13 Summary .................................................................................................. 115

Chapter 5. CG Molecular dynamics simulations of phospholipid bilayer with PCL ....... 117

5.1 Coarse-grained models ............................................................................. 117
5.1.1 CG-model development for PCL .......................................................... 118
5.1.2 CG-model development for POPC ......................................................... 119
5.1.3 PCL in solvate membrane ................................................................. 121
5.1.4 Simulation parameters ........................................................................ 124
5.2 Coarse-grained PCL-POPC interactions .................................................. 125
5.3 PCL and POPC analysis tools .................................................................. 127
5.3.1 Radius of gyration ............................................................................... 127
5.3.2 End-to-end distance ............................................................................. 130
5.3.3 Membrane thickness .......................................................................... 133
5.3.4 Area per lipid ....................................................................................... 136
5.3.5 Density profiles .................................................................................... 138
5.3.6 Solvent accessible surface area ............................................................ 142
5.4 Conclusions ............................................................................................ 145

Chapter 6. Summary and Prospects ............................................................... 146

6.1 Summary .................................................................................................. 146
6.2 Prospects ................................................................................................. 147

References ..................................................................................................... 151
List of Tables

3.1. Coarse grained parameters for non-bonded interactions in PS-water systems ..................62
3.2. Coarse grained parameters for bonded interactions in PS-water systems (ring bonds) ......62
3.3. Coarse grained parameters for bonded interactions in PS-water systems (ring angles) ......62
3.4. Averaged $R_g$ for AA non-accelerated MD simulations of equilibrated systems ..........63
3.5. Averaged $R_g$ for CG non-accelerated MD simulations of equilibrated systems ..........63
3.6. AA metadynamics simulation results for all single PS chains in water systems ............69
3.7. CG metadynamics simulation results for all single PS chains in water systems ..........73
3.8. Averaged SASA for AA non-accelerated MD simulations of equilibrated systems ..........77
3.9. Averaged SASA for CG non-accelerated MD simulations of equilibrated systems ..........77
4.1. PCL in water system sizes and setups ...........................................................................87
4.2. Averaged $R_g$ for PCL in POPC AA MD simulations (equilibrated systems) ............88
4.3. Averaged $R_g$ and Re for PCL in water AA MD simulations (equilibrated systems) ....94
4.4. Averaged $R_g$ and Re for PCL in water metadynamics .................................................97
4.5. Averaged SASA for AA MD simulations (equilibrated systems) .................................112
4.6. (a) Metadynamics: radius of gyration at minimum free energy for 5PCL ..................115
4.6. (b) Metadynamics: radius of gyration at minimum free energy for 17PCL ....................115
4.6. (c) Metadynamics: radius of gyration at minimum free energy for 58PCL ....................115
5.1. PCL and water CG mapping ......................................................................................118
5.2. POPC CG functional groups .....................................................................................121
5.3. System sizes and setups ..........................................................................................122
5.4. Bond stretching parameters ....................................................................................123
5.5. Angle bending parameters ....................................................................................124
### List of Figures

2.1. Illustration of different polymers, in which $m$ represents the degree of polymerization ............................................. 7

2.2. Freely jointed model (FJC) of a polymer chain ............................................................................................................. 8

2.3. $R_g$ According to Flory’s mean field theory .............................................................................................................. 13

2.4. Biological membrane with lipid bilayer and other membrane constituents .......................................................... 16

2.5. Structure of a phospholipid ......................................................................................................................................... 18

2.6. Structural representation of a sterol .......................................................................................................................... 19

2.7. The lipid bilayer .......................................................................................................................................................... 19

2.8. Lennard-Jones pair potential ..................................................................................................................................... 24

2.9. Molecular geometry of a chain molecule .................................................................................................................. 26

2.10. Periodic boundary conditions ................................................................................................................................. 31

2.11. The Verlet list on its construction .......................................................................................................................... 32

2.12. The cell structure ....................................................................................................................................................... 32

2.13. Multiscale modeling .................................................................................................................................................... 49

2.14. A CG polyalanine molecule ..................................................................................................................................... 50

2.15. The coarse grain model is shown superimposed over the atomistic model ...................................................... 55

2.16. VMD snapshots of AA (upper right) and CG (lower right) simulations .............................................................. 57

3.1. AA and CG PS chain representation .......................................................................................................................... 59

3.2. Snapshots of AA PS in water at the minimum energy configuration (metadynamics) ................................ 65

3.3. Single AA PS chains in water (equilibrated systems) ............................................................................................... 65

3.4. Radius of gyration at minimum energy (AA, metadynamics) .................................................................................. 67

3.5. Radius of gyration vs. number of monomers (AA, equilibrated systems) ............................................................ 68

3.6. Snapshots of CG PS in water at the minimum energy configuration (metadynamics) ................................ 70
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.7</td>
<td>Single CG PS chains in water (equilibrated systems)</td>
</tr>
<tr>
<td>3.8</td>
<td>Radius of gyration at minimum energy (CG, metadynamics)</td>
</tr>
<tr>
<td>3.9</td>
<td>Radius of gyration vs. number of monomers (CG, equilibrated systems)</td>
</tr>
<tr>
<td>3.10</td>
<td>AA-CG comparison ( R_g ) vs. number of monomers at minimum energy</td>
</tr>
<tr>
<td>3.11</td>
<td>SASA vs. time (AA, equilibrated systems)</td>
</tr>
<tr>
<td>3.12</td>
<td>SASA vs. time (CG, equilibrated systems)</td>
</tr>
<tr>
<td>3.13</td>
<td>SASA’s standard deviation vs. number of monomers ( N ) (AA and CG)</td>
</tr>
<tr>
<td>3.14</td>
<td>SASA vs. number of monomers (AA and CG) for equilibrated systems</td>
</tr>
<tr>
<td>3.15</td>
<td>SASA vs. number of monomers (AA and CG) for metadynamics</td>
</tr>
<tr>
<td>3.16</td>
<td>Average number of hydration waters per SASA</td>
</tr>
<tr>
<td>3.17</td>
<td>Local water density profiles</td>
</tr>
<tr>
<td>4.1</td>
<td>Snapshots of PCL in POPC membrane at 0ns, 75 ns and 150ns (5, 17 and 57PCL)</td>
</tr>
<tr>
<td>4.2</td>
<td>Snapshots of PCL in POPC at the end of simulations for all PCL lengths</td>
</tr>
<tr>
<td>4.3</td>
<td>Time evolution of the radius of gyration for 5, 10, 13, 22, 30, 44 PCL in water</td>
</tr>
<tr>
<td>4.4</td>
<td>Time evolution of end-to-end distance for 5, 10, 13, 22, 30, 44 PCL in water</td>
</tr>
<tr>
<td>4.5</td>
<td>( R_g ) versus ( N ) for PCL in water (equilibrated systems)</td>
</tr>
<tr>
<td>4.6</td>
<td>( R_e ) versus ( N ) for PCL in water (equilibrated systems)</td>
</tr>
<tr>
<td>4.7</td>
<td>Snapshots of PCL in water (equilibrated systems)</td>
</tr>
<tr>
<td>4.8</td>
<td>Snapshots of PCL in water (metadynamics)</td>
</tr>
<tr>
<td>4.9</td>
<td>( R_g ) versus ( N ) for PCL in water (metadynamics)</td>
</tr>
<tr>
<td>4.10</td>
<td>( R_e ) versus ( N ) for PCL in water (metadynamics)</td>
</tr>
<tr>
<td>4.11</td>
<td>Free energy vs. ( R_g ) for PCL in water</td>
</tr>
<tr>
<td>4.12</td>
<td>Time evolution of radius of gyration of various single PCL chains in POPC</td>
</tr>
</tbody>
</table>
4.13. Time evolution of end-to-end distance of various single PCL chains in POPC ..........101
4.14. Comparison of radius of gyration of PCL in POPC and PCL in water .................102
4.15. Comparison of end-to-end distance of PCL in POPC and PCL in water .................102
4.16. Time evolution of membrane thickness due to insertion of various single PCL chains ...104
4.17. Average membrane thickness and area per lipid vs. number of monomers ............104
4.18. PCL chain positioning with respect to membrane-water interface ......................106
4.19. Time evolution of area per lipid due to insertion of various single PCL chains ........106
4.20. Average mass density distributions for of the PCL, water, and POPC ...................107
4.21. POPC tail order parameter for single PCL chains in solvated POPC .................109
4.22. POPC sn-1 ad sn-2 tail order parameters ............................................................110
4.23. Time evolution of SASA due to insertion of various single PCL chains .............111
4.24. SASA and SASA’s standard deviation vs. number of monomers .......................112
4.25. AA metadynamics free energy distribution contour graphs ...............................113
4.26. AA metadynamics snapshots taken at minimum energy ......................................114
5.1. Coarse grained models of PCL ...........................................................................117
5.2. Schematic definition of CG sites .......................................................................119
5.3. Snapshots of CG PCL in POPC membrane at 0, 75 and 150ns (short chains) ........125
5.4. Snapshots of CG PCL in POPC membrane at 0, 75 and 150ns (long chains) ..........126
5.5. Snapshots of CG PCL in POPC membrane at the end of simulations ................126
5.6. Rg time evolution .........................................................................................128
5.7. AA-CG time evolution Rg comparison ..............................................................129
5.8. AA-CG average Rg vs. N comparison ..............................................................130
5.9. Re time evolution .........................................................................................131
5.10. AA-CG time evolution Re comparison ...........................................................132
5.11. AA-CG average Re vs. N comparison ............................................................133
5.12. Membrane thickness time evolution ................................................................134
5.13. AA-CG time evolution membrane thickness comparison .........................................................135
5.14. AA-CG average membrane thickness vs. N comparison ..............................................................135
5.15. Area per lipid time evolution ........................................................................................................136
5.16. AA-CG time evolution area per lipid comparison ........................................................................137
5.17. AA-CG average area per lipid vs. N comparison ........................................................................137
5.18. CG density profiles (all) ..............................................................................................................139
5.19. CG density profiles (PCL) .........................................................................................................140
5.20. CG density profiles (POPC) .......................................................................................................141
5.21. CG density profiles (water) ........................................................................................................142
5.22. SASA time evolution ..................................................................................................................143
5.23. AA-CG time evolution SASA comparison ....................................................................................144
5.24. AA-CG average R\textsubscript{e} vs. N and SASA vs N comparison .................................................145
CHAPTER 1

Introduction

1.1. Research premises

Polymers are a large class of materials consisting of many repeating units (monomers) linked to form long chains that may include tens of thousands of monomers. Naturally, occurring polymers such as rubber have been used long before chemists first developed the production of first synthetic polymers (1). Polymers, such as polystyrene, have diverse applications in industry, ranging from polymer melts (2) used in plastic and rubber manufacturing industries (3) to biomedical applications of polymer-composites (4). Furthermore, advances in synthetic polymer chemistry have allowed for the synthesis of hybrid polymers (block or multiblock copolymers) that can self-assemble into ordered structures of various morphologies (spheres, cylinders, bicontinuous structures, lamellae, vesicles and other assemblies), that possess a wide range of applications in soft materials (5,6) as well as biomaterials (7). Self-assemblies with polystyrene in particular often utilize the miscibility of polystyrene in different solvents (dimethylformamide, tetrahydrofuran, dioxane) to control the kinetics and resulting shape of the self-assembly (6,8,9). For example, solvent effects have been shown to control the end to end vs. side-to-side aggregation of polystyrene functionalized nanorods (10). Additionally, the aggregate shape of diblock functionalized nanoparticles self-assemblies, including ellipsoidal micelles, curved lamellae, and nanospheres, has been shown to be strongly dependent on the solvent and water content (8). Thus, multiscale computational models that accurately describe the behaviour of polystyrene in water (and other solvents) can add further insight into the self-assembly behaviour of these PS-based assemblies. Another example is polycaprolactone (PCL), which has been approved by the Food and Drug Administration (FDA) in bio-medical applications and technologies, such as drug
delivery devices, (7), sutures (known as Monocryl) and adhesion barriers (11). It is being used in tissue engineering research (12) and as the hydrophobic block of amphiphilic synthetic block copolymers in forming the vesicle membrane of polymersomes (13,14). A variety of drugs have been encapsulated within PCL beads for controlled release and targeted drug delivery, such as naltrexone hydrochloride (15). PCL’s rheological and viscoelastic properties allow for manufacturing of a great variety of scaffolds (16-20).

Polymers’ various functions and properties are best illustrated in their interactions with biological membranes, such as lipid bilayers. The lipid bilayer (or phospholipid bilayer) is a polar membrane (21,22) arranged in two layers of lipid molecules (21). The membrane acts as a barrier of selective permeability (23,24) surrounding cells of living organisms, viruses and sub-cellular structures, and preventing various chemicals to diffuse through the membrane. Even with their nanoscale size (25), bilayers are not permeable by hydrophilic molecules, such as sugar and particularly to ions, but compounds with hydrophobic properties such as carbon dioxide can pass through.

Lipid bilayers are invisible to the traditional microscope and are very fragile, therefore are difficult to study. The lipid interactions with one another or with a solvent can be modeled as hydrated lipid bilayers, in which the complexities of the interaction analysis are addressed by computational methods.(26) Also, advanced experimental techniques, such as AFM (27)and EM (28), NMR (29) and molecular simulations (30) can offer insights into the forces that govern microscopic processes in bilayers.

Lipid bilayers, in their liposome form (31) are used successfully for drug delivery. Drugs in solutions are encapsulated within the liposomes, and once inside a patient, they transport to the
target area, where the drug is released (32). A supported lipid bilayer (33) can measure permeability of drugs (34).

Lipid membrane investigations are difficult to approach directly, due to their complex nature. However, model membranes assist essentially in the study of membrane characterization (35). One such model refers to lipid membrane interacting with polymers (36-39), in which polymer chains are inserted into the membrane, upon which the effect of molecular interaction is studied. Particularly, the hydrophobic interactions between hydrophobic lipid tails in the membrane and anchoring groups are of importance in the process of binding of the polymer chain to the membrane. Crystal microbalance with dissipation experiments (35) showed that a higher polymer concentration favors insertion of polymer chains in the lipid membranes.

In this thesis models for polymers polystyrene (PS) and polycaprolactone (PCL) in different environments will be developed theoretically and analyzed with molecular dynamics simulations. The model of PS in water characterizes the immiscibility of this polymer in the context of shape transition as a result of the hydrophobic interactions between individual chains of different sizes and water. For PCL model in a biological membrane environment, the repulsive interactions between individual chains PCL and POPC (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine) bilayer determines deformations of the membrane and the increase in chain entropy. POPC (40,41) (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine) is a phospholipid that is major component of biological membranes. (42,43) and it is present naturally in eukaryotic cell membranes (44). The theoretical framework of drawing up the two models of PS in water and PCL in a lipid bilayer are based on force fields for molecular dynamics. The computer simulation packages associated with these force fields will serve to analyze these models as means of understanding the structural properties of polymer chains, membrane and interfaces and the microscopic interactions between
them. Aside from conventional experiments, computer simulations provide unique information that cannot be obtained in other way, due to complexities of self-assembly behavior of PS in water and PCL diffusion through a bilayer. Accurate predictions of bulk properties as well as otherwise inaccessible details of bulk measurements can be obtained. Theoretical techniques, such as mean field theory (45) can be used to describe the statistical behavior of a polymer system and, while such techniques will continue to improve their applicability, simulations are the optimum tool to bridge the gap between microscopic length and time scales and macroscopic measurements in a lab, and between theoretical and experimental results. Ultimately, the simulation model of molecular interactions does not necessarily have to be perfectly realistic. For example, a coarse-grained model (46) contains the essential physics and has the advantage to allow for longer simulation times than in classical atomistic models. Computer simulation techniques will provide insightful details into the layout and mechanics of polymers and lipid membranes for various applications in medicine, medical technologies and biology.

### 1.2 Thesis organization

Chapter 2 begins with an overview of selected topics in the theory of polymer and phospholipid bilayer that will be utilized in this thesis. Background information is provided for computer simulation techniques, specifically molecular dynamics, as applied in molecular modeling of macromolecules.

In Chapter 3, a theoretical model for polystyrene (PS) is introduced. This model is used to characterize the hydrophobic collapse of single PS chains in water using molecular dynamics simulations. Specifically, potential of mean force for the collapse of a single polystyrene chain in water is calculated using metadynamics, comparing the results between all atomistic with coarse-
grained (CG) molecular simulation. Investigations of the scaling behavior of the collapsed globular shape at the minimum energy configuration, show that the exponent the power fit of radius of gyration as a function of chain length, is close to one third, consistent with that predicted for a polymer chain in bad solvent. Analog, investigations of scaling behavior of the solvent accessible surface area (SASA) as a function of chain length, reveal a similar exponent for both all atomistic and CG simulations. Furthermore, calculation of the local water density as a function of chain length near the minimum energy configuration suggests that intermediate chain lengths are more likely to form dewetted states, as compared to shorter or longer chain lengths.

Next in Chapter 4, the idea of solvated phospholipid morphology modifications due to insertion of polymer chains is presented. In order to investigate the microscopic interactions between single hydrophobic polymer chains and lipids in biological membranes and at lipid membrane/solvent interface, a series of molecular dynamics simulations were carried out first for PCL in water, then for PCL in a small membrane. PCL in water simulations were run for 50 ns in each case, and for PCL in lipid membrane simulations were run for 150 ns each, with hydrated 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) bilayers of 59 lipids and single polycaprolactone (PCL) chains of various sizes. PCL length was the parameter by which the structural properties and dynamics of the bilayer were determined as well as comparing the effects of solvent on chain conformation. The effects on the membrane include an increased disorder of the POPC chains (tail order parameters), membrane thinning, an increase in area per lipid with minimal fluctuations and the corresponding increase in POPC mass density. Altogether, the PCL interactions with POPC bilayer were characterized by the increase in molecular entropy of mixing between the PCL/POPC and PCL/water.
In Chapter 5, the investigations from Chapter 4 are replicated using a coarse-grain method. The results are compared and contrasted in terms of the changes in PCL chain conformations during simulations, POPC deformation and computation time advantages in each case.
CHAPTER 2

Comprehensive Background

2.1. Polymer science

Polymers as solutions or mixtures are a large category of soft condensed matter. Many materials naturally occurring are made of polymers, such as nucleic acids or starches. Polymers can be synthetic in origin, such as plastics. Regardless of their composition or origin, polymers properties are derived from a common structural characteristic: a long chain like macromolecule composed of linearly connected units (or monomers) produced by the chemical bonding of the monomers (47). Addressing the correlations between monomers is the primary task in polymer physics.

Polymer atomic composition is based on carbon. Polymer chains can have a simple structure, in which each carbon in the main chain is coupled with two hydrogens (polyethylene), or more complicated, to include other atoms (nylon), or may not include carbon atoms polydimethylsiloxane, as illustrated in Figure 2.1

![Figure 2.1. Illustration of different polymers, in which m represents the degree of polymerization. a) Polyethylene, (CH\textsubscript{2})\textsubscript{m}, simple chemical structure of carbon main chain with two hydrogen atoms per carbon. b) Nylon 6-10, (C\textsubscript{10}H\textsubscript{18}O\textsubscript{4}·C\textsubscript{6}H\textsubscript{16}N\textsubscript{2})\textsubscript{m}, with other atoms aside carbon on the main chain. c) Polydimethylsiloxane, (C\textsubscript{2}H\textsubscript{6}OSi)\textsubscript{m}, with no carbon atoms in the main chain.](image_url)

According to positioning of the side groups, polymers can be classified as isotactic (side groups are on the same side of the chain), syndiotactic (groups arrange on alternates sides), or at random (atactic). In particular, the atactic arrangement is related to the disorder of polymer chain due to high barriers of rotation of odd side groups, resulting in the absence of crystallization at low temperatures.
Polymers chains can be linear or branched. Linear polymers are defined by the degree of polymerization, or the number of monomers present in the chain, which is proportional to the relative molecular mass of the chain. One type of repeating unit in a polymer chain classifies it as homopolymer, and copolymer, respectively, when more types of units are present in the chain. At a given temperature, polymeric material can be in liquid state (melts and solutions), glass (un crystalized polymers), crystalline (incomplete crystallization of liquid or glassy synthetic or natural materials) and liquid crystalline (rigid materials).

### 2.1.1. Flory theory for a single polymer chain

For an isolated polymer of N + 1 monomers at positions \{\mathbf{r}_0, \mathbf{r}_1, ..., \mathbf{r}_N\} in space (Figure 2.2), the distance between adjacent monomers is called Kuhn length \(b\),

\[
b = \mathbf{r}_j - \mathbf{r}_{j-1} = \mathbf{\tau}_j
\]

(2.1)

where \(\mathbf{\tau}_j\) is the bond vector (48). The special path of the polymer chain is a random walk, or "self-avoiding walk" (SAW), allowing no double occupancy of sites or crossing. End-to-end distance \(\mathbf{R}\), is the direction and magnitude of each bond vector:

\[
\mathbf{R} = \mathbf{r}_N - \mathbf{r}_0 = \sum_{j=1}^{N} \mathbf{\tau}_j
\]

(2.2)

*Figure 2.2. Freely jointed model (FJC) of a polymer chain formed by a random walk of rigid segments of the same length in two-dimensional space. For an isolated chain, there is no orientation preference of bond angles (equal probability for all bond angle values and uncorrelated to the orientations of all other bonds in the chain); bond junctions undergo free rotation; there is no self-interactions of segments, such as excluded volume effects (Drawn from Grosberg, A.(2012). Introduction to Polymer Physics, Lecture 1, Boulder Summer School [PowerPoint Slides].)*
The mean end-to-end distance is:

\[
\langle \mathbf{R} \cdot \mathbf{R} \rangle = \left\langle \left( \sum_{i=1}^{N} \mathbf{\tau}_i \right) \cdot \left( \sum_{j=1}^{N} \mathbf{\tau}_j \right) \right\rangle
\]  

(2.3)

\[
\langle \mathbf{R}^2 \rangle = N b^2 + \left\langle \sum_{i \neq j} \mathbf{\tau}_i \cdot \mathbf{\tau}_j \right\rangle
\]  

(2.4)

A freely joined chain is a flexible polymer in which the bond vectors can orient in any directions in space, for any uncorrelated bond vectors \( \mathbf{\tau}_i \) and \( \mathbf{\tau}_j \), the cross-terms disappear when the average is taken. So for a random walk, the root mean squared end-to-end-distance is:

\[
\langle \mathbf{R}^2 \rangle = N b^2
\]  

(2.5)

\[
R \sim b N^\nu
\]  

(2.6)

in which \( \nu \) is the scaling coefficient, which has the value of \( \frac{1}{2} \) for a freely joined chain.

More complex polymer chain structures appear as networks. Particularly, when networks scale with the distance the chain is defined by a fractal dimension, \( d \). The total length of the polymer inside of a sphere of radius \( R \), with its center on any position in the chain is \( R^d \), where \( d=1/\nu \). So for simple random walk \( \nu = 1/2 \) and \( d = 2 \). In the limit of large \( N \), the distribution of end-to-end distance is Gaussian (47). A polymer chain characterized by a Gaussian distribution is called an ideal chain, with a probability distribution given by:

\[
P(\mathbf{R}, N) \approx \left( \frac{3}{2\pi N b^2} \right)^{3/2} \exp \left[ -\frac{3\mathbf{R}^2}{2N b^2} \right]
\]  

(2.7)
The configurational entropy as a function of end-to-end distance is given by Boltzmann’s equation, \( S = k_B \ln \Omega(R) \), in which \( \Omega(R) \) is the number of monomers’ conformations that would arrange to give the chain end-to-end distance \( R \):

\[
S(R, N) = -\frac{3k_B R^2}{2Nb^2} + S_0
\]

(2.8)

Stretching the chain results in the decrease of its entropy, with less chain conformations than in an unperturbed state. Consequently, the free energy increases:

\[
F(R, N) = \frac{3k_B T R^2}{2Nb^2} + F_0
\]

(2.9)

A stretched polymer chain behaves like a spring, being acted upon by a restoring force as a result of being extended beyond the random walk value. Such force is entropic in nature, as the system tends to acquire a state of maximum entropy as a result of thermal fluctuations. The chain becomes stiffer with increased temperature. When the chain conformation is significantly perturbed, that is, when the deformation is larger than the contour length of the polymer chain, then the chain experiences a non-Gaussian behavior, subject to non-Gaussian statistics. Monomers’ interaction becomes repulsive and cannot interpenetrate. The minimization of the free energy corresponds to a size of the chain in which the polymer unconstrained.

In a real polymer, chain-neighboring bonds are constrained to have specific bond angles, such the cross-terms in root mean square end-to-end distance do not disappear. The random-walk characteristics of the polymer are not changed by these short-range correlations of the neighboring bonds, but the length of bond vector changes. In an isolated chain, the interactions between distant bonds influence the size of the chain. These interactions are repulsive in nature, resulting in a non-Gaussian behavior of the polymer chain. Considering the fact that the chain cannot intersect itself.
and that the monomers cannot overlap, the path of the polymer can be defined in this case as a self-avoiding walk. From

\[
\langle R^2 \rangle^{1/2} = bN^\nu
\]  \hspace{1cm} (2.10)

The scaling coefficient \( \nu > \frac{1}{2} \), showing that the effect of no-overlapping of monomers is the swelling of the chain beyond the random walk value. Flory’s theory and the idea of excluded volume aids in determining the value of the scaling coefficient. If \( N \) monomers of the chain are considered a gas of \( N \) atoms, each of volume \( v \), inaccessible to other atoms, all together occupying volume \( V \), in the shape of a coil of radius \( R \), the entropy per atom is reduced by a factor \( k_B v N / V \), such that the free energy due to excluded volume,

\[
F_r = k_B T v \frac{N^2}{2R^3}
\]  \hspace{1cm} (2.11)

The excluded volume effect causes the monomers to repel each other, causing chain expansion, with an elastic contribution to free energy of:

\[
PE_{el} = k_B T \frac{R^2}{Nb^2}
\]  \hspace{1cm} (2.12)

From the minimization of total energy \( F = F_r + PE_{el} \), and considering the excluded volume to be \( v = b^3 \), the size of the size of the polymer coil, \( R \) is given by:

\[
R \sim bN^{3/5}
\]  \hspace{1cm} (2.13)

By generalization, the free energy of repulsion,

\[
F_r = k_B T v \frac{N^2}{2R^d}
\]  \hspace{1cm} (2.14)

where \( R^d \) is the volume occupied by the polymer. The minimization of the total free energy gives a scaling coefficient \( \frac{3}{d+2} \). For \( d \leq 4 \), the chain is in swollen phases, while if \( d > 4 \), the chain is compact.
In the case of attractive interaction between monomers, with the core repulsion, the monomers in the chain tend to be closely packed together. The polymer chain is now in a collapsed, globular state, in which

\[ R \sim N^{1/d} \]  

(2.15)

2.1.2. Polymer in a solvent: Edwards model

Aside from the macroscopic interactions between monomers, the polymer chain is subject to interactions with the molecules of the solvent. When polymer is dissolved in a solvent, the monomers are surrounded in great part by the molecules of the solvent, their mutual positions are as if they repel each other. This is the situation of a polymer in good solvent, described by effective repulsion among monomers. If the polymer separates from a solution, the monomers tend to avoid the solvent molecules. This is the case of polymer in a bad solvent, characterized by attraction between the monomers. Thus, the interaction between monomers depends on temperature, solvent quality and other parameters.

A polymer in good solvent is described by a pairwise contact repulsive two-body interaction, in which a polymer with excluded volume interaction with a self-avoiding walk path in space is represented by Edwards Hamiltonian (49). For the poor solvent case of polymer collapse, a three-body potential is used.

It has been observed that a polymer can be separated from a solution by cooling, so there exists a temperature \( T_0 \) such that for any temperature \( T > T_0 \), the polymer is in a good solvent (repulsive case), and for \( T < T_0 \), the polymer is in a bad solvent (attractive case). The idealized case is when \( T = T_0 \), the transition point.
2.1.3. Flory theory predictions

In summary, Flory’s theory predicts the following (50), Figure 2.3.:

1. Flory regime, $T > T_\theta$: maximization of monomer-solvent contacts; chain expansion
   \[ \nu = \frac{3}{d+2}, \quad 2 \leq d \leq 4 \] (swollen phase)
   \[ \nu = \frac{1}{2}, \quad d > 4 \] (coiled phase)

2. Theta regime, $T = T_\theta$: balanced interactions; unperturbed chain as an ideal random walk, and excluded volume repulsion is canceled by the attractive effect of polymer-solvent interaction
   \[ \nu = \frac{2}{d+1}, \quad 2 \leq d \leq 3 \] (cross-over behavior from theta temperature to good solvent limit)
   \[(51).\]
   \[ \nu = \frac{1}{2}, \quad d > 3 \] (Gaussian chain)

3. Bad solvent, compact regime, $T < T_\theta$: minimization of monomer-solvent contact; globular state of chain collapse, when the attraction effect of solvent-polymers interaction is much greater than the excluded volume repulsion
   \[ \nu = \frac{1}{d} \] for any $d$.

---

Figure 2.3. According to Flory’s mean field theory, the size of the polymer $R$ depends on the number $N$ of bond segments (or monomers in the chain) by a power law, $R \sim N^\nu$, where $\nu$ is the scaling exponent. a) long range repulsion, polymer chain is not expanded $\nu = 1$, $R \approx L \approx bN$; b) good solvent (b), $\nu = 3/5$ and the polymer behave like a one dimensional line object; c) $\theta$ solvent, $\nu = 1/2$, in which the polymer assumes the random walk coil dimensions; d) bad solvent, $\nu = 1/3$, in which the polymer shrinks in size and behaves like a globule. (Drawn from Grosberg, A. (2012). Introduction to Polymer Physics, Lecture 1, Boulder Summer School [PowerPoint Slides].)
2.2. Phospholipid bilayers

2.2.1. Biomembranes

Biomembranes (52) represent the external boundary of cells, controlling the movement of molecular constituents across this boundary and separating various internal spaces and processes. They are composed of one or more layers of combinations between lipids and other immiscible compounds (50) with reducing surface tension properties; are very thin relative to their curvature, as to be consider two-dimensional; they act as a barrier between liquid media and supporting expansion/compression, bending and sheer stresses caused by interface forces; they play a central role in structure and function of cells (53), facilitating the passage of ions or molecules between inside and outside of a cell. Physical properties of membranes determine their biological functions, such as toughness, flexibility and selective permeability to polar solutes. Cells can grow and move due to flexibility of the membrane, can fuse as a result of sealing fissures and can retain specific compounds and ions and exclude others. Moreover, biomembranes are responsible for the information transmitted through conformational changes induced in membrane components and are involved in bio-chemical functions in cells, such as prokaryotic DNA replication, protein biosynthesis, protein secretion, bioenergetics.

2.2.1.1. Molecular constituents of biomembranes

Biomembrane composition is essential in determining membrane function as well as in the study of membrane structures and constructs of membrane models that must validate membrane composition. Polar lipids and protein are the most important and massive constituents of a biomembrane, followed by carbohydrates in glycoproteins and glycolipids. The number of proteins and lipids vary in different biomembranes, reflecting a great variety of biomembrane roles.
2.2.1.2. Lipid composition of biomembranes

Biomembranes from various sources have similar characteristics, but specific lipid composition according to species of origin, tissue and cell type. (54). The types and quantities of lipids synthesized are controlled by cells, which can also direct a particular class of lipids to specialized cellular part. Such lipid combinations support development and evolution of cells without revealing their functionality.

2.2.1.3. Protein composition of biomembranes

There is a great variation in protein composition of biomembranes, more than the lipid part (53). Such variation confirms different functions that proteins have in a biomembrane (55), such as absorbing nutrients, expelling wastes, maintaining the proper intracellular ionic composition.

2.2.1.4. Multimolecular structure of biomembranes

Common fundamental properties of all biomembranes include impermeability to polar or charged solutes, permeability to nonpolar compounds; thickness of 5-8 nm; two layers appearance in cross section, indicated by electron microscopy (56). A proposed molecular model for biomembranes is the fluid mosaic model (57), supported by biomembrane’s chemical structure, electron microscopy and experimental evidence of protein and lipid permeability and movement in the membrane (Figure 2.4).
Figure 2.4. Biological membrane with lipid bilayer and other membrane constituents. On the inner surface of the membrane, the complex network of filaments and tubules from the cytoplasm influence mechanical properties of the membrane. On the outer surface, the carbohydrate network causes cell recognition and adhesion. (Drawn from W.H. Freeman and Company (2008). Molecular Cell Biology, Sixth Edition)

2.2.1.5. Fluidity of the lipid bilayer

The lipid bilayer is a stable structure in which individual molecules of lipid and sterol can move freely in the plane of the biomembrane. The acyl chains rotate about the carbon-carbon bonds, such that the hydrocarbon chain of fatty acids are in constant motion. Thus, the interior of the bilayer is also fluid. Lipid composition and temperature determine the state of fluidity. At low temperature, there is no much motion within the lipids, keeping the bilayer in a state of paracrystalline array (58). Above a certain specific temperature to each membrane, lipids move fast, causing a transition from paracrystalline structure to fluid. The transition temperature is determined by the distribution of saturated fatty acids and by the sterol content. Lipid bilayers play the role of biological barriers between aqueous compartments in the cells of living organisms (59). Another function comprises signaling through the cell, for example – in triggering phagocytosis (60), and in permeability to polar molecules, particularly to water (61). Apolar molecules gave greater mobility through the lipid bilayer (62).
Mechanical properties of lipid bilayers are mostly the same as those of liquids and solids (63), such as area compression modulus $K_a$, bending modulus $K_b$, edge energy $\Lambda$, shear modulus $G$ (for solid bilayers). These quantities measure molecular ability to attach and cross the bilayer and the bilayer’s deformations when subjected to stress factors. $K_a$ measures the energy needed to stretch the bilayer by measuring the area of lipids in contact with water (50), while $K_b$ measures the bending energy of the bilayer (64). $\Lambda$ measures the energy of bilayer tearing such that it would be exposed to water (65).

### 2.2.2. Phospholipids

The lipid bilayer is a core component of a biomembrane having a thickness is of order of nanometers (14,25). The outer portion, consisting of hydrophilic head group has a thickness of ~0.9 nm and it is completely hydrated. Next to it there is a region less hydrated of 0.3 nm, followed by the hydrophobic core of ~4nm thickness. Two types of molecules make the lipid bilayers: amphipathic phospholipids and sterols (66).

A phospholipid (67) (Figure 2.5) is a lipid molecule containing a hydrophobic core composed of two acyl chains linked through a glycerol backbone to a polar hydrophilic head group. The glycerol backbone is derived from a three-carbon molecule. The fatty acid chains are long nonpolar chains of carbon atoms ending in a carboxyl (-COOH) group. The last backbone carbon is attached to a polar alcohol that forms hydrogen bonds with water. This alcohol is attached by a phosphate group that gave the name “phospholipid” to this particular lipid molecule. Thus one end of the phospholipid is strongly nonpolar - the fatty acid hydrocarbon tail has no charge and is repelled by water (water-insoluble). The other end is strongly polar - the phospholipid head containing negatively charged phosphate group, attracted to water due to its polarity (water-soluble). Such a phospholipid, characterized a hydrophilic head and a hydrophobic tail is called amphipathic.
Phospholipids are one of the principal components of cell membranes (in addition to membrane proteins). All phospholipids have the following common basic structure: polar organic molecule (e.g. choline, serine); phosphate group; glycerol molecule two fatty acid tails (may be saturated or unsaturated). (Drawn from Moore R., Clark D., Vodopich D. (1998). Botany Visual Resources Library The McGraw-Hill Companies, Inc. All rights reserved.)

Sterols are a subgroup of steroids with four rings hydrocarbon framework arranged, having a hydroxyl group at position 3 of A ring (68) (Figure 2.6). Sterols can be found in the plasma membranes of eukaryotic cells and in minute amounts in some prokaryotic cells. In particular, cholesterol affects the structure and fluidity of the bilayer (69).
When placed in water, the phospholipids form a double layer, or bilayer (Figure 2.7) in which the nonpolar regions are facing each other at the core, while the polar heads facing outwards, in contact with water. The bilayer center excludes water or water-soluble molecules. This is a very stable structure surrounding the cell due to hydrophobic interactions between tails and hydrophilic interactions between tails and water. Proteins are embedded in the bilayer at positions held by hydrophobic interactions with the lipids. The lipid interactions and lipid-protein interactions are noncovalent, allowing for lateral motion of the lipids and proteins in the plane of the membrane.

2.2.2.1. Phospholipid bilayer sheets

When phospholipid molecules are placed in water, polar water molecules repel nonpolar tails of the phospholipids, as water molecules interact to form hydrogen bonds. Polar hydrophilic heads orient toward water, with nonpolar tails away from water, in a densely packed pattern and facing in other. In such a bilayer structure, tails do not come in contact with water. Lipid packing in the
bilayer can affect membrane mechanical properties (70), such as stretching and bending (71). Experimental studies have been conducted on model vesicle bilayer in drug delivery devices (72).

Water molecules have a tendency to form hydrogen bonds with other molecules (adhesion), leading to the spontaneous formation of lipid bilayers. Any water-soluble solutes cannot pass through the nonpolar bilayer core. Membrane proteins embedded in biomembrane, however, allow hydrophilic compounds pass through the membrane (73).

The hydrophobic effect (74), characterized by interactions between hydrophobic molecules drives the assembly of phospholipids into a bilayer sheet. These interaction cause delimitations of hydrophobic regions in which water molecule bond freely with each other, thus increasing the entropy (75) of the system. The assembly process includes non-covalent interactions, such as van der Waals and hydrogen bonds. Bilayer structure is maintained under variations in the external aqueous environment, such as pH and ion concentration (76).

2.2.2.2. Phospholipid bilayer fluidity

As the water’s tendency for hydrogen bonding is continuous, hydrogen water bonds holds together the biomembrane. While this process takes place spontaneously, relative position of specific phospholipids with respect to their neighbors cannot be determined. As a consequence, phospholipids and unfixed proteins can move within the membrane, making phospholipid bilayer fluid, with an increasing viscosity for lower temperatures. Upon closed alignment, individual tails are attracted to one another, thus restricting free movement inside the bilayer and decreasing fluidity of the membrane. Certain double bonds between carbon atoms in the tails cause misalignment and, therefore, less fluidity of the membrane.
2.3. Polymer interactions with lipid membranes

Nanosized external materials interacting with a biomembrane alter its functioning, composition and structure, along with dynamic properties and motility. Nanosized particle are characterized by a large surface to volume ratio, which increases their exposure to the environment. More surface is available to react at faster rates, which is a remarkable property with applications in medical technologies of particle delivery to cells. Particularly, understanding the molecular mechanics of nanoparticle entry into the cell without damaging it could help in creating materials able to reach safely inside cells (77).

2.3.1 Size and shape

The size of the interacting nanoparticles determines the means of entry into the membrane. Particles of size less than 10 nm transport across the membrane without disrupting the membrane. Changes of in size and shape affect the interactions between nanomaterials and the membrane, thus changing membrane properties.

2.3.2 Polymers

When polymers interact with lipid membranes, both undergo morphological changes according to their charge, composition and size (78). Physical characteristics such as hydrophobicity, hydrophilicity, and surface charge of polymer coatings can determine their interactions with lipids. For example, poly (amidoamine) dendrimers disrupt lipid bilayers by forming holes (79,80). If the diameter of the polymer particle is less than the membrane thickness, then the polymer is absorbed more easily into the lipid bilayer (81).
2.4. Simulation techniques

2.4.1 Introduction: theoretical underpinning of molecular dynamics

Molecular dynamics (MD) simulation at present time is highly useful in a variety of ways which can otherwise be inaccessible through experiment. It can provide accurate information between macroscopic properties of matter in liquid, solid and gaseous state (i.e. equation of state, transport coefficients, structural order parameter, etc.) and microscopic details of the system (i.e. masses of atoms or molecules, electrostatic interactions, molecular geometry, etc.). Accuracy of results larger systems is improving with the continuous increase in computing power. MD simulations have many applications in chemistry, biochemistry, physics, materials science and engineering (82).

2.4.2 Molecular Dynamics

Computer simulations complete the experimental and theoretical methods of investigating molecular assemblies. Aside from molecular dynamics (MD), there is Monte Carlo (MC) technique, having the advantage of providing a direct way to determining dynamical and rheological properties of a system. Additionally, a variety of hybrid methods can be combined to suit particular lines of research. Thus, simulations serve as a bridge between theory and experiment: a theory can be tested and compared with the results of an experiment. It is also possible to simulate extreme conditions that are hard to obtain in a lab for special materials, and using an appropriate molecular model, containing the essential physics (83).
2.4.3 Molecular interactions

Molecular dynamics predicts atomic trajectories by direct integration of Newton’s equations of motion for a simple atomic system describing the state of a conservative system of atoms of identical masses $m_i$, defined with an appropriate inter-atomic potential and initial and boundary conditions. Newton’s second law for each atom can be written as:

$$m_i \ddot{r}_i = F_i$$  \hspace{1cm} (2.16)

So that the corresponding force component for each atom is:

$$F_i = - \frac{\partial U(r^N)}{\partial r_i}$$  \hspace{1cm} (2.17)

Forces $F_i$ acting on the atoms, are usually derived from a potential energy:

$$U(r^N) = U(r_1, r_2, ..., r_N)$$  \hspace{1cm} (2.18)

in which $r^N = (r_1, r_2, ..., r_N)$ is the complete set of 3N atomic coordinates.

2.4.3.1 Non-bonded interactions

The non-bonded interactions between atoms potential energy $U^\text{non-bonded}$ is split into 1-body, 2-body, 3-body . . . terms:

$$U^\text{non-bonded}(r^N) = \sum_i u(r_i) + \sum_i \sum_{j>1} v(r_i, r_j) + ...$$  \hspace{1cm} (2.19)

where $u(r)$ is an external potential energy, which is dropped for fully periodic simulations of bulk systems. Only pair potential interactions of continuous nature, $v(r_i, r_j) = v(r_{ij})$ are considered in these simulations, while any higher order are neglected.

Lennard-Jones potential is given by:

$$v^{LJ}(r_{ij}) = \varepsilon \left[ \left( \frac{\sigma}{r_{ij}} \right)^{12} - \left( \frac{\sigma}{r_{ij}} \right)^6 \right]$$  \hspace{1cm} (2.20)
where $\sigma$ is the finite distance at which the inter-particle potential is zero, and $\varepsilon$ the depth of the potential well.

The term $\left(\frac{\sigma}{r_{ij}}\right)^{12}$ models a strong repelling force at very short distances, while the term $\left(\frac{\sigma}{r_{ij}}\right)^{6}$ models an attractive force at very long interaction distances. Lennard-Jones potential was used, initially to study the properties of liquid argon (Figure 2.8). For simulations in which attractive interactions matter less than the excluded volume effects concerning molecular packing, the potential can be truncated at its minimum position, and shifted upwards. This is the WCA model.

![Figure 2.8. Lennard-Jones pair potential, with $\left(\frac{\sigma}{r_{ij}}\right)^{12}$ and $\left(\frac{\sigma}{r_{ij}}\right)^{6}$ contribution and WCA shifted repulsive potential.](image)

The force on particle i by particle j due to Lennard-Jones potential is given by:

$$F_{ij} = -\nabla V_{LJ}(r_{ij}) = 4\varepsilon \left[12 \left(\frac{\sigma}{r_{ij}}\right)^{12} - 6 \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] \left(\frac{r_{ij}}{r_{ij}^2}\right)$$

In the presence of electrostatic charges, Coulomb potentials are given by:
where \( Q_i, Q_j \) are the charges of particles I and j, and \( \varepsilon_0 \) is the permittivity of free space.

The electrical force on particle i due to particle j is given by:

\[
F_{ij}^{\text{Coulomb}} = \frac{Q_i Q_j r_{ij}}{4\pi \varepsilon_0 r_{ij}^3}
\]

\[ (2.23) \]

### 2.4.3.2 Bonded Interactions

Molecules are subject to intramolecular bonded interactions. The chemical bonds between molecules, and are not created or broken during a simulation. The simplest molecular model will include terms of

\[
U^{\text{intramolecular}} = \frac{1}{2} \sum_{\text{bonds}} K_b (r_{ij} - r_0)^2 + \frac{1}{2} \sum_{\text{bend angles}} K_\theta (\theta_{ijk} - \theta_0)^2 + \frac{1}{2} \sum_{\text{torsion angles}} \sum_{m} K_{\varphi_m} (1 + \cos(m \varphi_{ijkl} - \gamma_m))
\]

\[ (2.24a,b,c) \]

The bonded interactions (Figure 2.9) involve a separation \( r_{ij} = |r_i - r_j| \) between adjacent pair of atoms, in which equilibrium separation is \( r_0 \). The bonded potential energy may be considered harmonic, with a stiff linear spring between particle i and j, with a spring constant \( K_b \).

The bond-angle interactions are three particle interactions in which bend angles \( \theta_{ijk} \) are between successive bond vectors such as \( r_i - r_j \) and \( r_j - r_k \), which act as a torsion spring with an angular displacement from the equilibrium value \( \theta_0 \) and a spring constant \( K_\theta \). For \( \hat{F} = \frac{F}{r} \),
\[
\cos(\theta_{ijk}) = \hat{\mathbf{r}}_{ij} \cdot \hat{\mathbf{r}}_{jk} = \frac{\mathbf{r}_{ij} \cdot \mathbf{r}_{jk}}{\sqrt{(\mathbf{r}_{ij} \cdot \mathbf{r}_{ij})(\mathbf{r}_{jk} \cdot \mathbf{r}_{jk})}}
\]  

(2.25)

The dihedral-angle interactions are four particle interactions, in which the torsion angles $\phi_{ijkl}$ are defined for three connected bonds. For $\hat{\mathbf{n}} = \frac{n}{n}$, the unit normal to the plane defined by each pair of bonds,

\[
\cos(\varphi_{ijkl}) = -\hat{\mathbf{n}}_{ijk} \cdot \hat{\mathbf{n}}_{jkl} ; \ n_{ijk} = \mathbf{r}_{ij} \times \mathbf{r}_{jk} ; \ n_{jkl} = \mathbf{r}_{jk} \times \mathbf{r}_{kl}
\]

(2.26)

The torsional potential is expressed in periodic functions of order $m = 1, 2, \ldots$

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure2.9.png}
\caption{Molecular geometry of a chain molecule, illustrating the distance between adjacent pair of atoms $r_{23}$, bend angle $\theta_{234}$ and torsion angle $\varphi_{1234}$. (Drawn from Allen MP. Introduction to molecular dynamics simulation. Computational soft matter: from synthetic polymers to proteins. 2004;23:1-28.)}
\end{figure}

Bonded interactions potential can be specified in a simulation package force field, including strength parameters K.

2.4.3.3 Force Calculations

The forces on each atom can be calculated using the potential energy function $U(r^N)$,
\[ f_i = -\frac{\partial}{\partial r_i} U(r^N) \]  

(2.27)

### 2.4.4 MD Algorithm

For a system of \( N \) atoms of coordinates \( r^N = (r_1, r_2, \ldots, r_N) \), with potential energy \( U(r^N) \), momenta of all atoms is given by \( p^N = (p_1, p_2, \ldots, p_N) \), with kinetic energy:

\[
K(p^N) = \sum_{i=1}^{N} \frac{|p_i|^2}{2m_i}
\]

(2.28)

Thus, the total energy (Hamiltonian) is given by:

\[
H = U(r^N) + K(p^N)
\]

(2.29)

From Newton’s classical equation of motion,

\[
\dot{r}_i = \frac{p_i}{m_i}; \quad \dot{p}_i = f_i
\]

(2.30)

a system of coupled differential equations is obtained, which can be solved by step-by-step numerical integration. The numerical algorithm for solving these ODE involves an approximate method that can approach the right solution in the limit of extremely small step size. Also, force calculations are expensive and difficult to program (84), involving sums over pair of atoms. Along with coordinate calculations for each atom, dynamical properties of the system are outputted over times comparable to correlation times of a specific property. The trajectory generated by the system time evolution must sample the constant energy hypersurface, in which all points on this surface correspond to the same momenta (85) for time runs larger than time correlation of particular property of interest. To fulfil these conditions, the time step of the algorithm must be sufficiently large, such that the phase space is sampled at a fast rate. A low order algorithm would allow for time step increase while preserving energy conservation. Verlet algorithm most closely concurs with these requirements.
2.4.4.1 Verlet Algorithm

For an atom i, momentum and position steps in Verlet algorithm can be written as:

\[ p_i \left( t + \frac{1}{2} \delta t \right) = p_i(t) + \frac{1}{2} \delta f_i(t) \]  
(2.31a)

\[ r_i(t + \delta t) = r_i + \delta p_i(t + \frac{1}{2} \delta t) \]  
(2.31b)

\[ p_i(t + \delta t) = p_i \left( t + \frac{1}{2} \delta t \right) + \frac{1}{2} \delta f_i(t + \delta t) \]  
(2.31c)

Given a time step \( \delta t \), for the position of an atom, \( r_i(t + \delta t) \) in equation (2.30b), force is calculated for the same time step, \( \delta t \), \( f_i(t + \delta t) \). Verlet algorithm is time reversible, symplectic (provides enhanced stability over long-time simulations) (86), low order, it calculates one force evaluation per step and it is easy to program.

2.4.4.2 Constraints

The constrains are defined in Lagrangian or Hamiltonian mechanics (87,88). Bonds in the potential energy function are constrained to have fixed length. For a fixed bond length \( b \) between two atoms 1 and 2, the constraint equation and it time derivative can be written as:

\[ \chi(r_1, r_2) = (r_1 - r_2) \cdot (r_1 - r_2) - b^2 = 0 \]  
(2.32a)

\[ \dot{\chi}(r_1, r_2) = (v_1 - v_2) \cdot (r_1 - r_2) = 0 \]  
(2.32b)

For each atom i, Newton’s equation with Lagrange multiplier, \( \Lambda \), can be written as

\[ m_i \ddot{r}_i = f_i + \Lambda g_i \]  
(2.33)

where
\[ g_1 = \frac{\partial \chi}{\partial \mathbf{r}_1} = -2(\mathbf{r}_1 - \mathbf{r}_2); \quad g_2 = \frac{\partial \chi}{\partial \mathbf{r}_2} = 2(\mathbf{r}_1 - \mathbf{r}_2) \] 

(2.34)

And exact expression for multiplier \( \Lambda \) can be calculated from the above equations. However, the equations of motion are solved approximately in discrete time steps \( \delta t \), so the constraints will fail during the course of simulation. The solution is to determine the constraining forces such that at the end of each time step constraints are exactly satisfied (89-91). This is the SHAKE version for the original Verlet algorithm, while for velocity the version used is called RATTLE. Combining \((\mathbf{r}_1, \mathbf{r}_2)\) into \(\mathbf{r}\) and \((\mathbf{p}_1, \mathbf{p}_2)\) into \(\mathbf{p}\), Verlet algorithm can be written as:

\[
\mathbf{p}\left(t + \frac{1}{2}\delta t\right) = \mathbf{p}(t) + \frac{1}{2}\delta t\mathbf{f}(t) + \lambda \mathbf{g}(t)
\]

(2.35)

\[
\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \frac{\delta t\mathbf{p}(t + \frac{1}{2}\delta t)}{m}
\]

(2.36)

\[
\mathbf{p}(t + \delta t) = \mathbf{p}\left(t + \frac{1}{2}\delta t\right) + \frac{1}{2}\delta t\mathbf{f}(t + \delta t) + \mu \mathbf{g}(t + \delta t)
\]

(2.37)

where constants \( \lambda \) and \( \mu \) are chosen such that

\[ 0 = \chi(\mathbf{r}(t + \delta t)) \] 

(2.38a)

\[ 0 = \dot{\chi}(\mathbf{r}(t + \delta t), \mathbf{p}(t + \delta t)) \] 

(2.38b)

The following unconstrained variables are defined for equation 2.38a:

\[
\bar{\mathbf{p}}\left(t + \frac{1}{2}\delta t\right) = \mathbf{p}(t) + \frac{1}{2}\delta t\mathbf{f}(t); \quad \bar{\mathbf{r}}(t + \delta t) = \mathbf{r}(t) + \frac{\delta t\bar{\mathbf{p}}(t + \frac{1}{2}\delta t)}{m}
\]

(2.39)

that allows solving for Lagrange multiplier \( \lambda \):

\[ \chi(t + \delta t) = \chi(\bar{\mathbf{r}}(t + \delta t)) + \lambda \frac{\delta t\mathbf{g}(t)}{m} = 0 \]

(2.40)

So:
\[
\mathbf{p}\left(t + \frac{1}{2}\delta t\right) = \mathbf{\bar{p}}\left(t + \frac{1}{2}\delta t\right) + \frac{1}{2}\delta t + \lambda \mathbf{g}(t); \quad \mathbf{r}(t + \delta t) = \mathbf{\bar{r}}(t + \delta t) + \frac{\delta t \lambda \mathbf{g}(t)}{m}
\] (2.41)

For equation 2.38b, the following unconstrained variable is defined:

\[
\mathbf{\bar{p}}(t + \delta t) = \mathbf{p}\left(t + \frac{1}{2}\delta t\right) + \frac{1}{2}\delta t \mathbf{f}(t + \delta t)
\] (2.42)

That allows for solving of Lagrange multiplier \(\mu\):

\[
\dot{\chi}(t + \delta t) = \dot{\chi}\left(\mathbf{r}(t + \delta t)\right), \quad \mathbf{\bar{p}}(t + \delta t) + \mu \mathbf{g}(t + \delta t) = 0
\] (2.43)

So:

\[
\mathbf{p}(t + \delta t) = \mathbf{\bar{p}}(\delta t) + \mu \mathbf{g}(t + \delta t)
\] (2.44)

SHAKE algorithm calculates constraint forces \(\lambda g_i\), ensuring that position \(r_i\) satisfy constraint equation (2.32a), while RATTLE algorithm calculate constraint forces \(\mu g_i\), such that momenta \(p_i\) satisfy (2.32b).

### 2.4.4.3 Periodic boundary conditions

For a small sample size in which surface effects can be ignored, it is necessary to use periodic boundary conditions. A non-trivial amount of atoms are always located on the outer faces of the unit cell, affecting measured properties of the system. Replicas of the cell resolves this problem. Minimum-image convention states that each individual particle in the simulation interacts with the closest image of the remaining particles in the system (92), for an intermediate potential range. If an atom leaves the unit cell, it can be replaced by the incoming image (Figure 2.10).
Figure 2.10. Periodic boundary conditions. As a particle moves out of the simulation box, an image particle moves in to replace it. In calculating particle interactions within the cutoff range, both real and image neighbors are included (Drawn from Allen MP. Introduction to molecular dynamics simulation. Computational soft matter: from synthetic polymers to proteins. 2004;23:1-28.)

2.4.4.4 Neighbor lists

Non-bonded potentials calculations involve a large number of pairwise calculations. For each atom i, looping over all other all other atoms j, minimum image separation $r_{ij}$ are calculated. For short range interaction potentials, $v(r_{ij}) = 0$ if $r_{ij} > r_{cut}$, in which $v(r_{cut})$ is the cutoff potential. In this case force calculations are skipped, and potential calculations proceed for the next atom j. In an N atom system there are $n = N(N - 1)/2$ distinct atom pairs, and the time to process all pair separations is proportional with n, which is a lot of time, adding to the computations of at least $r_{ij}^2$. Time can be saved by using the list of nearby pairs of atoms. The shape of the cutoff potential is that of a sphere if radius $r_{cut}$ surrounding the atom, while $r_{list}$ is in the exterior of $r_{cut}$. The list of all neighbors of each atom for which the pair separation list is constructed during the first step of a simulation. Only pairs in the list are used for the force calculations over the next steps of the simulations. The list is reconstructed periodically, before any unlisted pairs came within interaction range. The cutoff distance $r_{cut}$ can be obtained by experimentation in terms of computation time.
Larger than 1000 atoms systems require a cubic or non-cubic simulation box (Figure 2.12). Cells are chosen such that each side of the cell is greater than the potential cutoff distance. Given that there are atoms lists in each cell, the neighbor atoms in the same cell and nearest neighbor cells can be sufficiently fast. First, atoms are sorted in their cells at every step, followed by pair force calculations.

Figure 2.11. The Verlet list on its construction, later, and too late. The potential cutoff range (solid circle), and the list range (dashed circle), are indicated. The list must be reconstructed before particles originally outside the list range (black) have penetrated the potential cutoff sphere. (Drawn from Allen MP. Introduction to molecular dynamics simulation. Computational soft matter: from synthetic polymers to proteins. 2004; 23: 1-28.)

Figure 2.12. The cell structure. The potential cutoff range is indicated. In searching for neighbors of an atom, it is only necessary to examine the atom’s own cell, and its nearest-neighbor cells (shaded). (Drawn from Allen MP. Introduction to molecular dynamics simulation. Computational soft matter: from synthetic polymers to proteins. 2004; 23: 1-28.)
2.4.5 Time dependence

In computer simulations, time dependent statistical mechanics has three areas of importance: a) molecular dynamics algorithms have advanced significantly in the past years as a result of applying operator approach to classical mechanics; b) connection with experiment by elucidation of equilibrium time correlation functions in relation with dynamical properties of the simulation system, particularly transport coefficients; c) fast development of non-equilibrium molecular dynamics in the past decade, such as the connections between dynamical algorithm, dissipation, chaos and fractal geometry.

Time evolution of a mechanical system outlines a phase space trajectory through the multi-dimensional space, representing states compatible with a given initial set of coordinates. Shape analysis of the phase diagram can reveal system characteristics and properties unavailable by other means. A phase space can be defined by many dimensions, such as positions, momenta, etc. For a classical system, any set of generalized coordinates for position and conjugate generalized momenta (87) defines coordinates in phase space.

2.4.5.1 Liouville’s Theorem

For a classical system of particles, the complete state of a particle is given by its coordinates and momenta. In three-dimensional space, for example, there are three position coordinates and three conjugate momenta. A point in a six dimensional phase space represents the state of a particle. The equations of motion provide coordinates for the particle at any time, past or future. Particles follow individual, non-intersecting paths through the six dimensional phase space.

A large system of particles can be described by one point in phase space per particle. An ensemble of particles can be defined by a density or phase-space distribution function, \( \rho (q_i, p_i) \), in
terms of position and momenta. Liouville’s theorem stats that the local density of points in the system traveling through phase space is constant, \( \frac{d\rho}{dt} = 0 \). The point at which the density is calculated moves with the particle, whose coordinate is highly correlates with its momentum such that the density in phase space stays constant.

The Liouville equation describes the time evolution of the phase space distribution function, \( \rho(r_N, p_N, t) \) for a classical system and it can be derived from Hamiltonian mechanics:

\[
\frac{\partial \rho}{\partial t} = -\left\{ \sum_i \dot{r}_i \cdot \frac{\partial}{\partial r_i} + \dot{p}_i \cdot \frac{\partial}{\partial p_i} \right\} \rho \equiv -iL\rho
\]  

(2.45)

Where \( iL \) is the Liouville operator,

\[
\sum_i \dot{r}_i \frac{\partial}{\partial r_i} + \dot{p}_i \frac{\partial}{\partial p_i}
\]  

(2.46)

With the solution:

\[
\rho(t) = e^{-iLt} \rho(0)
\]  

(2.47)

where the exponential operator \( e^{-iLt} \) is propagator.

Applied to an equilibrium ensemble, Liouville theorem implies that distribution function \( \rho \) is constant, \( \frac{\partial \rho}{\partial t} = 0 \). Given any phase space volume element, the rate of incoming particles equals the rate of outgoing particles, such that \( \rho \) is a function of the constants of motion.

Deriving Liouville’s equation (93) implies calculation of the rate of change of number of particles in an infinitesimal hypercube in phase space. Given face dimensions \( q_k \) and \( q_j, j \neq k \), the cube face can be placed perpendicular to \( q_k \), such that the flow of particles through the face is:

\[
\rho \dot{q}_k dq_k dp_k dq_j dp_j
\]  

(2.48)

The flow of particle through this face and its opposite determines the rate of change of number of particles in the hypercube:
\[
\frac{\partial N}{\partial t} = - \frac{\partial (\rho \dot{q}_k)}{\partial q_k} \prod_j dq_j dp_j
\]  
\hspace{1cm} (2.49)

For all hypercub’s faces, the net flow is:
\[
\frac{\partial N}{\partial t} = - \sum_k \left[ \frac{\partial (\rho \dot{q}_k)}{\partial q_k} + \frac{\partial (\rho \dot{p}_k)}{\partial p_k} \right] \prod_j dq_j dp_j
\]  
\hspace{1cm} (2.50)

Density rate of change is obtained by dividing the net flow by the hyprecub’s volume:
\[
\frac{\partial \rho}{\partial t} = - \sum_k \left[ \frac{\partial (\rho \dot{q}_k)}{\partial q_k} + \frac{\partial (\rho \dot{p}_k)}{\partial p_k} \right]
\]  
\hspace{1cm} (2.51)

Adding the rate of change due to one moving particle,
\[
\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \sum_k \left[ \frac{\partial \rho}{\partial q_k} \dot{q}_k + \frac{\partial \rho}{\partial p_k} \dot{p}_k \right]
\]  
\hspace{1cm} (2.52)

\[
\frac{d\rho}{dt} = - \sum_k \left[ \frac{\partial (\rho \dot{q}_k)}{\partial q_k} + \frac{\partial (\rho \dot{p}_k)}{\partial p_k} \right] + \sum_k \left[ \frac{\partial \rho}{\partial q_k} \dot{q}_k + \frac{\partial \rho}{\partial p_k} \dot{p}_k \right]
\]  
\hspace{1cm} (2.53)

\[
\frac{d\rho}{dt} = \sum_k \left[ - \frac{\partial \rho}{\partial q_k} \dot{q}_k - \frac{\partial \rho}{\partial p_k} \dot{p}_k \rho - \frac{\partial \rho}{\partial q_k} \dot{p}_k - \frac{\partial \rho}{\partial p_k} \dot{q}_k + \frac{\partial \rho}{\partial p_k} \dot{p}_k \right]
\]  
\hspace{1cm} (2.54)

\[
\frac{d\rho}{dt} = \sum_k \left[ - \frac{\partial \dot{q}_k}{\dot{q}_k} - \frac{\partial \dot{p}_k}{\dot{p}_k} \right] \rho
\]  
\hspace{1cm} (2.55)

Given Hamilton’s equations of motion,
\[
\dot{q}_i = \frac{\partial H}{\partial \dot{p}_i} ; \dot{p}_i = - \frac{\partial H}{\partial q_i}
\]  
\hspace{1cm} (2.56)

Density rate of change can be expressed as
\[ \frac{d\rho}{dt} = \sum_k \left[ -\frac{\partial^2 H}{\partial q_k \partial p_k} + \frac{\partial^2 H}{\partial p_k \partial q_k} \right] \rho = 0 \] (2.57)

Density in phase space of a system of particle cannot be changed, and it is independent of external forces.

### 2.4.5.2 Propagators and the Verlet Algorithm

All particles in the system must be assigned initial position, velocities and accelerations by integration of the equations of motions of particles, carried out numerically by Verlet algorithm.

For velocities, Verlet algorithm can be obtained using standard approximate decomposition of the Liouville operator, assuring reversibility and volume conservation in phase space. The simulation time \( t \) is divided in a large number \( n \) of small time steps, \( \delta t \to 0 \), \( (\delta t = t/n) \). Equations of motion are approximated at each time step. Liouville operator \( e^{iLt} \) can be expressed as:

\[ e^{iLt} = (e^{iL\delta t})^n + \mathcal{O}(n\delta^3) \] (2.58)

Operator \( iL \) can be written as a sum of two non-commuting parts (94),

\[ iL = iL_p + iL_r \] (2.59)

\[ iL_p = \sum_i \dot{p}_i \cdot \frac{\partial}{\partial p_i} = \sum_i f_i \cdot \frac{\partial}{\partial p_i} \] (2.60)

\[ iL_r = \sum_i \dot{r}_i \cdot \frac{\partial}{\partial r_i} = \sum_i \frac{p_i}{m_i} \cdot \frac{\partial}{\partial r_i} \] (2.61)

Leading to the following approximation, exact in the limit of \( \delta t \to 0 \):

\[ e^{iL\delta t} = e^{(iL_p+iL_r)\delta t} \approx e^{iL_p\delta t} e^{iL_r\delta t} e^{\frac{iL_p\delta t}{2}} \] (2.62)
This approximate equation skips kinetic and potential energy part of the Hamiltonian, one at a time, providing velocity Verlet algorithm in three successive steps. Given a dynamical variable \( A(r^N, p^N) \), operators \( iL_p \) and \( iL_r \) applied to \( A \) advance separately position and momenta of each particle (95):

\[
e^{iL_r \delta t} A(\mathbf{r}, \mathbf{p}) = A(\mathbf{r} + \mathbf{p} \frac{\delta t}{m}, \mathbf{p}) \tag{2.63a}
\]

\[
e^{iL_p \delta t} A(\mathbf{r}, \mathbf{p}) = A(\mathbf{r}, \mathbf{p} + f \delta t) \tag{2.63b}
\]

Verlet algorithm provides approximate trajectories in which Hamiltonian is not conserved. However, when Hamiltonian is changed by a small amount which disappears as \( \delta t \to 0 \), this Hamiltonian is conserved such that the system in phase space stays close to a constant energy hypersurface. This is the stability condition, important in molecular dynamics for sampling constant energy states.

### 2.4.6 Rigid Molecule Rotation

In the molecular model for simulation of polymers, non-spherical rigid molecules must be included, implying calculations of intermolecular toques aside forces and using classical rotational dynamics equations of motion. Intermolecular forces will be expressed as center of mass forces, while torques will be around the center of mass. For two molecules a and b, with center of mass positions \( \mathbf{r}_a \) and \( \mathbf{r}_b \), with intermolecular vector \( \mathbf{r}_{ab} = \mathbf{r}_a - \mathbf{r}_b \), for any sites i in a and j in b, the interaction potential can be expressed as:

\[
v_{ab} = \sum_{i \in a} \sum_{j \in b} v(r_{ij}) \tag{2.64}
\]

Forces and torques are computed as follows:
\[ F_{ab} = \sum_{i \in a} \sum_{j \in b} f_{ij} = -F_{ba} \]  
(2.65)

\[ \tau_{ab} = \sum_{i \in a} \sum_{j \in b} r_{ia} \times f_{ij} \]  
(2.66)

\[ \tau_{ba} = \sum_{i \in a} \sum_{j \in b} r_{jb} \times f_{ji} \]  
(2.67)

Where \( F_{ab} \) is the force on \( a \) due to \( b \), \( F_{ba} \) is the force on \( b \) due to \( a \), \( \tau_{ab} \) is the torque on \( a \) due to \( b \), \( \tau_{ba} \) is the torque on \( b \) due to \( a \), \( r_{ia} = r_i - r_a \) is site \( i \) position relative to molecule \( a \) center and \( r_{jb} = r_j - r_b \) is the site \( j \) position relative to molecule \( b \) center. The system is in rotational equilibrium as long as forces \( f_{ij} = -f_{ji} \) and are a parallel to vectors \( r_{ij} \). Potential energy \( v_{ab} \) is invariant to a rotation of coordinates, resulting in conservation of local angular momentum:

\[ \tau_{ab} + \tau_{ba} + r_{ab} \times f_{ab} = 0 \]  
(2.68)

### 2.4.7 Molecular dynamics in different ensembles

In molecular dynamics simulations microscopic information, such as atomic positions and velocities, is utilized in determination of macroscopic observables that characterize large systems of particles, defined by statistical mechanics quantities of pressure, energy, volume, temperature, heat capacity, etc.

The state of a classical system of particles is completely specified by the measurements of three-position components \( q \), and three momentum components \( p \), a total of six quantities, composing a six dimensional phase space for a system of \( N \) particles at temperature \( T \) and pressure \( P \).

For a particular observable \( A \), as a function of position and momenta, \( A = A(r^N, p^N) \), the ensemble average is defined as:
\[ < A > = \int \int dr^N dp^N A(r^N, p^N) \rho(r^N, p^N) \quad (2.69) \]

in which \( \rho(r^N, p^N) \) is the probability density over phase space, given by:

\[ \rho(r^N, p^N) = \frac{1}{Z} \exp \left( -\frac{H(r^N, p^N)}{k_B T} \right) \quad (2.70) \]

for a temperature \( T \), where \( k_B \) is Boltzmann’s constant, \( H \) is Hamiltonian and \( Z \) is the partition functions of the ensemble.

A partition function describes the statistical properties of an ensemble of particles in thermodynamic equilibrium, as dependent on state variables, such as temperature and volume. Thermodynamic quantities, such as total energy, free energy, entropy and pressure can be expresses or derived from the partition functions.

A partition function can be defined according to the type of statistical ensemble it represents, which corresponds to a particular type of free energy, describing how probability is partitioned among the available microstates compatible with the constraints imposed on the ensemble. The general, a partition function can be defined as:

\[ Z = \int \int dp^N dr^N \exp \left[ -\frac{H(p^N, r^N)}{k_B T} \right] \quad (2.71) \]

Ensembles are subject to various macroscopic constrains, with specific statistical characteristics. The most important types of ensembles are described below.

Microcanonical ensemble represents possible states of a mechanical system with exactly specified total energy (96). The system cannot exchange energy or particles with environment (i.e., the system is isolated). The system's energy (\( E \)), composition, or number of particles (\( N \)), volume (\( V \)), and shape constant in all possible states. Such an ensemble (NVE) has fluctuations temperature (\( T \)) and pressure (\( P \)), where temperature can be calculated from kinetic energy:
and pressure is calculated via Clausius virial theorem (97):

$$P = \frac{Nk_B T}{V} - \frac{1}{3V} \sum_i \sum_{j>i} r_{ij} \frac{\partial U_{ij}}{\partial r_{ij}}$$

where \(V\) is the volume, \(N\) is the number of particles, \(T\) is the temperature of the system, \(k_B\) is Boltzmann’s constants, \(r_{ij}\) is the vector joining particle \(i\) and \(j\), and \(U_{ij}\) is the interatomic potential between \(i\) and \(j\).

In a microcanonical ensemble, every microstate is assigned an equal probability, given that its energy falls within a certain range \(E\). As all probabilities add up to 1, the probability \(p\) is the inverse of the number of microstates \(\Omega(E)\) within energy range \(E\), \(p = 1/\Omega(E)\), in which all microstates are equally likely to occur. This is the macroscopic state of maximal entropy, \(S\):

$$S = -k_B \sum_{i=1}^{\Omega(E)} \left\{ \frac{1}{\Omega(E)} \ln \left( \frac{1}{\Omega(E)} \right) \right\} = k_B \ln(\Omega(E))$$

Canonical ensemble represents a system in thermal equilibrium that can exchange energy with a heat bath, such that its possible states have different total energy, but the same temperature and volume (NVT ensemble).

According to Boltzmann’s distribution, the probability that each microstate has energy \(E_i\), can be expressed in terms of canonical partition function,

$$Z = \sum_{i=1}^{i_{\text{max}}} e^{-\beta E_i} = e^{-\beta F}$$

where \(\beta = 1/k_B T\) and \(F\) is the Helmholtz free energy of the system, \(F = U - TS\), in which \(U\) is the internal energy of the system \(S\) is the entropy and \(T\) is the temperature.

Thus, probability is given by:
\[ p_i = \frac{e^{-\beta E_i}}{Z} \]  

(2.76)

Partition function can be used to calculate averages of any microscopic property of the system, which can be related to macroscopic quantities.

Average total energy \( E \) can be calculated as derivative of \( Z \) with respect to \( T \):

\[ < E > = -\frac{1}{Z} \frac{dZ}{d\beta} \]  

(2.77)

and internal energy \( U \) can be expressed as

\[ U = -\frac{d\ln Z}{d\beta} \]  

(2.78)

Gibbs entropy is:

\[ S = -k_B \sum_i p_i \ln(p_i) = k_B \sum_i \frac{e^{-\beta E_i}}{Z}(\beta E_i + \ln Z) = k_B (\ln Z + \beta U) \]  

(2.79)

implying that

\[ -\frac{\ln(Z)}{\beta} = U - TS = F \]  

(2.80)

Expressions for internal energy \( U \), entropy \( S \) and free energy \( F \), allow for deriving other thermodynamic quantities, such as pressure \( P \) and heat capacity \( C \) and chemical potential \( \mu \).

Helmholts free energy, \( F = -\frac{\ln Z}{\beta} \)  

(2.81)

Internal energy, \( U = -\left(\frac{\partial \ln Z}{\partial \beta}\right)_{N,V} \)  

(2.82)

Pressure, \( P = -\left(\frac{\partial F}{\partial V}\right)_{N,T} = \frac{1}{\beta} \left(\frac{\partial \ln Z}{\partial V}\right)_{N,T} \)  

(2.83)

Entropy, \( S = k_B (\ln Z + \beta U) \)  

(2.84)

Gibbs free energy, \( G = F + PV = -\frac{\ln Z}{\beta} + \frac{V}{\beta} \left(\frac{\partial \ln Z}{\partial V}\right)_{N,T} \)  

(2.85)
Enthapy, \( H = U + PV \) \hspace{1cm} (2.86)

Constant volume heat capacity, \( C_V = \left( \frac{\partial U}{\partial T} \right)_{N,V} \) \hspace{1cm} (2.87)

Constant pressure heat capacity, \( C_P = \left( \frac{\partial H}{\partial T} \right)_{N,P} \) \hspace{1cm} (2.88)

Chemical potential, \( \mu_i = -\frac{1}{\beta} \left( \frac{\partial \ln Z_i}{N_i} \right) T,V,N \) \hspace{1cm} (2.89)

Grand canonical ensemble represents a system of particles maintained in thermal and chemical equilibrium. The system can exchange energy and particles with a reservoir (i.e. the system is open), such that system’s state can differ in total energy and number of particles. The probability assigned to each distinct microstate by grand canonical ensemble can be expressed as:

\[
p = e^{\frac{\Phi + \mu N - E}{k_B T}}
\]

where \( N \) is the number of particles in the microstate, \( E \) is the total energy and \( \Phi \) is the grand potential, constant for the ensemble. If different \( \mu, V, T \) are selected, \( p \) and \( \Phi \) will be different. The grand potentials, \( \Phi \) provides a normalization factor for the probability distributions (adding up to one over all microstates). Ensemble averages can be derived from \( \Phi \) (\( \mu, V, T \)). The system’s volume (\( V \)) and temperature \( e \) (\( T \)) are kept constant. Grand canonical partition function is defined as:

\[
\Xi(V, T, \mu) = \sum_i \exp \left( \beta \left[ \sum_{j=1}^{N} \mu_j N_{ij} - E_i \right] \right)
\]

where \( N_{ij} \) is the number of the \( j \)th particle species at \( i \)th configuration.

\[
\text{Internal energy, } U = \sum_i E_i \frac{\exp(-\beta(E_i - \sum_j \mu_j N_{ij}))}{\Xi}
\]

\hspace{1cm} (2.92)
Number of particles, \( N_j = \sum_i N_{ij} \exp(-\beta(E_i - \sum_j \mu_j N_{ij})) \) (2.93)

Grand potential, \( \Phi = -\frac{\ln \Xi}{\beta} \) (2.94)

Internal Energy, \( U = -\left( \frac{\partial \ln \Xi}{\partial \beta} \right)_{\mu} + \sum_i \frac{\mu_i}{\beta} \left( \frac{\partial \ln \Xi}{\partial \mu_i} \right)_{\beta} \) (2.95)

Number of particles, \( N_i = \frac{1}{\beta} \left( \frac{\partial \ln \Xi}{\partial \mu_i} \right)_{\beta} \) (2.96)

Entropy, \( S = k_B \left( \ln \Xi + \beta U - \beta \sum_i \mu_i N_i \right) \) (2.97)

Helmholtz free energy, \( F = \Phi + \sum_i \mu_i N_i = -\frac{\ln \Xi}{\beta} + \sum_i \frac{\mu_i}{\beta} \left( \frac{\partial \ln \Xi}{\partial \mu_i} \right)_{\beta} \) (2.98)

In isothermal–isobaric ensemble, the number of molecules (N), pressure (P) and temperature (T) are held constant (NPT ensemble) (98), whose partition function can be written as a weights sum of canonical partition function, \( Z \):

\[
\Delta(N, P, T) = \sum_i \sum_V Z(N, V, T) \exp(-\beta PV)
= \sum_i \sum_V \Omega(N, V, E_i) \exp(-\beta E_i) \exp(-\beta PV)
\]

Gibbs free energy, \( G(N, P, T) = -k_B T \ln \Delta = F + PV \) (2.100)

where \( F \) is Helmholtz free energy.

2.4.8. Steps in molecular dynamics simulations

A simulation run has parameters that’s specify its conditions, such as initial temperature, number of particles, time step, which are read first. Then initialization of the system consists of
reading initial positions and velocities selected such that they are compatible to the simulated structure. To generate a Maxwell-Boltzmann velocity distribution, a random number generator can be used such that the normal distribution has zero mean and unit variance, and scaling the resulting velocities (multiplication by \(\sqrt{\frac{k_B T}{m_i}}\)) to adjust the mean kinetic energy to the desired value. Also, velocities must be shifted so there is no overall momentum. An ensemble with such a set of initial velocities satisfy the condition for average kinetic energy, \(<KE> = \frac{Nk_BT}{2}\). However, the distribution will change during simulation, particularly if initial configuration is far from equilibrium. If the system is equilibrated, of thermodynamic quantities such as potential energy, temperature, etc. as function of times should fluctuate around their averages, without drifting. Different initial coordinates and velocities should produce the same thermodynamic quantities after equilibrium.

Next comes force calculations and integrating Newton’s equation of motion. For close pair of particles that interact, forces on these particles will be computed, along with contributions to potential energy. For example, x-component of the force can be calculated from potential energy:

\[
F_x(r) = -\frac{\partial U(r)}{\partial x} = -\frac{x \frac{\partial U(r)}{\partial r}}{r} \tag{2.101}
\]

The equation of motion will be integrated using the algorithms described above (2.31).

Following after these steps, are calculations for thermodynamic properties obtained from ensemble averages.
2.4.9 Free energy methods

Free energy is the principal thermodynamic quantity, to the extent that thermodynamics is applied to the analysis of physical equilibrium (99), and it is defined as Helmholtz or Gibbs free energy:

\[
\text{Gibbs free energy } G(N, P, T) = U - TS + PV, \quad G = \sum_{N} \mu_{i}N_{i} \quad (2.102)
\]

Helmholtz free energy: \( F(N,V,T) = U-TS \) \( (2.103) \)

Where \( U \) is internal energy, \( T \) is temperature, \( S \) is entropy, \( P \) is pressure, \( V \) is volume, \( N \) is the number of particle in the system and \( \mu \) is chemical potential.

Free energy techniques are applied to study phase transitions, critical phenomena and other transformations. Free energy minimization, applied with specific constrains (reaction coordinates), is basic to understanding, modeling and predicting phase equilibria and stability of different phenomena. A system is stable if free energy is lowest at predetermined conditions.

Free energy methods consist of two overlapping concepts. First is the function of state (\( A, G, \Omega \ldots \)) and its change during a process: \( \Delta A, \Delta G, \Delta \Omega \ldots \) Second is calculating free energy profiles with respect to an internal constraint (reaction coordinate).

Without an appropriate reference state, calculating absolute free energies is impossible. However, computational methods can be used to calculate relative free energies.

Helmholtz free energy can be calculated by evaluating the following integral:

\[
A = k_{B}T \left( \iint dp^{N}dr^{N}\exp[-\beta H(p^{N}, r^{N})] \right) \quad (2.104)
\]

where \( H = H(p_{1}, p_{2}, ..., p_{N}, r_{1}, r_{2}, ... r_{N}) \) is the Hamiltonian of a system of \( N \) interacting particles. Since high energy regions are not adequately sampled, it is very difficult to evaluate such integral using molecular dynamics (or Monte Carlo simulations). Accurate calculations of absolute free energy are not possible due to insufficient sampling obtained during simulations. However, free
energy differences can be used to calculate ensemble averages. Since free energy is a state function, given two states X and Y, the difference in energy between these two states is:

\[ \Delta A = -k_B T \ln \left( \frac{\iint dp^N dr^N \exp[-\beta H_Y(p^N, r^N)]}{\iint dp^N dr^N \exp[-\beta H_X(p^N, r^N)]} \right) \] (2.105)

Multiplying the numerator by a factor \( \exp(\beta H_x) \exp(-\beta H_x) = 1 \), \( \Delta A \) can be expressed as:

\[ \Delta A = -k_B T \ln \left( \frac{\iint dp^N dr^N \exp[-\beta(H_Y - H_x)] \exp[-\beta H_X]}{\iint dp^N dr^N \exp[-\beta H_X]} \right) \] (2.106)

This is an ensemble average taken over state X or Y:

\[ \Delta A = -k_B T \ln(\exp[-\beta(H_Y - H_x)])_X, \, \Delta A = -k_B T \ln(\exp[-\beta(H_X - H_Y)])_Y, \] (2.107)

Which can be evaluated by running a simulation for state X or Y, then collect statistics.

Most used free energy methods include thermodynamics integration, free energy perturbation, umbrella sampling, potential mean force, metadynamics.

### 2.4.9.1 Metadynamics

Metadynamics is an algorithm used to compute free energy and other state functions of a system, where ergodicity is hindered by the form of the system's energy landscape. It consists in the inserting of memory into an enhanced sampling (100) by local elevation (101) and conformational flooding (102) such resampling of previous states is avoided and computational resources are redirected to a wider investigation of free energy landscape. Because metadynamics can search through all energy landscape, Darve and Phorriile (103), suggested that this method can be applied to study biological systems such as protein folding (104), molecular docking (105), phase transitions (106,107) and conformational changes (108). Metadynamics does not depend on an accurate description of the explored potential energy surface. Possible misevaluated conformations can be recalculated, with leveled error (100). In order to produce an accurate
description of the free energy surface, the system analyzed with metadynamics must be relatively small. Also using a small set of collective coordinate is essential. These features of metadynamics method allow for quickly obtaining information about the topology of the investigated free energy surface.

Metadynamics method uses an artificial energy term (external potential) to obtain potential mean force \( \text{PMF} = A(\xi) \) from previous simulations on collective variables (colvars), representing differentiable functions of atomic Cartesian coordinate for \( N \) atoms, \( \xi = (\xi_1, \xi_2, \ldots, \xi_N) \). Most commonly used colvars include radius of gyration of a group of atoms, center-of-mass distance between two groups, angle between three groups, coordination number between two groups, orientation from reference coordinates, etc. The history dependent potential \( V_{\text{meta}} \) acting on current values of colvars \( \xi \) is defined as:

\[
V_{\text{meta}}(\xi(t)) = \sum_{t' < t} W \prod_{i=1}^{N} \exp \left( -\frac{(\xi_i(t) - \xi_i(t'))^2}{2\sigma_{\xi_i}^2} \right)
\]

\( V_{\text{meta}} \) is obtained by summing up \( N \)-dimensional repulsive Gaussian energy “hills” of constant valued \( W \), centered at previous configurations \( (\xi(\delta t), \xi(2\delta t), \ldots) \).

The effective pmf \( \bar{A}(\xi) \), which corresponds to the nearest minimum that the system progresses to, represents current pmf \( A(\xi) \) plus the metadynamics potential \( V_{\text{meta}}(\xi) \). The probability of observing a configuration \( \xi^* \) is defined by:

\[
\xi^* \sim \frac{\bar{A}(\xi^*)}{k_B T}
\]

Representing the probability of adding a new Gaussian hill energy to the configuration. For sufficiently long simulations, the sum of Gaussian energy hills cancels out all local minima, and the effective pmf, \( \bar{A}(\xi) \) is constant. Thus \( -V_{\text{meta}}(\xi) \) becomes approximates accurately pmf \( A(\xi) \):
A(\xi) \approx -V_{\text{meta}}(\xi) + K \quad (2.110)

where \( K \) is an additive constant. If collective variables are assumed to include all relevant degrees of freedom, the estimation error is a function of correlation times of colvars, energy hills and energy widths (109). Width of hills \( \sigma_{\xi} \) is chosen as half of the width of the colvar \( \xi \), while the rest of the parameters (height of each hill, \( W \) (kcal/mol), frequency of hill creation, \( \delta t \), etc.) are provided in the metadynamics block. Also, the ratio \( W/\delta t \ll k_B T/\tau_{\xi} \), where \( \tau_{\xi} \) is the longest correlation time of the colvars \( \xi \)'s. The resolution of the calculated pmf is determined by the width of the Gaussian hill, \( \sigma_{\xi} \).

2.4.10 Coarse grained simulations

2.4.10.1 Multiscale simulations

Sometimes larger systems require running simulations for longer times than ab initio methods can achieve (110). For example, complex polymeric and biological systems are difficult to represent accurately and completely. Using simpler models is beneficial because quantities and phenomena of interest can manifest at larger length (hundreds of nanometers to micrometers) and time scales without including all atomistic (AA) levels of detail.

Multiscale modeling describes simulation approaches to systems over different scales of resolution (Figure 2.13).
2.4.10.2 Coarse-grained (CG) models

In the bottom-up approach, information at smaller scales is used for models at larger scales (coarse-grained models). Such models are discrete and derived from atomistic classical models. An effective CG model would be easy to simulate and would reproduce all physical characteristics as all-atomistic one. When designing a CG model, three factors must be considered: a) Number of pseudoatom sites, representing groups of atoms and their connections. b) Energy function for CG model defining interactions between pseudoatoms, reproducing static thermodynamic properties of the atomistic system. c) Effective dynamic equations describing time evolution of the CG system. Removing degrees of freedom from all atomistic system implies removing time scales associated with them. Getting the correct CG dynamics depends on building these times scales back into the dynamical equations.

The goals in building a CG model are to develop a cheaper, easier-to-simulate CG model that has a smoother underlying energy landscape, it equilibrates faster with longer time steps and
maintains correct physical behavior. However, finding a CG model that adequately reproduces the properties of the atomistic system is not guaranteed.

### 2.4.10.3 Pseudoatoms

Pseudoatoms are defined as groups of atoms of common chemistries that can contain functional units. Pseudoatoms can have of various resolutions (Figure 2.14).

![Figure 2.14. A polyalanine molecule composed of 103 atoms is mapped to a 51-bead coarse-grained representation. Each heavy atom with its attached hydrogen(s) is mapped to one coarse-grained bead. Next resolution maps polyalanine’s functional groups: α-amino, α-carboxylic acid, methyl groups. In the highest resolution, amino acid residues beads have the largest size. (Drawn from Shell, M.S. (2009). Coarse graining and multiscale techniques. [Lecture notes]).](image)

A mapping function $M$, takes a set of atomic coordinates in the atomistic system and associates it (maps) to a pseudoatom configuration (bead) in the CG system:

$$M(r) = R$$  \hspace{1cm} (2.111)

where $r = r_A^N_A$ (A - all atomistic system); $R = r_{CG}^N_{CG}$ (CG- coarse-grained system)  \hspace{1cm} (2.112)

For the same CG mapping, there may be more than one all-atomistic configuration.
Pseudoatoms sites are defined as the center-of-mass coordinates of groups of atoms in all-atom representation:

\[ R_l = \left( \sum_{j \in \text{atoms for } l} m_j \vec{r}_j \right) \left( \sum_{j \in \text{atoms for } l} m_j \right)^{-1} \]  \hspace{1cm} (2.113)

CG coordinates that can be expressed as a linear combination of atomic coordinates, where the mapping function becomes a matrix of dimensions $3N_{CG} \times 3N_A$:

\[ \mathbf{R} = \mathbf{M} \mathbf{r} \]  \hspace{1cm} (2.114)

### 2.4.10.4 Finding optimum energy function $U_{CG}$

If interactions between pseudoatoms sites are not specified, or if they are pairwise additive, then $U_{CG}$ is optimal.

For the canonical ensemble using $U_{CG}$, CG and AA configurational probabilities must be equal,

\[ p_{CG}(\mathbf{R}) = p_A(\mathbf{R}) \]  \hspace{1cm} (2.115)

In AA ensemble, $p_{CG}$ is the sum of all $p_{AA}$, where the mapping function is:

\[ p_A(\mathbf{R}) = \int p_A(\mathbf{R}) \delta[\mathbf{M}(\mathbf{r}) - \mathbf{R}] d\mathbf{r} \]  \hspace{1cm} (2.116)

where delta function eliminates AA configurations $\mathbf{r}$ that map to the same CG configuration $\mathbf{R}$.

CG and AA probabilities are defined as:

\[ p_{CG}(\mathbf{R}) = \frac{e^{-\beta U_{CG}(\mathbf{R})}}{Z_{CG}} \quad p_A(\mathbf{r}) = \frac{e^{-\beta U_A(\mathbf{r})}}{Z_A} \]  \hspace{1cm} (2.117)

And now relation 2.109 becomes:

\[ \frac{e^{-\beta U_{CG}(\mathbf{R})}}{Z_{CG}} = \int \frac{e^{-\beta U_A(\mathbf{r})}}{Z_A} \delta[\mathbf{M}(\mathbf{r}) - \mathbf{R}] d\mathbf{r} \]  \hspace{1cm} (2.118)
Taking the logarithm of this expression will give $U_{\text{CG}}$. Partitions function will be absorbed into a normalization constant:

$$U_{\text{CG}}(R) = -k_B T \ln \int e^{-\beta U_A(r)} \delta[\mathbf{M}(r) - \mathbf{R}] \, d\mathbf{r} + \text{const} = F_A(\mathbf{R}) + \text{const} \quad (2.119)$$

where $F_A(\mathbf{R})$ is the multidimensional potential of mean force or free energy surface along reduced degrees of freedom $\mathbf{R}$. CG energy function must reproduce the free energy surface along CG degrees of freedom, such that CG thermodynamic averages should be the same as AA averages, for the same configurational distributions. $U_{\text{CG}}(R)$ would deviate from the correct $F_A(R)$ because modelling CG energy functions involves summation of pairwise terms, in which CG eliminated degrees of freedom are not included in the effective remaining interactions.

### 2.4.10.5 Iterative Boltzmann inversion

$U_{\text{CG}}$ is defined as a sum of pairwise potentials (111,112), based on Henderson’s uniqueness theorem (113), saying that for a given pair radial distribution function $g(r)$, there is a unique underlying pair potential $u(r)$ producing it.

$$U_{\text{CG}}(R) = \sum_{i<j} u_{\text{CG}}(R_{ij}) \quad (2.120)$$

This theorem along with iterative simulations constitutes iterative Boltzmann inversion for determining effective CG potentials, using the following steps: 1) Perform AA simulation. 2) Design CG model of pseudoatoms and compute $g_A(R)$ for each pair by converting AA trajectory frames to CG. 3) Estimate pairwise pseudoatom potential:

$$u_{\text{CG}} = -k_B T \ln g_A(R) \quad (2.121)$$

4) Perform canonical ensemble CG simulation using $U_{\text{CG}}(R)$ potentials and measure pseudoatoms
5) Update pairwise potential:

\[
\begin{align*}
\text{u}_{CG}(R) & \leftarrow \text{u}_{CG}(R) - k_B T \ln \left( \frac{g_A(R)}{g_{CG}(R)} \right) \\
\end{align*}
\]  

(2.122)

If \( g_A(R) = g_{CG}(R) \), then CG pairwise structure is identical to AA pairwise structure. If \( g_A(R) > g_{CG}(R) \), \( u_{CG} \) underestimates correlation between pseudoatoms at \( r \), lowering pairwise CG energy at \( r \) for a favourable interaction. If \( g_A(R) < g_{CG}(R) \), \( u_{CG} \) overestimates correlation between pseudoatoms at \( r \), increasing pairwise CG energy.

6) Repeat (4) CG simulation, continuing until \( U_{CG}(R) \) converges. The final \( u_{CG} \) differs from 2.114 because summed PMF between two pseudoatoms does not produce total energy surface:

\[
F_A(R) \neq \sum_{i<j} -k_B T \ln g_A(R_{ij})
\]

(2.123)

Multidimensional free energy cannot be expressed as a sum of pairwise PMFs, but the above iterating procedure for \( u_{CG} \) produces a better pairwise decomposing of the total energy surface.

2.4.10.6 Force matching

As an alternative to inverse MC force matching was developed (114) to extract classical potential from ab initio calculations, which later (115) was generalized to derive pseudoatom CG models from AA classical models. It also involves an initial AA simulation followed by an iteration such that the average force on pseudoatoms matches the force expected in AA system for the entire trajectory:

\[
< F >_A = F_{CG} = \frac{\partial U_{CG}(R)}{\partial R_i}
\]

(2.124)

for pseudoatoms sites \( i = 1 \ldots N \)

Matching forces requires definition of a least-square objective function:

\[
Y = \sum_{k=1}^{n} \sum_{i=1}^{N} (F_{LA}(R_{A,k}) - F_{i,CG}(R_{A,k}))^2
\]

(2.125)
where \( n \) is the number of AA trajectories and \( N \) is the number of pseudoatom sites. Minimization of \( Y \) with respect to all \( U_{CG} \) parameters determines optimal parameters for \( U_{CG} \).

A typical CG energy function for pseudoatom sites is of form:

\[
u_{ij}(R) = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{R} + f_{ij}(R)
\]

(2.126)

where \( f_{ij}(R) \) is a cubic spline function for atom types I and j, containing knot values of \( f_{ij} \) at predetermined discrete distances \( R \), linear in all knot values and net charge equals that of AA system. Potential is also linear in all combinations of all pseudoatoms charges \( q_i q_j \).

### 2.4.10.7 CG models for polymers

A coarse-grained model of a polymer represents groups of atoms or monomers known as super atoms, between which interactions are defined. The number of degrees of freedom is reduced, such that about ten atoms are collected into one super atom. Thus, calculations of the number of pair interactions are reduced. Coarse-grained potentials are softer than atomistic potentials. Time step can increase to 10-100 fs. A generic polymer model, such as a lattice, continuous or analytical, describes any polymer according to chain topology. Prefactors are used to differentiate between polymers. A coarse-grained model of a polymer must be specific to that polymer. For example, for polystyrene (PS) (Figure 2.15), coarse-graining consists of two types of beads: one bead located at the center of mass of phenyl ring and another centered on the alkyl carbon bonded to the phenyl ring (116).
Figure 2.15. The coarse grain model is shown superimposed over the atomistic model. Backbone type beads are centered on the alkyl carbon bonded to phenyl ring (grey). Side-chain beads are located at the center of mass of the phenyl ring (blue). (Drawn from Hsu DD, Xia W, Arturo SG, Keten S. Thermomechanically consistent and temperature transferable coarse-graining of atactic polystyrene. Macromolecules. 2015 Apr 27;48(9):3057-68.)

Some of the generally chosen features to be reproduced are structural or thermodynamical (117).

Polymer chains in an environment can be described by geometric intramolecular quantities (e.g. distance between two super atoms) or intermolecular (e.g. distance between super atoms of different chains), that can be achieved by the coarse-grained model, according to the desired features. Total potential energy, $U_{CG}$ for a coarse-grained polymer chain can be separated into two parts, bonded and nonbonded (118):

$$U_{CG} = \sum U_{CG}^{bonded} + \sum U_{CG}^{nonbonded}$$ (2.127)

Bonded interactions are the potentials of mean force of the CG degrees of freedom (bond lengths, angles, dihedrals), obtained by sampling distribution functions $p_{CG}$ from atomistic simulations of isolated polymer random walks.

CG procedure consists of four steps: 1. Performing atomistic molecular dynamics (MD) simulations of isolated random walks. 2. Independent conformations for each polymer random walk at a given temperature $T$ will serve to sample histogram distribution functions $PCG$, which are factorized unknown functions of the CG bond lengths $r$, bending angles $\theta$, and dihedral angles $\phi$, for which the internal CG degrees of freedom are uncorrelated:

$$p_{CG}(r, \theta, \phi, T) = p_{CG}(r, T)p_{CG}(\theta, T)p_{CG}(\phi, T)$$ (2.128)

3. CG bonded potentials are provided by inverse Boltzman relations:
\begin{align}
U_{\text{CG}}(r, T) &= -k_B T \ln p_{\text{CG}}(r, T) \\
U_{\text{CG}}(\theta, T) &= -k_B T \ln p_{\text{CG}}(\theta, T) \\
U_{\text{CG}}(\varphi, T) &= -k_B T \ln p_{\text{CG}}(\varphi, T)
\end{align}

(2.129) \quad (2.130) \quad (2.131)

4. The mesoscopic force field is completed by adding a corresponding non-bonded interaction potential (soft repulsive short-range potential) such as one of Lennard–Jones (LJ) potentials with modified exponents:

\[ U_{\text{CG}}^{\text{nonbonded}} = 4\varepsilon \left[ \left( \frac{\sigma_i}{r} \right)^{n_i} - \left( \frac{\sigma_i}{r} \right)^{m_i} \right] + U_{\text{shift}} \]

(2.132)

where \( \sigma_i \) corresponds to one of the nonbonded interactions, \( n_i \) and \( m_i \) are power laws for the soft nonbonded potential; \( U_{\text{shift}} \) is the minimum potential (shifted to zero at cut-off). As CG beads are distinct, all interactions must be described by a repulsive term softer than 12-6 LJ.

For given values of \( \sigma_i \), the exponents \( n_i \) and \( m_i \) can be changed to give optimum agreement between the bead–bead radial distribution function (RDF) of the CG model RDF of the atomistic simulation.

\textbf{2.4.10.8 CG models for phospholipid membranes}

Lipid assemblies are of great importance in biological and nanotechnology applications, featuring various conformations within 10 nm length scales (e.g. bilayers, micelles, vesicles), reason for which CG lipid model is studied to a great extent (119). Compared to the size of the membranes, system sizes that can be treated in an all-atom configuration in reasonable time are limited to short length and time scale of about 10 nm and tens of nanoseconds (120). Longer length
and times scale experiments on membranes can offer some information, but without explaining the phenomenon at molecular level. CG lipid model bridges the gap between atomistic simulations and experiments on longer time and length scale within fair molecular detail. The intermediate resolution of CG MD simulations serves to balance chemical detail with system size (121) (Figure 2.16). The size is sufficient to study membrane morphologies in the process of self-assembly.

Figure 2.16. Visual Molecular Dynamics snapshots of AA (upper right) and CG (lower right) simulations of a 4:1 DOPC/DOPS bilayer. A single CG DOPC lipid (upper left) is showed with beads colored by bead type (gray for hydrocarbons-, pink for glycerol-, brown for phosphate- and blue for choline-type).(Drawn from Bradley R, Radhakrishnan R. Coarse-grained models for protein-cell membrane interactions. Polymers. 2013 Jul 2;5(3):890-936.)

There are several MD CG approaches connecting CG description to the implicit atomic structures of molecules (115,122-127), such as reverse Monte Carlo (112), where CG polymer modeling was extended to lipids; force matching (115), using atomic interaction between CG segments; and CG force fields (125), based on empirical parametrization of CG segment using functional forms
CHAPTER 3
Hydrophobic Collapse of Polystyrene in Water

3.1 Polystyrene Molecular Model

Polystyrene (PS), with chemical formula \((\text{C}_8\text{H}_8)_n\), \((n\) is the number of repeated subunits – monomers) is a glassy hydrophobic polymer of alternating carbon centers with every other carbon attached to a phenyl group. Three distinct structures can describe the configuration of polymer chain: isotactic, in which all substituents are on the same side of polymer chain, syndiotactic, when the polymer chain is composed of alternating groups of opposite configurations, and atactic, when the constituents are in a random combination. A licorice representation of an atactic PS chain is shown in Figure 3.1(a). The sidechain configuration determines physical properties of the polymer in the melt phase such as melting point and moisture resistance \(^1\). While computational models of PS (and additives) possess a long history, \((128-130)\) it is noted that PS has been increasingly well studied by both all-atomistic (AA) \((131)\) and coarse-grained molecular dynamics simulations \((132)\). Numerous molecular-based models for PS have been developed, utilizing both structural \((132)\) and thermodynamic \((133)\) mapping methodologies. A combined structural and thermodynamic approach for force field development has been shown to be transferrable to similar molecular groups over a range of temperatures close to room temperature. For example, the MARTINI force field \((134)\) has been applied to simulations of lipids, proteins, carbon-based allotropes, as well as polymers such as PS \((133)\) and polyethylene oxide \((135)\). The Shinoda–DeVane–Klein (SDK) approach, \((136)\) which provides the parameters for a coarse-grained model of PS as outlined in this paper, has been applied to lipids, \((137)\) proteins, \((138)\) and surfactants, \((139)\) and polymers, such as and polyethylene oxide \((136)\) and polycaprolactone \((140)\). Mesoscale simulation techniques, such as dissipative particle dynamics (DPD) \((141,142)\), are another class
of simulation techniques that have also been used to characterize the microphase behavior of block copolymers that accurately predicts properties of copolymer self-assembly such as micelles or polymersomes (143,144).

Figure 3.1. (a) An atactic PS chain is displayed in a licorice representation, with carbon atoms in grey color and hydrogen in white and (b) CG mapping for SDK-type PS model as applied to an isotactic PS chain. The backbone mapping includes a bead centered between the two carbons. Only one bead type is included for both backbone and ring beads, for four beads in total.

Herein we describe and compare results of AA and coarse-grained (CG) molecular dynamics simulations of single PS chains of varying lengths in a solution of water. We utilize enhanced sampling techniques in molecular dynamics, specifically metadynamics, to improve sampling of the conformation of the polymer chain in both AA and CG molecular dynamics. Metadynamics (100) is a methodology whereby a Gaussian potential is applied with time to the free energy surface until it can overcome free energy minima; it has been extensively applied to accurate sampling of the free energy surfaces of protein configurations, (104) chemical reaction pathways, (145) etc. Advanced sampling techniques such as umbrella sampling (146) have been previously utilized to characterize the collapse of 25-mer hydrophobic polymer chains, finding that the contribution of hydration energy favors compact folded states. Near the folded states, Athawale et al. (147) found that there is a corresponding increase in the fluctuations of local water density, as well as a weak dewetting of the polymer surface. This finding was consistent to previous observations by ten Wolde et al. (148) that during the collapse of a hydrophobic polymer chain the polymer goes
through dry or dewetted transition states. Within this study, calculations of the local hydration water density as a function of chain length suggest that the minimum energy conformation for intermediate chain lengths has a comparably lower hydration water density, indicating that intermediate chain lengths are more likely to form dewetted, globular states, as compared with shorter or longer chain lengths.

3.1.1. All-atomistic simulations

Multiple lengths of atactic PS chains (5, 10, 13, 22, 30, 44 and 58 monomers) were solvated with TIP3P water in boxes of size $100 \times 100 \times 100 \text{ Å}^3$. The simulation boxes containing each chain were initially equilibrated at $T = 300 \text{ K}$ for 50 ns, then the box size was reduced for each PS length and simulated for 2 additional ns. Force field parameters were based on CHARMM27 (149). Simulations were performed with the NAMD2 software package (150,151). A cut-off of 12 Å was used for the calculations of non-bonded interactions, with a LJ switching function of 10 Å. Simulations were performed with a time step of 2 fs and carried out in the NPT ensemble, with temperature held constant by Langevin dynamics with a coupling constant of 5 ps$^{-1}$. Similarly, the pressure was maintained anisotropically by a Langevin piston at 1.01325 bar. After the initial equilibration phase of PS chains, metadynamics was utilized to explore conformational space of each polymer chain, using radius of gyration as a reaction coordinate. Metadynamics simulations were run for 10 ns for each polymer length using the Collective Variables Module (152) in NAMD2. Metadynamics parameters of hill energy (hill weight) was set to 0.01 kcal/mol, hill creation frequency, (i.e. integration steps required for addition of a new hill to the metadynamics potential) was set to 100, and radius of gyration fluctuation amplitude was set to 0.5 Å.
3.1.2 Coarse-grained simulations

To compare with AA molecular dynamics results, coarse-grained (CG) simulations of PS in water were run using parameters based on SDK force field parameters (136). Specific molecular chemistry of PS determines the coarse-grained model structure. We mapped the phenyl rings onto three coarse-grained beads, and the backbone onto a bead that is centered between the two backbone carbons as shown in Figure 3.1(b). CG intramolecular interactions are modelled via harmonic potentials based on AA simulations of a 40 ns simulation of a 10-monomer PS in a 50:50 Molar mixture of THF and water in a box of approximately 800 nm$^3$. For the coarse-grained force field, the intramolecular interactions are modelled via harmonic potentials given by

$$V(r)_{\text{bond}} = K_b(r - r_0)^2$$  
$$V(r)_{\text{angle}} = K_a(\theta - \theta_0)^2$$

Here, $K_b$ and $r_0$ represent the equilibrium force constant and distance for bonds, and $K_a$ and $\theta_0$ represent the equilibrium force constant and angle for angles. These constants are obtained from the respective simulations using an inverse Boltzmann technique, such that

$$U(r)_{\text{bond}} \propto -k_B T \ln \frac{P(r)}{r^2}$$
$$U(r)_{\text{angle}} \propto -k_B T \ln \frac{P(\theta)}{\sin(\theta)}$$

Non-bonded interactions are set by a pairwise additive potential based on one of the Lennard-Jones (LJ) potential (136):

$$U_{9-6}^{\text{LJ}} = \frac{27}{4} \varepsilon \left\{ \left( \frac{\sigma}{r} \right)^9 - \left( \frac{\sigma}{r} \right)^6 \right\}$$
$$U_{12-4}^{\text{LJ}} = \frac{3\sqrt{3}}{2} \varepsilon \left\{ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^4 \right\}$$

Interactions between the PS beads and water were scaled such that the radius of gyration of the polymer as a function of polymer length matched results obtained using AA simulations. Values
for inter- and intramolecular potentials are shown in Tables 3.1–3.3. All system sizes and set-ups are included in Tables 3.4 and 3.5.

<table>
<thead>
<tr>
<th>Atom Type</th>
<th>Atom Type</th>
<th>ε (kcal/mol)</th>
<th>LJ type</th>
<th>σ (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>W</td>
<td>0.895</td>
<td>12-4</td>
<td>4.371</td>
</tr>
<tr>
<td>B</td>
<td>B</td>
<td>0.312</td>
<td>9-6</td>
<td>4.221</td>
</tr>
<tr>
<td>B</td>
<td>W</td>
<td>0.24</td>
<td>9-6</td>
<td>4.296</td>
</tr>
</tbody>
</table>

Table 3.1. Coarse grained parameters for non-bonded interactions in PS-water systems. Pairwise force-field coefficients for bead type ‘B’ represent both backbone and ring type beads. ‘W’, the SDK water model, represents 3 water beads. ε is the depth of the Lennard-Jones potential well and σ is the finite distance at which the inter-particle potential is zero.

<table>
<thead>
<tr>
<th>Bond Type</th>
<th>Atoms</th>
<th>Spring Constant (kcal/mol/Å²)</th>
<th>K_b (kcal/mol/Å²)</th>
<th>Equilibrium Distance r_0 (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>B1-B2</td>
<td>22.1</td>
<td>2.56</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>B1-B4</td>
<td>27.3</td>
<td>2.69</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>B2-B3</td>
<td>50</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>B3-B4</td>
<td>50</td>
<td>2.89</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>B4-B2</td>
<td>50</td>
<td>2.57</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2. Coarse grained parameters for bonded interactions in PS-water systems. The following representations have been used for backbone and ring bonds: B1-B2, backbone-ring; B1-B4, backbone-backbone; B2-B3, B3-B4, B4-B2, ring-ring. K_b and r_0 are the elastic constant and equilibrium bond length of the harmonic bond potential.

<table>
<thead>
<tr>
<th>Angle Type</th>
<th>Atoms</th>
<th>Prefactor K_a (kcal/mol/radian²)</th>
<th>Equilibrium Angle θ_0 (°)</th>
<th>Angle θ_0</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>B1-B2-B3</td>
<td>15.5</td>
<td>83.99561277</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>B2-B3-B4</td>
<td>50</td>
<td>51.50890578</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>B4-B2-B1</td>
<td>17</td>
<td>145.53128</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>B4-B2-B3</td>
<td>50</td>
<td>61.30648408</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>B2-B1-B1</td>
<td>50</td>
<td>87.66254266</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>B1-B1-B1</td>
<td>7.65</td>
<td>173.6062119</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>B3-B4-B2</td>
<td>8.79</td>
<td>65.89014644</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>B1-B1-B2</td>
<td>50</td>
<td>85.94366927</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.3. Coarse grained parameters for bonded interactions in PS-water systems. The following representations have been used for backbone and ring angles: B1-B2-B3, backbone-ring-ring; B2-B3-B4, B4-B2-B3, B3-B4-B2, ring-ring-ring; B4-B2-B1, ring-ring-backbone; B2-B1-B1, ring-backbone-backbone; B1-B1-B1, backbone-backbone-backbone; B1-B1-B2, backbone-backbone-ring. K_a and θ_0 are the elastic constant and equilibrium angle of the harmonic angle potential.
<table>
<thead>
<tr>
<th>PS Length</th>
<th>$&lt;R_g&gt;$ (Å)</th>
<th>$\delta R_g$</th>
<th>Box Size (Å³)</th>
<th>Water Atoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5.11</td>
<td>0.09</td>
<td>26.350x20.440x22.030</td>
<td>876</td>
</tr>
<tr>
<td>8</td>
<td>5.89</td>
<td>0.1</td>
<td>25.798x21.276x30.115</td>
<td>1236</td>
</tr>
<tr>
<td>10</td>
<td>6.47</td>
<td>0.14</td>
<td>30.834x23.732x22.143</td>
<td>1107</td>
</tr>
<tr>
<td>13</td>
<td>6.4</td>
<td>0.05</td>
<td>25.722x28.784x26.256</td>
<td>1374</td>
</tr>
<tr>
<td>17</td>
<td>7.54</td>
<td>0.1</td>
<td>37.700x32.098x41.298</td>
<td>4104</td>
</tr>
<tr>
<td>22</td>
<td>7.87</td>
<td>0.08</td>
<td>38.482x28.238x32.720</td>
<td>2667</td>
</tr>
<tr>
<td>30</td>
<td>9.3</td>
<td>0.13</td>
<td>30.073x37.775x37.236</td>
<td>3123</td>
</tr>
<tr>
<td>44</td>
<td>13.69</td>
<td>0.19</td>
<td>30.967x59.608x44.452</td>
<td>6471</td>
</tr>
<tr>
<td>58</td>
<td>12.57</td>
<td>0.1</td>
<td>64.847x34.640x34.117</td>
<td>5673</td>
</tr>
</tbody>
</table>

Table 3.4. Radius of gyration averaged over the last 0.2 ns of all atomistic non-accelerated MD simulations of equilibrated systems of 5, 8, 10, 13, 17, 22, 30, 44 and 58 single PS chain lengths in water, with standard deviations. Water box sizes and number of water atoms are also included.

<table>
<thead>
<tr>
<th>PS Length</th>
<th>$&lt;R_g&gt;$ (Å)</th>
<th>$\delta R_g$</th>
<th>Box Size (Å³)</th>
<th>Water Atoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5.38</td>
<td>0.23</td>
<td>100x100x100</td>
<td>96447</td>
</tr>
<tr>
<td>8</td>
<td>6.87</td>
<td>0.17</td>
<td>100x100x100</td>
<td>95703</td>
</tr>
<tr>
<td>10</td>
<td>6.12</td>
<td>0.25</td>
<td>100x100x100</td>
<td>96309</td>
</tr>
<tr>
<td>13</td>
<td>6.77</td>
<td>0.13</td>
<td>100x100x100</td>
<td>96195</td>
</tr>
<tr>
<td>17</td>
<td>7.2</td>
<td>0.14</td>
<td>100x100x100</td>
<td>95703</td>
</tr>
<tr>
<td>22</td>
<td>7.93</td>
<td>0.15</td>
<td>100x100x100</td>
<td>95940</td>
</tr>
<tr>
<td>30</td>
<td>8.76</td>
<td>0.08</td>
<td>100x100x100</td>
<td>95703</td>
</tr>
<tr>
<td>44</td>
<td>9.9</td>
<td>0.1</td>
<td>100x100x100</td>
<td>95331</td>
</tr>
<tr>
<td>58</td>
<td>10.96</td>
<td>0.07</td>
<td>100x100x100</td>
<td>94896</td>
</tr>
</tbody>
</table>

Table 3.5. Radius of gyration averaged over last 5 ns of CG non-accelerated MD simulations of equilibrated systems of 5, 8, 10, 13, 17, 22, 30, 44 and 58 single PS chain lengths in water, with standard deviations. Water box sizes and number of water atoms are also included.

The LAMMPS molecular dynamics code (153) was used in both the NVT and NPT ensemble with Nosé-Hoover dynamics (154) and particle-particle-particle mesh (PPPM) (155) for the electrostatic interactions. The pressure and temperature were set to 1 atm and 300 K, respectively. Lennard-Jones pair potential cut-off distance was set to 15 Å. To begin with, the system was run in the canonical (NVT) ensemble for 1 ns at 300 K. Following, the system was run for 200 ns in isothermal-isobaric (NPT) ensemble also at 300 K and 1 bar pressure with a drag factor of 0.01 added to the barostat. A time step of 10 fs was used for all simulations, and the neighbor list for
non-bonded interactions was updated at every step. Next, we continued with simulation runs using metadynamics for 50 ns for all polymer lengths maintained in previous solvation boxes of 100 × 100 × 100 Å3 in the NPT ensemble. For this, we used the LAMMPS USER-COLVARS (156) package, thus obtaining calculations of the potentials of mean force (PMFs) for radius of gyration as a collective variable and using metadynamics as sampling method. Radius of gyration fluctuations amplitude of 0.5 Å and energy hill weight of 0.01 kcal/mol remained the same as in AA simulations.

### 3.2 Collapsed Globular States

#### 3.2.1 All-atomistic simulations

To begin with, we compare the conformations of the collapsed globular states of PS for the polymer chains of a range of lengths from 5 to 58 monomers, using both accelerated (metadynamics) and non-accelerated molecular dynamics simulations. We refer to conventional molecular dynamics as non-accelerated molecular dynamics. We compare conformations of the AA PS chains in water at the minimum energy conformation from metadynamics calculations as shown in Figure 3.2 (a)–(i) to conformations of the PS chains after 50 ns non-accelerated molecular dynamics as shown in Figure 3.3 (a)–(i). The PS polymer chain is displayed in a liquorice representation. The head and tail hydrogens in the hydroxyl group are highlighted using a blue van der Waals representation. Water is shown by a transparent cyan background. As shown in Figure 3.2, it is clearly seen that for polymer lengths of eight monomers or longer the chain collapses into a globular state. The characteristic size of the globular state increases as a function of monomer length. For chains shorter than eight monomers, i.e. five monomers, the polymer chain does not collapse and remains in a more extended conformation. In comparison, as shown in Figure 3.3, for all polymer lengths, the PS approaches a loosely collapsed state, but does not reach the minimum
energy conformation after 50 ns simulation.

Figure 3.2. Snapshots of AA PS in water at the minimum energy configuration calculated using metadynamics for lengths of PS: (a) 5, (b) 8, (c) 10, (d) 13, (e) 17, (f) 22, (g) 30, (f) 44 and (i) 58 monomers. The polymer chain is displayed in a licorice representation. The head and tail hydrogens in the hydroxyl group are highlighted using a blue van der Waals representation. Water is indicated by transparent cyan background.

Figure 3.3. Single PS chains in water of the following lengths: a) 5, b) 8, c) 10, d) 13, e) 17, f) 22, g) 30, h) 44 and i) 58 PS monomers. These snapshots were taken at the end of all atomistic non-accelerated MD simulations of equilibrated PS-water systems. The polymer chain is displayed in a licorice representation. The head and tail hydrogens in the hydroxyl group are highlighted using a blue van der Waals representation. Water is indicated by transparent cyan background.

According to Flory (157), we can write the free energy of a polymer chain as a sum of the
elasticity of the chain and the internal energy. The degree of swelling of the chain can be represented by the parameter $\alpha$ where

$$\alpha^2 = \frac{R_g^2}{R_{g0}^2} \quad (3.7)$$

where $R_g^2$ is the mean square radius of gyration and $R_{g0}^2$ is the equilibrium mean square radius of gyration of a Gaussian chain of length $N$ monomers. The limit $\alpha > 1$ corresponds to a swollen polymer chain in a good solvent, while the limit $\alpha < 1$ corresponds to a compressed polymer chain due to poor solvent. Summing the elastic and internal energy contributions to the free energy in terms of $\alpha$, and then taking the appropriate limits for strong swelling or strong compression, it is found that $R_g \sim N\nu$, where $\nu$ is the scaling exponent. For a good solvent, $\nu = 3/5$ and the polymer behaves like an extended coil. In a bad solvent $\nu = 1/3$, in which the polymer collapses from a coil to a globule to minimize contact with the solvent. For an ideal $\theta$ solvent $\nu = 1/2$, in which the polymer chain dimension scales like a random walk.

To quantify the scaling of the radius of gyration of the collapsed PS state as a function of chain length, we characterise the scaling of the average radius of gyration, comparing the results from non-accelerated molecular dynamics with results from metadynamics. For the metadynamics calculations, we plot average radius of gyration $R_g$ of PS at the minimum energy state vs. numbers of monomers $N$ as shown in Figure 3.4 (a) compared with Figure 3.5. To calculate the radius of gyration, we sum over all atoms where $R_g$ is defined as:

$$R_g^2 = \frac{1}{N} \sum_{k=1}^{N} (r_k - r_{\text{min}})^2 \quad (3.8)$$

and $r_k$ is the position of the kth atom. By fitting $\langle R_g \rangle \propto N^{\nu}$, we find that the average radius of
gyration for polymer chain at the minimum energy state from metadynamics calculations, $R_g = 2.31 \, N^{0.36}$, providing a scaling coefficient $\nu = 0.36 \pm 0.02199$ close to $1/3$, expected for collapsed states of long hydrophobic polymer chains. In comparison, for non-accelerated molecular dynamics, we calculate $R_g$ as an average over the last $0.2$ ns after $50$ ns equilibration. In this case, we obtain $R_g = 2.26 \, N^{0.42}$, giving a scaling coefficient $\nu = 0.42 \pm 0.03882$, where $1/2 > 0.42 > 1/3$, indicating that the polymer chain does not reach the minimum globule state without enhanced sampling techniques. The scaling coefficient $\nu = 0.42$ indicates a departure from predicted polymer behaviour in a poor solvent. The polymer does not reach collapsed globular shape, and we can attribute this result to the fact that approaching equilibrium in polymer–water solutions is difficult and requires longer simulation timescales or else lower resolution techniques such as coarse graining.

Figure 3.4. (a) the average radius of gyration $R_g$ of PS chains at the minimum energy state vs. numbers of monomers $N$ and (b) the free energy (kcal/mol) as a function of radius of gyration ($R_g$) for nine PS chain lengths of $5, 8, 10, 13, 17, 22, 30, 44$ and $58$ monomers after $10$ ns metadynamics calculations.
Figure 3.5. Radius of gyration $R_g$ dependence of number of monomers $N$ (including standard deviation) in all atomistic non-accelerated MD simulations of equilibrated single PS chain and water systems.

3.2.2 PS free energy characterization in all-atomistic simulations

Next, we examine the shape of the potential of mean force as a function of radius of gyration obtained from metadynamics calculations as shown in Figure 3.4(b) for chain lengths from 5 to 58 monomers. All curves exhibit a characteristic minimum $R_g$, indicating a collapsed globular state. The increase in free energy at smaller radii of gyration corresponds to a barrier for compressing the globule, while the increase in free energy corresponding to increasing radius of gyration corresponds to stretching the chain away from its equilibrium radius of gyration. Moreover, several potential of mean force curves (for 8, 10, 13 monomers) exhibit a secondary minimum at a higher radius of gyration and correspondingly higher free energy, thus exhibiting a local minimum or unstable chain conformation during the transition between coil and globular state. The characteristic $R_g$ for the globular state is increasing with the number of monomers in the chain, ranging from 4.75–6.75 Å to 8.25–10.25 Å (for 30–58 monomers). The secondary minimum at higher radius of gyration (Table 3.6) can be attributed to a local minimum during the folding process with a lower local hydration water density, and will be discussed further later.
Table 3.6. All atomistic metadynamics simulation results for all single PS chains in water systems. PMF minima and radius of gyration were recorded from the last pmf file output of each simulation. More radii of gyration within ±0.05Å of this value were sought throughout the trajectory and averaged for each minima. R\textsubscript{g} standard deviations, water box sizes and number of water atoms are also included.

### 3.2.3 Coarse-grained simulations

We next compare results from coarse grain simulations with all atomistic simulations of the same PS chains lengths of 5, 8, 10, 13, 17, 22, 30, 44 and 58 monomers in water. Initially, we compare conformations of the coarse-grained PS chains in water at the minimum energy conformation from metadynamics calculations as shown in Figure 3.6(a)–(i) to conformations of the PS chains after 50 ns non-accelerated molecular dynamics as shown in Figure 3.7(a)–(i). The PS polymer chain is displayed in a liquorice representation. The head and tail hydrogens in the hydroxyl group are highlighted using a blue van der Waals representation. Water is not shown. As shown in Figure 3.6, it is clearly seen that for polymer lengths of 8–10 monomers or longer the chain collapses into a globular state. This is complementary to what is seen with atomistic simulations. This characteristic size of the globular state increases as a function of monomer length. For chains shorter than eight monomers, i.e. five monomers, the polymer chain does not collapse and 44 remains in a more extended conformation. In comparison, as shown in Figure 3.7,
for all polymer lengths, after 200 ns non-accelerated molecular dynamics, the coarse-grained chain conformation also approaches a globular state.

Figure 3.6. Snapshots of the collapsed CG PS chains in water at the minimum energy configuration calculated using metadynamics for lengths of (a) 5, (b) 8, (c) 10, (d) 13, (e) 17, (f) 22, (g) 30, (h) 44 and (i) 58 PS monomers of in water. The polymer chain is displayed in a licorice representation. The head and tail hydrogens in the hydroxyl group are highlighted using a blue van der Waals representation. CG water not shown.

Figure 3.7. Single PS chains in water of the following lengths: a) 5, b) 8, c) 10, d) 13, e) 17, f) 22, g) 30, h) 44 and i) 58 monomers. These snapshots were taken at the end of CG non-accelerated MD simulations of equilibrated PS-water systems. The polymer chain is displayed in a licorice representation. The head and tail beads are highlighted using a blue van der Waals representation.
3.2.4 PS free energy characterization in coarse-grained simulations

To quantify the scaling of the radius of gyration of the collapsed coarse-grained PS state as a function of chain length, we characterise the scaling of the average radius of gyration, \( R_g \), comparing the results from non-accelerated molecular dynamics with results from metadynamics. For the coarse-grained metadynamics calculations, we plot average radius of gyration \( R_g \) of PS at the minimum energy state vs. numbers of monomers \( N \) as shown in Figure 3.8 (a) compared with Figure 3.9. To calculate \( R_g \) we include all CG beads in the calculation. By fitting \( \langle R_g \rangle \propto N^\nu \), we find that the average radius of gyration for coarse-grained polymer chain at the minimum energy state from metadynamics calculations, \( R_g(N) = 2.88 \, N^{0.32} \), with scaling coefficient \( \nu = 0.32 \pm 0.01865 \) close to 1/3 (Figure 3.8(a)). In comparison, for long-time non-accelerated coarse-grained molecular dynamics, we calculated \( R_g \) as an average over the last 5 ns after 200 ns equilibration. In this case, we obtain \( R_g(N) = 3.17 \, N^{0.3} \), \( \nu = 0.3 \pm 0.01875 \), extremely close to the atomistic and coarse-grained scaling coefficients.

Next, we examine the shape of the potential of mean force as a function of radius of gyration obtained from coarse-grained metadynamics calculations as shown in Figure 3.8(b) for chain lengths from 5 to 58 monomers. As compared with the potential of mean force obtained with AA calculations as shown in Figure 3.4(b), the wells in the free energy surface are slightly wider, due to the difference in simulation resolution. Furthermore, a larger number of potential of mean force curves exhibit a secondary minimum, as opposed to the curves from AA calculations. (Table 3.7). For example, polymer lengths ranging from 10 to 44 monomers all display a secondary minimum. This implies that during the collapse of the coarse-grained chains metastable conformations of the polymer chains are achieved that are not necessarily seen in atomistic simulations where water and
chain conformations are simulated in molecular level detail. This is most likely a result of the lower resolution model of the PS and, in particular, the lower resolution model of the coarse-grained water.

Figure 3.8 (a) the average radius of gyration $R_g$ of the CG PS chains at the minimum energy state calculated using metadynamics vs. numbers of monomers $N$ and (b) the free energy (kcal/mol) as a function of radius of gyration ($R_g$) for nine PS chain lengths of 5, 8, 10, 13, 17, 22, 30, 44 and 58 monomers after 50 ns metadynamics.

Figure 3.9. Radius of gyration $R_g$ dependence of number of monomers $N$ (including standard deviation) in CG non-accelerated MD simulations of equilibrated single PS chain and water systems.
Table 3.7. CG metadynamics simulation results for all single PS chains in water systems. PMF minima and radius of gyration were recorded from the last pmf file output of each simulation. More radii of gyration within ±0.05Å of this value were sought throughout the trajectory and averaged for each minima. $R_g$ standard deviations, water box sizes and number of water atoms are also included.

<table>
<thead>
<tr>
<th>PS Length</th>
<th>PMF 1st minimum</th>
<th>$&lt;R_g&gt;$ 1st</th>
<th>$\delta R_g$</th>
<th>PMF 2nd minimum</th>
<th>$&lt;R_g&gt;$ 2nd</th>
<th>$\delta R_g$</th>
<th>Box Size (Å$^3$)</th>
<th>Water Atoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0</td>
<td>4.75</td>
<td>0.04</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>100x100x100</td>
<td>96447</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>5.25</td>
<td>0.0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>100x100x100</td>
<td>95703</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>6.35</td>
<td>0.0</td>
<td>6.51067</td>
<td>8.31</td>
<td>0.0</td>
<td>100x100x100</td>
<td>96309</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>6.79</td>
<td>0.0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>100x100x100</td>
<td>96195</td>
</tr>
<tr>
<td>17</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>32.1962</td>
<td>7.24</td>
<td>0.01</td>
<td>100x100x100</td>
<td>95703</td>
</tr>
<tr>
<td>22</td>
<td>0</td>
<td>7.71</td>
<td>0.008</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>100x100x100</td>
<td>95940</td>
</tr>
<tr>
<td>30</td>
<td>0</td>
<td>8.6</td>
<td>0.0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>100x100x100</td>
<td>95703</td>
</tr>
<tr>
<td>44</td>
<td>0</td>
<td>9.81</td>
<td>0.0</td>
<td>5.97183</td>
<td>n/a</td>
<td>n/a</td>
<td>100x100x100</td>
<td>95331</td>
</tr>
<tr>
<td>58</td>
<td>0</td>
<td>10.27</td>
<td>0.13</td>
<td>95.264</td>
<td>n/a</td>
<td>n/a</td>
<td>100x100x100</td>
<td>94896</td>
</tr>
</tbody>
</table>

3.2.5 Scaling behaviour of the polymer chain from metadynamics calculations

The results for average $R_g$ vs. N, (Figure 3.10) show that from AA simulations, $R_g$ (N) = 2.31N$^{0.36}$, with $\nu = 0.36 \pm 0.02199$ and from CG simulations, $R_g$ (N) = 2.88N$^{0.32}$, $\nu = 0.32 \pm 0.01865$. Both all-atomistic and coarse-grained simulations exhibit a power law close to 1/3, which is what is expected for a hydrophobic chain in aqueous medium (poor solvent conditions).

![Figure 3.10](image-url)
3.3 Solvent Accessible Surface Area (SASA)

Solvent accessible area (SASA) is the area of potential contact between a biomolecule and a solvent (158). To calculate SASA, a sphere of a given radius probes the van der Waals surface of the molecule. Initially the concept of SASA was applied to protein folding to form globular structures. There is a nonlinear correlation between SASA and the molecular weight M of the protein, SASA = kM^γ, with multiple reported values of γ, ranging from 0.76 (159) to 0.87 (160). For monomeric proteins k = 6.3 (159). We expect a similar behavior for single PS polymer chains, as a measure of characterization of globular structures determined by hydrophobic collapse of the PS chains. Collapsed, globular states of single PS chains in water have a characteristic radius of

$$R = \left(\frac{3V}{4\pi}\right)^{1/3} \text{ or } R = \left(\frac{A}{4\pi}\right)^{1/2}$$

(3.9)

where V is solvent excluded volume and A is SASA (147). Next, from $R_g \propto N^\nu \propto (\text{SASA}/4\pi)^{1/2}$, we know that SASA $\propto N^{2\nu}$, which for an ideal polymer chain in bad solvent ($\nu = 1/3$), SASA $\propto N^{2/3}$.

Thus, in order to investigate the solvation effects during PS collapse, we calculate SASA for both AA and CG simulations, using a probe sphere of radius of 1.4 Å for AA simulations and a probe sphere of 4.2 Å for CG simulations. For the AA systems we characterize SASA time evolution over the initial 50 ns, while for the CG systems, we characterized SASA time evolution over the initial 200 ns as shown in Figures 3.11 and 3.12.
Figure 3.11. Solvent accessible surface area time evolution for all atomistic non-accelerated MD simulations of equilibrated single PS chains of the following lengths: a) 5PS; b) 8PS; c) 10PS; d) 13PS; e) 17PS; f) 2PS; g) 22PS; h) 30PS; h) 44PS; i) 58PS monomers and water systems.
Figure 3.12. Solvent accessible surface area time evolution for CG non-accelerated MD simulations of equilibrated single PS chains of the following lengths: a) 5PS; b) 8PS; c) 10PS; d) 13PS; e) 17PS; f) 2PS; g) 22PS; h) 30PS; h) 44PS; i) 58PS monomers and water systems.

From Figure 3.11 it is clear that for AA simulations the equilibration time for longer polymer chains (N ≥ 17 monomers) becomes significantly longer than shorter chains, however for CG simulations the equilibration time is insignificant. Overall, the fluctuations in SASA increase as function of N, for both AA (Figure 3.13 (a)) and CG (Figure 3.13(b)) simulations.
Figure 3.13. Standard deviation in SASA versus number of monomers $N$ computed for a) all atomistic and b) CG non-accelerated MD simulations of equilibrated single PS chain and water systems

For metadynamics calculations (for both AA and CG simulations), we also calculate SASA for the collapsed, globular state at the minimum of the PMF as a function of polymer length. In particular, for cases when more than one collapsed states were observed, a SASA average over the multiple minimum PMF states is calculated. Specifically, we calculate average SASA for the last 10 ns (Tables 3.8 and 3.9) of simulation time and plot these results against number of monomers, $N$.

<table>
<thead>
<tr>
<th>PS Length</th>
<th>&lt;SASA&gt; for last 10ns ($\text{Å}^2$)</th>
<th>$\delta$SASA</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>813.25</td>
<td>28.34</td>
</tr>
<tr>
<td>8</td>
<td>1112.62</td>
<td>48.21</td>
</tr>
<tr>
<td>10</td>
<td>1248.87</td>
<td>48.78</td>
</tr>
<tr>
<td>13</td>
<td>1510.22</td>
<td>53.57</td>
</tr>
<tr>
<td>17</td>
<td>1919.43</td>
<td>97.73</td>
</tr>
<tr>
<td>22</td>
<td>1973.15</td>
<td>70.86</td>
</tr>
<tr>
<td>30</td>
<td>2546.82</td>
<td>48.18</td>
</tr>
<tr>
<td>44</td>
<td>3502.06</td>
<td>63.76</td>
</tr>
</tbody>
</table>

Table 3.8. Solvent accessible surface area averaged over last 10 ns of all atomistic non-accelerated MD simulations of equilibrated systems of 5, 8, 10, 13, 17, 22 30, 44 and 58 single PS chain lengths in water, with standard deviations.

<table>
<thead>
<tr>
<th>PS Length</th>
<th>&lt;SASA&gt; for last 10ns ($\text{Å}^2$)</th>
<th>$\delta$SASA</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>843.98</td>
<td>35.95</td>
</tr>
<tr>
<td>8</td>
<td>1205.45</td>
<td>59.04</td>
</tr>
<tr>
<td>10</td>
<td>1326.35</td>
<td>86.71</td>
</tr>
<tr>
<td>13</td>
<td>1518.80</td>
<td>72.70</td>
</tr>
<tr>
<td>17</td>
<td>1827.67</td>
<td>55.16</td>
</tr>
<tr>
<td>22</td>
<td>2215.58</td>
<td>89.81</td>
</tr>
<tr>
<td>30</td>
<td>2709.12</td>
<td>89.82</td>
</tr>
<tr>
<td>44</td>
<td>3675.59</td>
<td>97.17</td>
</tr>
<tr>
<td>58</td>
<td>4489.74</td>
<td>146.98</td>
</tr>
</tbody>
</table>

Table 3.9. Solvent accessible surface area averaged over last 10 ns of CG non-accelerated MD simulations of equilibrated systems of 5, 8, 10, 13, 17, 22 30, 44 and 58 single PS chain lengths in water, with standard deviations.
We next fit $<\text{SASA}> \propto N^\mu$. For AA simulations we obtain $\mu = 0.68 \pm 0.02227$ (Figure 3.14 (a)), and for CG $\mu = 0.7 \pm 0.01464$ (Figure 3.14 (b)). In both cases, scaling coefficient $\mu$ is slightly larger than $2/3$, and within the range of previously reported values for collapsed protein structures. Next, using the results from metadynamics, we calculate the SASA at the collapsed, globular state for both AA and CG simulations as shown in Figure 3.15. For the PS chain lengths exhibiting a secondary minimum, we calculate an average SASA over the minimum energy states and it SASA
vs. number of monomers in each chain. For AA simulations, we obtain \( \text{SASA} (N) = 257.11 \ N^{0.67}, \mu = \pm 0.03675 \) and for CG, \( \text{SASA} (N) = 276.32 \ N^{0.67}, \mu = \pm 0.02352 \). In this case, the scaling coefficient \( \mu \approx 2/3 \) is nearly identical for AA and CG simulations. Furthermore, the SASA at the second minimum energy configuration is invariably larger that at the first minimum, indicating that as the polymer chain collapses, the chain samples multiple minima before finding a minimum conformation for the radius of gyration \( R_g \) of the PS, where the SASA is also at a minimum.

SASA time evolution for AA simulations (Figure 3.10) shows a definite decrease towards the end of simulation for intermediate and longer chains (22, 30, 44 and 58PS), attributed to the transition from disordered coils to the collapsed states of the polymer. As a consequence, the available configuration space of water molecules is increasing, a condition similar to that of protein folding, upon which the solvent gains translational entropy (161). In case of PS collapse, the newly configured entropy of the water molecules drives chain folds inside the globular structure to arrange in compressed configurations of low entropy (162-164).

### 3.4 Hydration Waters

For AA and CG simulations, we calculate the number of hydration waters corresponding to each SASA value determined previously during metadynamics calculations. We calculate \(< \text{N}_w/\text{SASA} >\) vs. \( N \) (number of monomers per polymer chain) and \(< \text{N}_w/\text{SASA} >\) vs. \(< \text{SASA} >\) as shown in Figure 3.16. The number of hydration waters per SASA at first decreases for short chains (5, 8, 10 and 13 monomers chains), then increases for longer chains (17, 22, 30 and 58 monomers) (Figure 3.16(a)). This suggests that compact, dewetted states are not readily achieved for shorter chains, but that dewetted states are more characteristic for intermediate chain lengths. The same behavior was observed in CG simulations (Figure 3.16(c). Likewise, \(<\text{N}_w/\text{SASA} >\) vs. SASA (Figure
3.16(b) and (d)) confirms a marked decrease with increasing SASA corresponding to short chains, showing that while surface accessible surface area is increasing for shorter chains (5, 8, 10 and 13 monomers), number of hydration waters per SASA is decreasing, only to increase later with SASA for longer chains (17, 22, 30 and 58 monomers). Thus, the folded conformations of shorter chains are wetted, a process that is reversed in the case of intermediate length chains, in which folding is accompanied by lower water densities around folded chains. Previously it was reported that in the process of folding to a compact, globular state, the polymer goes to dry transition states with low vicinal water densities (147,148). From the statistics of radius of gyration, we determine the $R_g$ range, distribute it into ten sections and average $N_w$/SASA for each $R_g$ range as shown in Figure 3.17. States with high water densities are globular states in which the polymer is wetted, vs. states with comparably lower water densities corresponding to dry states of the polymer chain. Also, the highest water densities’ radius of gyration in Figure 3.17 corresponds to the minimum free energy (Figure 3.3 (b)) of the collapsed chain conformations. For example, in AA simulations, for 30 monomer chain length, $N_w$/SASA = 0.06 occurs at $R_g = 8.91$ Å (Figure 3.17(g)), while in Figure 3.3(b), the local free energy minimum for 30 PS occurs at $R_g = 8.26$ Å (Table 3.5).
Figure 3.16 Average number of hydration waters per SASA ($N_w$/SASA) vs. (a) number of monomers $N$ and vs. (b) SASA at the minimum energy state calculated with metadynamics for atomistic simulations. Average number of hydration waters per SASA ($N_w$/SASA) vs. (a) number of monomers $N$ and vs. (b) SASA at the minimum energy state calculated with metadynamics for CG simulations.
3.5 Conclusions

Herein we describe and compare results of all-atomistic (AA) and coarse-grained (CG) molecular dynamics simulations of single PS chains of varying lengths in a solution of water. We utilize enhanced sampling techniques in molecular dynamics, specifically metadynamics, to improve sampling of the conformation of the polymer chain in both AA and CG molecular dynamics. Overall, these results indicate initial validation of the CG model of PS as compared with AA simulations for these simulations of single PS chains in water. By fitting $<R_g> \propto N^v$, we find that the scaling coefficients for average radius of gyration for AA and coarse-grained polymer chains at the minimum energy state from metadynamics calculations are comparable, and close to
1/3. Furthermore, by fitting $<\text{SASA}> \propto N^\mu$, we again find that the scaling coefficients for SASA for AA and CG polymer chains as a function of chain length are comparable, and close to 2/3. Furthermore, calculations of the local hydration water density as a function of chain length suggest that the minimum energy conformation for intermediate chain lengths has a comparably lower hydration water density. This suggests that intermediate chain lengths are more likely to form dewetted, globular states, as compared with shorter or longer chain lengths. Thus, SASA measurements provide another criteria to characterize the conformation of the collapsed states of hydrophobic polymers such as PS.
CHAPTER 4

Molecular Dynamics Simulations of Phospholipid Bilayers with Polycaprolactone

4.1 Lipid Bilayer Interactions with Polymers

The molecular interaction between common polymer chains used in biomedical applications and the cell membrane is unknown. This interaction may affect the biocompatibility of the polymer chains. Molecular dynamics simulations offer an emerge tool to characterize the interaction between common degradable polymer chains used in biomedical applications, such as polycaprolactone, and model cell membranes. Herein we systematically characterize with long-time all-atomistic molecular dynamics simulations (~150 ns) the interaction between single polycaprolactone chains of varying chain lengths with a model phospholipid membrane. We find that the length of polymer chain greatly affects the nature of interaction with the membrane, as well as the membrane properties. Furthermore, we next utilize advanced sampling techniques in molecular dynamics to characterize the two-dimensional free energy surface for the interaction of varying polymer chain lengths (short, intermediate, and long) with model cell membranes. We find that the free energy minimum shifts from the membrane-water interface to the hydrophobic core of the phospholipid membrane as a function of chain length. These results can be used to design polymer chain lengths and chemistries to optimize their interaction with cell membranes at the molecular level.

4.2 Introduction

The cell membrane acts as a barrier of selective permeability barrier, surrounding cells of living organisms. The membrane can allow the diffusion of certain hydrophobic molecules passively
into the membrane, but block the entry of hydrophilic molecules. Self-assembled block copolymers are common drug delivery vehicles (165). For example, a widely used, biodegradable polymer commonly used for drug delivery is polycaprolactone (PCL) (11,166-168). PCL is also used for development of other biomaterial applications such as stents (169) and is used in tissue scaffolds (12). PCL is a hydrophobic, semi-crystalline polymer (170,171). However, the nature of the interaction between PCL polymers and the cell membrane is unknown. For example, it is unknown if a single PCL polymer itself can cross the cell membrane due to its hydrophobic nature; moreover, if the polymer is intercalated into the membrane, the polymer chains may affect the membrane properties, such as the ordering of the phospholipid tails, the area per lipid, and membrane thickness. For example, block copolymer interactions with the membrane have been shown to perturb and even porate the membrane using simulation studies (78). Furthermore, simulations of the interaction of Pluronic copolymers interacting with the polar head groups of lipid molecules show that membrane properties are altered to allow permeation of drugs (77,172). Additional coarse-grained simulations have shown that Poloxamer interaction with lipid membranes showed that shorter chains insert partially into the membrane, while longer chains cross the membrane completely, inducing thinning of the membrane and increasing area per lipid (173).

In this study, we focus on the characterization of the interaction of hydrophobic polymers such as PCL with model cell membranes such as POPC using large-scale all atomistic simulations. POPC (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine) is chosen since it is a model lipid found in the cell membrane and is commonly used for biophysical measurements. PCL, as previously discussed, is synthetic hydrophobic aliphatic polyester with semi-crystalline properties that has been approved by the Food and Drug Administration (FDA) to be used in various
biomedical applications. A variety of hydrophobic drugs have been encapsulated within PCL self-assemblies for controlled release and targeted drug delivery, such as paclitaxel (7). Indeed, the shape of the self-assembled structure has been shown to influence the solubility of the hydrophobic drug in the hydrophobic environment of a micelle core (174).

In this study, we describe results of long-time all-atomistic molecular dynamics simulations of single polycaprolactone (PCL) chains of varying lengths in solvated POPC lipid membranes. We find that the nature of the interaction of the PCL chain with the lipid membrane varies as a function of chain length, finding that shorter chains prefer to sit at the membrane-water interface while longer chains prefer to behave as slightly distended hydrophobic globules in the center of the hydrophobic lipid membrane core. We find that this variation of the strength of interaction by polymer length also affects the membrane properties, such as area per lipid and membrane width. Surprisingly, we find that intermediate chain lengths thin the membrane the most during the length of these simulation studies. We next characterize the complete interaction free energy surface of the polymer with the membrane using advanced sampling techniques in molecular dynamics, namely metadynamics (100). We have previously used metadynamics to characterize the collapse of single hydrophobic chains of PS (polystyrene) in water (175). In this case, to characterize the interaction free energy surface for PCL with the model phospholipid membrane composed by POPC, we choose to characterize the interaction free energy as a function of two separate reaction coordinates: radius of gyration of the polymer and the distance between center of masses of the POPC membrane and the PCL chains. We find results that are consistent with previously described long-time all-atomistic simulations studies, with the free energy minimum that shifts from the membrane-water interface to the center of the hydrophobic inner core of the
4.3 Simulation Systems

Simulations were carried out for PCL sizes of 5, 10, 13, 22, 30, 44 and 58 monomers in water boxes of various sizes (Table 4.1). Boxes containing single chains were initially equilibrated at T = 300K for 50 ns. Force field parameters were based on CHARMM27 (149). Simulations were performed with the NAMD2 (150,151) software package. A cut-off 12Å was used for the calculations of non-bonded interactions, with LJ switching function of 10Å. Simulations were performed with a time step of 2fs and carried out in NPT ensemble, with a temperature held constant by Langevin dynamics with a coupling constant of 5 ps⁻¹. Similarly, the pressure was maintained anisotropically by a Langevin piston at 1.01325 bar.

<table>
<thead>
<tr>
<th>PCL monomers</th>
<th>Box Size [Å³]</th>
<th>water atoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>197.143x98.571x98.571</td>
<td>193593</td>
</tr>
<tr>
<td>10</td>
<td>118.058x59.029x59.029</td>
<td>41355</td>
</tr>
<tr>
<td>13</td>
<td>200x100x100</td>
<td>193953</td>
</tr>
<tr>
<td>22</td>
<td>235.688x39.281x58.922</td>
<td>55002</td>
</tr>
<tr>
<td>30</td>
<td>314.909x59.045x59.045</td>
<td>110781</td>
</tr>
<tr>
<td>44</td>
<td>413.890x59.127x59.127</td>
<td>145488</td>
</tr>
<tr>
<td>58</td>
<td>531.999x59.111x59.111</td>
<td>187161</td>
</tr>
</tbody>
</table>

*Table 4.1 PCL in water system sizes and set ups.*

POPC bilayer membrane systems consisting of 59 POPC and 3,222 TIP3P water molecules containing nine different monomer chain lengths of polycaprolactone (5, 8, 10, 13, 17, 22, 30, 44 and 58 monomers) (Table 4.2) were simulated with using a 0.2 M concentration of KCl salt in solution for the time period of 150 ns each.
Table 4.2. Radius of gyration averaged over the last 100 ns of all atomistic MD simulations of equilibrated systems of 5, 8, 10, 13, 17, 22, 30, 44 and 58 single PCL chain lengths in POPC membrane, with standard deviations. Water box sizes and number of water atoms are also included.

<table>
<thead>
<tr>
<th>PCL Length</th>
<th>&lt;R_g&gt; for last 100 ns (Å)</th>
<th>δR_g (Å)</th>
<th>Box Size (Å³)</th>
<th>Water Atoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>no PCL</td>
<td>42.62x45.63x87.71</td>
<td></td>
<td>9672</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>8.96</td>
<td>1.40</td>
<td>37.36x38.48x93.42</td>
<td>8277</td>
</tr>
<tr>
<td>8</td>
<td>11.50</td>
<td>1.64</td>
<td>36.86x36.58x93.71</td>
<td>7548</td>
</tr>
<tr>
<td>10</td>
<td>14.47</td>
<td>3.05</td>
<td>40.36x39.52x84.31</td>
<td>8277</td>
</tr>
<tr>
<td>13</td>
<td>17.11</td>
<td>1.53</td>
<td>40.52x39.27x85.92</td>
<td>8346</td>
</tr>
<tr>
<td>17</td>
<td>17.83</td>
<td>3.40</td>
<td>39.82x36.28x87.20</td>
<td>7548</td>
</tr>
<tr>
<td>22</td>
<td>13.47</td>
<td>1.23</td>
<td>39.73x36.11x88.06</td>
<td>7551</td>
</tr>
<tr>
<td>30</td>
<td>12.94</td>
<td>1.42</td>
<td>39.38x37.42x87.58</td>
<td>7548</td>
</tr>
<tr>
<td>44</td>
<td>16.11</td>
<td>0.88</td>
<td>37.24x36.73x91.68</td>
<td>7422</td>
</tr>
<tr>
<td>58</td>
<td>14.64</td>
<td>0.67</td>
<td>43.47x37.48x77.87</td>
<td>7422</td>
</tr>
</tbody>
</table>

Force field parameters (bond, angle, and dihedral) for the PCL chains as well as partial charges were based on the CHARMM27 force field (149). Simulations were performed with the NAMD2 software package (150,151). The polymer chain conformation was initially placed in the center of the POPC bilayer membrane as shown in Figure 1 (a), (d), and (g).
Figure 4.1. Snapshots of PCL in POPC membrane before simulation (a, d, g), at 75ns (d, e, f) and at 150ns (g, h, i) for 5 PCL (a, b, c), 17 PCL (d, e, f) and 58PCL (g, h, i). The polymer chain, potassium and chloride ions are displayed in a van der Waals representation while POPC membrane chains are displayed in a CPK representation. Water is indicated by transparent conformation background.

The CHARMM36 force field (176-178) was used for POPC (179) and the CHARMM TIP3P model (also known as the TIPS3P model) (180,181) were used for water. All systems used the NPT ensemble and Langevin dynamics (182) at a temperature of 300K with a damping coefficient $\gamma=5$ ps$^{-1}$, at a pressure of 1 atm using an anisotropic Langevin piston method (182,183) with a piston period of 200 fs and a damping time scale of 50 fs. The SHAKE algorithm was used to hold covalent bonds involving hydrogen rigid, allowing a 2 fs time step. The Particle Mesh Ewald (PME) (184) algorithm was employed to take full electrostatic interactions into account, with full periodic boundary conditions. The cut-off for van der Waals interactions was 12 Å with a smooth switching function at 10 Å used to truncate the van der Waals potential energy at the cutoff distance. Coordinates were saved every 2 ps for the trajectory analysis.
4.4 Free Energy Calculations.

Two-dimensional metadynamics simulations were run for 150 ns for three different polymer lengths (5, 17, and 58 monomers) using the Collective Variables Module (156) in NAMD2 using the radius of gyration of the polymer chain and the center of mass distance between the PCL chain and the phospholipid headgroups as reaction coordinates. Metadynamics (100) is a methodology whereby a Gaussian potential is applied with time to the free energy surface until it can overcome free energy minima. Metadynamics has been extensively applied to accurate sampling of the free energy surfaces of protein configurations (104), chemical reaction pathways (145), and other phenomena. Metadynamics parameters of hill energy (hill weight) was set to 0.01 kcal/mol, hill creation frequency, (i.e. integration steps required for addition of a new hill to the metadynamics potential) was set to 100, and radius of gyration fluctuation amplitude was set to 0.5 Å and center of mass fluctuation amplitude was set to 10 Å.

4.5 Chain Collapse

We observe that for all chain lengths of polycaprolactone (5, 8, 10, 13, 17, 22, 30, 44 and 58 monomers) the polymer chain opens up from its initial collapsed conformation in the membrane and moves towards the phospholipid membrane interface. This is highlighted in Figure 4.1, for three different chain lengths (5, 17, and 58 monomers). To begin with, Figure 4.1 (a), (d), (g) clearly shows the collapsed PCL chains of different lengths are in the center of the phospholipid membrane tails. After half of the total simulation time, the PCL membrane chain moves clearly towards the phospholipid headgroup-water membrane interface as shown in Figure 4.1 (b), (e), (h). The movement towards the phospholipid headgroup interface however is not as pronounced
for the longest PCL chain length of 58 monomers as it is for the shortest chain length of 5 monomers and intermediate chain length of 17 monomers. Indeed, the 58 monomer PCL chain length remains in a globular state in the center of the hydrophobic membrane core as shown in Figure 4.1 (h), (i), but the radius of gyration does expand from its initial conformation in Figure 4.1 (g). After the total simulation time of 150 ns the shortest chain length of 5 monomers remains at the membrane interface in an extended conformation as shown in Figure 4.1 (c), while the intermediate chain length of 17 monomers remains at the membrane interface, however it generates a hydrophobic tail that intercalates towards the center of the hydrophobic core of the membrane as shown in Figure 4.1 (f). Indeed, this trend for shorter to intermediate to longer chain lengths for the conformation of the polymer in the membrane at the end of the simulation time of 150 ns continues for all nine-polymer lengths simulated. We highlight the chain conformation at the end of the simulation time as shown in Figure 4.2 for all polymer chain lengths (5, 8, 10, 13, 17, 22, 30, 44 and 58 monomers).
Figure 4.2. Snapshots of PCL in solvated POPC membrane at the end of 150 ns simulation PCL: a) 5 PCL; b) 8 PCL; c) 10 PCL; d) 13 PCL; e) 17 PCL; f) 22 PCL; g) 30 PCL; h) 44 PCL; i) 58PCL. The polymer chain, potassium and chloride ions are displayed in a van der Waals representation while POPC membrane chains are displayed in a CPK representation. Water is indicated by transparent conformation background.

Clearly, we see a trend that for short chains, the PCL chain is in an extended conformation close to the phospholipid membrane-water interface as shown in Figure 4.2 (a), (b) for chain lengths of 5, 8 monomers. For intermediate chain lengths (10, 13, 17, 22, 30 monomers) the polymer chain moves towards the membrane interface, however chain ends wrap around or else intercalate into the center of the hydrophobic core of the membrane as shown in Figure 4.2 (c) – (g). For the longest chain lengths as shown in Figure 4.2 (h), (i) the chain remains in a collapsed conformation that slightly opens up and maintains contacts with the membrane-water interface. Overall, we find that—for the timescales simulated—shorter PCL chains move towards the membrane-water interface in an extended conformation, while longer chains remain in a more collapsed conformation in the center of the phospholipid membrane.
4.6 Radius of Gyration and End-to-End Distance for PCL in POPC

4.6.1 Initial simulations: PCL in water

Initial all-atomistic simulations were performed for single PCL chains of 5, 10, 13, 122, 20, 44 and 58 monomers were solvated in water boxes of various sizes for 50ns. For system sizes and setups see Table 4.3. Time evolution of $R_g$ (Figure 4.3) indicates that short chains (5, 10, 13PCL) have high fluctuations, while large chains (30, 44 58PCL) exhibit $R_g$ stability of chain conformation, corresponding to collapsed states. End-to-end distance time evolutions conforms $R_g$ results, indicating large fluctuations for small chains and conformation stability for long chains (Figure 4.4)

![Images of Figure 4.3](image_url)  
*Figure 4.3. Time evolution of the radius of gyration ($R_g$) for PCL5 (a), PCL10 (b), PCL13 (c), PCL22 (d), PCL30 (e), PCL44 (f) (PCL in water).*
Figure 4.4. Time evolution of end-to-end distance ($R_e$) for PCL5 (a), PCL10 (b), PCL13 (c), PCL22 (d), PCL30 (e), PCL44 (f) (PCL in water).

Table 4.3. Radius of gyration and end-to-end distance averaged over the last 5ns of all atomistic MD simulations of equilibrated systems of 5, 10, 13, 22, 30, 44 and 58 single PCL chain lengths in water, with standard deviations. Water box sizes and number of water atoms are also included.

<table>
<thead>
<tr>
<th>Monomers</th>
<th>$\langle R_g \rangle$</th>
<th>$\delta R_g$</th>
<th>$\langle R_e \rangle$</th>
<th>$\delta R_e$</th>
<th>Box Size</th>
<th>Water Atoms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[Å]</td>
<td>[Å]</td>
<td>[Å]</td>
<td>[Å]</td>
<td>[Å³]</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5.84</td>
<td>1.12</td>
<td>13.14</td>
<td>4.45</td>
<td>197.143x98.571x98.571</td>
<td>193593</td>
</tr>
<tr>
<td>10</td>
<td>7.75</td>
<td>1.94</td>
<td>17.13</td>
<td>9.25</td>
<td>118.058x59.029x59.029</td>
<td>41355</td>
</tr>
<tr>
<td>13</td>
<td>11.01</td>
<td>0.37</td>
<td>24.82</td>
<td>3.7</td>
<td>200x100x100</td>
<td>193953</td>
</tr>
<tr>
<td>22</td>
<td>8.3</td>
<td>0.3</td>
<td>14.11</td>
<td>5.46</td>
<td>235.688x39.281x58.922</td>
<td>55002</td>
</tr>
<tr>
<td>30</td>
<td>9.8</td>
<td>0.29</td>
<td>11.79</td>
<td>5.62</td>
<td>314.909x59.045x59.045</td>
<td>110781</td>
</tr>
<tr>
<td>44</td>
<td>10.76</td>
<td>0.32</td>
<td>21.02</td>
<td>5.12</td>
<td>413.890x59.127x59.127</td>
<td>145488</td>
</tr>
<tr>
<td>58</td>
<td>11.2</td>
<td>0.32</td>
<td>20.11</td>
<td>5.91</td>
<td>531.999x59.111x59.111</td>
<td>187161</td>
</tr>
</tbody>
</table>

$R_g$ and $R_e$ were averaged for the last 5ns of simulation time (Table 4.3). Small chains (5, 10, 13 PCL) show an increase of $R_g$ with number of monomers N. However, for PCL 22, radius of gyration is decreased, then, for longer chains, 30, 44, 58 radius of gyration is increasing again. There is no linear increase of $R_g$ with N (Figure 4.5. (a)). The power fit of average $R_g$ versus number of monomers N gives $R_g(N) = 6.65N^{0.12}$, where the scaling coefficient $\nu = 0.120667 \pm 0.08977$ is inconclusive in terms of an indication of chain conformations (Figure 4.5.(b))
Average end-to-end distance also increases with number of monomers for short chains (5, 10, 13 PCL) (Figure 4.6.(a)), then decreases abruptly for intermediate chains (22, 30PCL), after which Re increases for longer chain 44 and 58PCL, however less than for smaller chains. Power fit (Figure 4.6.(b)) gives \( R_e(N) = 15.4N^{0.06} \), \( \nu = 0.0603588 \pm 0.151 \), in which scaling coefficient is inconclusive. These results illustrate the insufficiency of conformation space sampling for accurate representation of collapse state of PCL in water.

Moreover, equilibration snapshots (Figure 4.7) show that in 50ns simulation time small chains extend, while large PCL chains in water attain states of loosely collapsed globules.
4.6.1.1 Initial simulations: PCL in water – metadynamics

All simulations of equilibrated systems of single PCL chains in water were followed by 50 ns metadynamics (Table 4.4). Averages of radius of gyration and end-to-end distance were calculated for values close within ± 0.05Å to minima found at 0 kcal/mol from the last pmf output file. PCL chains attain collapsed conformations (Figure 4.8).
<table>
<thead>
<tr>
<th>PCL monomers</th>
<th>$&lt;R_g&gt;$ [Å]</th>
<th>$\delta R_g$ [Å]</th>
<th>$&lt;R_e&gt;$ [Å]</th>
<th>$\delta R_e$ [Å]</th>
<th>Box Size</th>
<th>Water Atoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>4.26</td>
<td>0.03</td>
<td>5.51</td>
<td>2.12</td>
<td>197.206x98.603x98.603</td>
<td>193593</td>
</tr>
<tr>
<td>10</td>
<td>5.28</td>
<td>0.02</td>
<td>9.57</td>
<td>1.15</td>
<td>118.058x59.029x59.029</td>
<td>41355</td>
</tr>
<tr>
<td>12</td>
<td>6.26</td>
<td>0.03</td>
<td>10.97</td>
<td>1.31</td>
<td>200x100x100</td>
<td>193953</td>
</tr>
<tr>
<td>22</td>
<td>7.26</td>
<td>0.03</td>
<td>14.12</td>
<td>2.21</td>
<td>235.688x39.281x58.922</td>
<td>55002</td>
</tr>
<tr>
<td>30</td>
<td>7.77</td>
<td>0.02</td>
<td>18.56</td>
<td>1.42</td>
<td>315.358x59.130x59.130</td>
<td>110781</td>
</tr>
<tr>
<td>44</td>
<td>9.26</td>
<td>0.03</td>
<td>14.68</td>
<td>4.89</td>
<td>413.890x59.127x59.127</td>
<td>145488</td>
</tr>
<tr>
<td>58</td>
<td>10.27</td>
<td>0.03</td>
<td>8.47</td>
<td>4.25</td>
<td>531.999x59.111x59.111</td>
<td>187161</td>
</tr>
</tbody>
</table>

Table 4.4. All atomistic metadynamics simulation results for all single PCL chains in water systems. For PMF minima of zero kcal/mol and corresponding radius of gyration were recorded from the last pmf file output of each simulation. Corresponding end-to-end distance was calculated also. More radii of gyration within ±0.05 Å of $R_g$ at PMF= zero kcal/mol were sought throughout the trajectory and averaged. Analog, end-to-end distance values were sought and calculated within ±0.05 Å from the $R_e$ value at PMF=0kcal/mol. $R_g$ and $R_e$ standard deviations, water box sizes and number of water atoms are included.

Figure 4.8. Snapshots of PCL in water at minimum energy: a) 5 PCL; b) 10 PCL; c) 13 PCL; d) 30 PCL; e) 44 PCL; f) 58 PCL. The polymer chain is displayed in a licorice representation. The head and tail hydrogens in the hydroxyl group are highlighted using a blue van der Waals representation. Water is indicated by transparent cyan background.

Average radius of gyration at minimum energy increases with the number of monomers (Figure 4.9 (a)). The power fit, $R_g(N) = N^{0.36}$, gives a scaling coefficient $\nu = 0.356397 \pm 0.01983$ (Figure 4.9. (b)), indicating that the polymer chains achieve collapse states during metadynamics simulations.
Figure 4.9. a) Average radius of gyration at minimum energy vs. number of monomers; b) Power fit of radius of gyration vs. number of monomers at minimum energy, $R_g(N) = 2.38N^{0.36}$, $\nu = 0.36 \pm 0.01983$.

However, average end-to-end distance at minimum energy, is increasing for PCL chains 5-30, and then decreasing for 44-58PCL chains (Figure 4.10 (a)). A power fit of end-to-end distance (Figure 4.10. (b)), $R_e(N) = 4.0N^{0.40}$, $\nu = 0.401019 \pm 0.1432$ show a scaling coefficient greater than 1/3, indicating also collapsed states.

Figure 4.10. a) Average end-to-end distance at minimum energy vs. number of monomers; b) Power fit of end-to-end distance vs. number of monomers, $R_e(N) = 4.0N^{0.40}$, $\nu = 0.40 \pm 0.1432$
Next, we analyze the potential mean force as a function of radius of gyration obtained from metadynamics calculations (Figure 4.11) for chain lengths from 5 to 58 monomers. All curves exhibit a characteristic minimum $R_g$, indicating a collapsed globular state. The increase of free energy at smaller radii of gyration corresponds to a barrier for compressing the globule. Such barrier has the highest values for 44 PCL, 10PCL and 30PCL and much smaller values for 5PCL, 58PCL and 13PCL. The increase in free energy at larger radii of gyration corresponds to stretching the chain away from its equilibrium radius of gyration. Some chains (5, 13 and 58PCL) have very little increase of free energy with the radius of gyration, indication the tendency of the chain to remain in globular collapsed state, rather than stretch from that position. Intermediate chains (10, 30, 44PCL) show the same pattern, however, the free energy is much higher for these chains, both for compressing and stretching the chain, an indication that globular states for these chain are less stable than for the chain with lower free energy (5, 13, 58PCL). The characteristic radius of gyration is for the globular states is increasing with the number of monomers, from 4.25Å (5PCL) to 10.25Å (58PCL).

![Figure 4.11. The free energy (kcal/mol) as a function of radius of gyration ($R_g$) for six PCL chain lengths of 5, 10, 13, 30, 44 and 58 monomers after 50 ns metadynamics calculations.](image-url)
4.6.2 PCL in POPC membrane

We next characterize the size of the PCL chain by calculating its radius of gyration ($R_g$) and end-to-end distance ($R_e$) and averaging over the last 100 ns simulation time for the range of chain lengths from 5-58 monomers. The time evolution of the $R_g$ and $R_e$ is shown in Figure 4.12 and 4.13. We compare radius of gyration calculated for PCL in POPC with the radius of gyration ($R_g$) of a single PCL chain in water.

![Figure 4.12. Time evolution of radius of gyration of various single PCL chains in solvated POPC membrane. a) 5PCL b) 8PCL c) 10PCL d) 13PCL e) 17PCL f) 22 PCL g) 30PCL h) 44PCL i) 58PCL.](image)
Figure 4.13. Time evolution of end-to-end distance of various single PCL chains in solvated POPC membrane. a) 5PCL b) 8PCL c) 10PCL d) 13PCL e) 17PCL f) 22 PCL g) 30PCL h) 44PCL i) 58PCL.

Comparing $R_g$ for PCL in the membrane and PCL in water as a function of chain length as shown in Figure 4.14, there is a clear trend. For larger chain sizes of more than 20 monomers, $R_g$ values are higher for the simulation performed for PCL placed in POPC membrane compared with the simulations in which PCL is placed in water only. Indeed, this trend is also true for smaller and intermediate chain lengths.
The $R_g$ of the PCL chain is always larger in the POPC membrane than in pure water. For smaller chains (5, 8, 10, 13, 17 PCL), $R_g$ average increases with the number of monomers in the chain, from 8.96Å to 17.83Å. Larger chains (22, 30, 44, 58PCL), however, display a decrease in average $R_g$, with values close to each other, fluctuating around 15Å. This tendency is also observed for PCL in water, for larger chains leveling of $R_g$ around 10Å. While PCL chains evolve to a globular state when placed in water, the POPC membrane environment opens up the collapsed globule, with smaller chains opening up to an extended conformation and moving towards the interface as shown in Figure 4.2. This effect is exemplified in Figure 4.15, which shows that the end-to-end distance of smaller chain lengths (up to 17 monomers) has a large $R_e$ that longer chain lengths, due to the extended conformation of shorter chain lengths on the membrane-water interface. In water, longer chains tend to behave similarly, attaining compact globular states, with smaller $R_e$’s.

![Figure 4.15. Comparison between last 100 ns average PCL end-to-end distance versus number of monomers for single PCL chains in POPC and in water (total 150 ns simulation time) and last 5 ns PCL end-to-end distance versus number of monomers for single PCL chains in water (total 50 ns simulation time).](image-url)
4.7 Membrane Thickness

In the absence of polymer chain, the solvated POPC membrane has very small fluctuations in thickness, between 38-40Å. The solvation of the membrane causes it to undergo slight undulatory motions around equilibrium (Figure 4.16). However, when single PCL chains are inserted into the membrane, the effect of the polymer chain on membrane thickness varies according to the chain length as shown in Figure 4.17 (a). Small chains cause some thinning of the membrane, from an average of 38.26Å in the absence of the polymer to an average of 35.29Å (5-monomer PCL chain) and 35.99Å (8-monomer PCL chain). However, we notably find that intermediate chains (10, 13PCL) cause a much more substantial membrane thinning as compared to shorter chains, from 32.84 Å thicknesses for a 10 monomer PCL, 33.20 Å thicknesses for a 13 monomer PCL. Other intermediate chains (17, 22 30PCL) display a membrane thickness in a range similar to shorter chains (35-36Å). Maximum thickness of the membrane of 37.95Å is achieved by 44PCL chain, followed by an abrupt decrease to 34.76Å for 58PCL. Membrane thinning is associated with chain conformation and movement through the membrane.
Figure 4.16. Time evolution of membrane thickness for various single PCL chains in solvated POPC membrane. a) no PCL b) 5PCL c) 8PCL d) 10PCL e) 13PCL f) 17PCL g) 22 PCL h) 30PCL i) 44PCL j) 58PCL.

Figure 4.17. a) last 100 ns average POPC membrane thickness versus number of monomers for single PCL chains in POPC; b) last 100 ns average POPC area/lipid versus number of monomers for single PCL chains in POPC.
Membrane thinning is associated with chain conformation and movement through the membrane. The intermediate chain lengths, which thin the membrane the most are partially absorbed to the interface, with characteristic chain ends that sometimes insert themselves into the hydrophobic core of the membrane (Figure 4.2). The polar head groups of the membrane are highly solvated, acting as a barrier that impedes the movement of PCL chains to move towards water region. The energy required to remove hydration water is higher than the interactions between POPC head groups and non-polar PCL chains, which explains why short chains, in extended conformations or long, gathered-disordered chains remain within the membrane, without crossing the membrane-water interface. However, the loose ends of intermediate chains, once they come in contact with the top potion of membrane’s hydrocarbon chains, can overcome the hydrophilic barrier of membrane’s head groups. The interaction of the PCL chain with the membrane interface is most clearly shown in Figure 4.18.
Figure 4.18. PCL chain positioning with respect to membrane-water interface at 150ns. a) no PCL b) 5PCL c) 8PCL d) 10PCL e) 13PCL f) 17PCL g) 22 PCL h) 30PCL i) 44PCL j) 58PCL.

4.8 Area per Lipid

The area per lipid as a function of polymer chain length is shown in Figure 4.17 (b). The time evolution is shown in Figure 4.19.

![Graphs showing time evolution of POPC area/lipid for various single PCL chains in solvated POPC membrane.](image)

Figure 4.19. Time evolution of POPC area/lipid for various single PCL chains in solvated POPC membrane. a) 5PCL b) 8PCL c) 10PCL d) 13PCL e) 17PCL f) 22 PCL g) 30PCL h) 44PCL i) 58PCL.

Here, we see that with the membrane thinning of intermediate chain lengths comes an increase in are per lipid. After an initial increase to 71.88 Å² (5 monomers PCL), 70.95Å² (8 monomers PCL) and 83.96Å² (10 monomers PCL) area per lipid values decrease for intermediate chains to 79.56Å² (13 monomers PCL) and 72.24Å² (17 monomers PCL). For larger chains, area per lipid has a tendency to plateau around ~77Å², after which area per lipid increases to 95.82 Å² for 58
monomers PCL. Clearly, such pattern indicates that with the thinning of the membrane at intermediate chain lengths comes an increase in area per lipid. However, this trend does not hold at larger chain lengths.

4.9 Density Profiles

We next computed mass density profiles for the last 10 ns of each trajectory (Figure 4.20).

As would be expected, the PCL chain addition shifts the mass distribution of the membrane. While the water density distribution changes very little, POPC distributions show a decrease on the location where PCL chain density increases. Indeed, the mass density of the POPC phospholipid headgroup decreases in the upper leaflet for all chain lengths. For 17 PCL, the

![](figure.png)

*Figure 4.20. Average mass density distributions for of the PCL (green), water (cyan) and POPC (magenta) along the Z axis of the simulation box, for the last 10 ns of simulations. a) 5PCL b) 8PCL c) 10PCL d) 13PCL e) 17PCL f) 22 PCL g) 30PCL h) 44PCL i) 58PCL.*
maximum PCL density at 15 Å of 0.08 amu/Å³ corresponds to a decrease in POPC density to 0.38 amu/Å³, from the initial value of 0.46 amu/Å³. A similar pattern is observed for 22 PCL, showing a maximum PCL density at 13 Å of 0.1 amu/Å³, for which POPC decreased to 0.38 amu/Å³, and 22 PCL, with PCL maximum density at 10 Å of 0.2 amu/Å³, where POPC has a local minimum of 0.36 amu/Å³. For larger chains (44, 58 PCL), the PCL polymer density spreads throughout the membrane (Figure 4.20 (h), (i)), while the density of the membrane as a whole decreases with the increase of PCL chain length. At the point of maximum 44 PCL chain density occurring at 11 Å of 0.14 amu/Å³, POPC density is 0.39 amu/Å³, and for 58 PCL, maximum chain density is at 7 Å of 0.2 amu/Å³, for which POPC density is 0.3 amu/Å³.

4.10 Tail Order Parameters

In order to measure the molecular order of the POPC bilayer in the presence of various PCL chain lengths, we computed deuterium tail order parameters (S_CD) for palmitoyl (sn-1) and oleoyl (sn-2) chains. The palmitoyl chain is fully saturated, while the oleoyl chain has a single double bond. In the absence of PCL chain, our results for sn-1 and sn-2 (absolute value) show a close resemblance to previous results obtained with CHARMM36 force field (178) and with experiments (185). However, in our simulations, chain sn-2 shows reduced values for the last three Carbon atoms: 0.1 (C16), 0.08 (C17) and 0.03 (C18). Addition of PCL chains causes shifts in sn-1 and sn-2 values (Figure 4.21).
For the upper and middle section of lipid chains (Carbon atoms 2-10), short PCL chains (5, 8 PCL, Figure 4.21 (a), (b)) sn-1 values gather around the sn-1 of POPC in the absence of PCL chain. In this plateau region, the tail order parameter varies slightly, which is typical for a phospholipid bilayer (178). The intermediate PCL chain lengths (10, 13, 17 PCL) that approach the membrane-lipid interface, have lower values in the same region (Figure 4.21 (c), (d), (e)), while larger chains (22, 30, 44 PCL) have slightly greater values than in the absence of PCL. Only for 58 PCL sn-1 in the same region is decreased significantly, more than in the case of intermediate chains (Figure 4.21 (i)). For the lower section (Carbon atom 11-15), sn-1 for shorter and intermediate chains overlap, but large chains (44, 58PCL) are distinct with sn-1 for 44 PCL above sn-1 for 58 PCL (Figure 4.22 (a)).
Lipid chain sn-2 shows patterns of increase and decrease with the number of PCL monomers (Figure 4.21). For Carbon atoms 2-7, nearly overlapped sn-2 values for 5, 8, 13, 22, 30, 44 PCL are above those of 10, 58 PCL (Figure 4.22 (b)). Overlapping merges for Carbon atoms 7-11, after which, (Carbon atoms 11-17) distinct patterns emerge according to the length of PCL chains, with sn-2 for 5PCL on top and sn-2 for 58PCL at the base. All sn-2 values for any PCL length in between does not follow a particular pattern. For the last portion (Carbon atoms 17-18), sn-2 values merge towards a value close to zero. Figure 4.22 (a) shows that the lowest sn-1 is for the largest chain (58 PCL) or for intermediate chains that approach the membrane-water interface (10, 13PCL). These chains produce most disorder in the lipid chains of POPC.

4.11 Solvent Accessible Surface Area

We also investigate the characteristic shape change of PCL chain, when solvated by the POPC membrane by calculations of Solvent Accessible Surface Area (SASA) using a probe sphere of radius of 1.4 Å. We characterized SASA time evolution over 150 ns, as shown in Figure 4.23.
Figure 4.23. Solvent accessible surface area (SASA) time evolution for all atomistic equilibrating systems of single PCL chains placed POPC membrane: a) 5PS; b) 8PS; c) 10PS; d) 13PS; e) 17PS; f) 2PS; g) 22PS; h) 30PS; h) 44PS; i) 58PS.

For short chains (5, 8, 10 PCL monomers) (Figure 4.23 (a), (b), (c)), SASA is nearly constant for all simulation time. Chains of intermediate length (13, 17, 22 PCL monomers) (Figure 4.23 (d), (e), (f)) show slight fluctuations in SASA, while for longer chains (30, 44, 58 monomers PCL) (Figure 4.23 (g), (h), (i)) SASA fluctuations are more pronounced. Average SASA for the last 100 ns (Table 4.5) versus number of monomers in each chain fitted to a power law <SASA> ∝ N^μ (Figure 4.24 (a)) shows a scaling coefficient μ = 0.702528 ± 0.04458 for PCL in POPC. The value of μ =2/3 was reported for collapsed protein structures, so for PCL in POPC, μ > 2/3, so the chain inside the membrane does not collapse, but rather progresses towards to a state of increased disorder and expansion. SASA’s standard deviation (Figure 4.24 (b)) increases with the number
of monomers, showing that for larger chains there are greater changes in the surface area of the hydrophobic polymer.

<table>
<thead>
<tr>
<th>PCL Length</th>
<th>&lt;SASA&gt; for last 100 ns (Å²)</th>
<th>δSASA (Å²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>1057.70</td>
<td>55.69</td>
</tr>
<tr>
<td>8</td>
<td>1599.50</td>
<td>85.57</td>
</tr>
<tr>
<td>10</td>
<td>1956.69</td>
<td>149.37</td>
</tr>
<tr>
<td>13</td>
<td>2545.07</td>
<td>103.74</td>
</tr>
<tr>
<td>17</td>
<td>3135.97</td>
<td>237.55</td>
</tr>
<tr>
<td>22</td>
<td>3641.45</td>
<td>234.62</td>
</tr>
<tr>
<td>30</td>
<td>3993.75</td>
<td>266.10</td>
</tr>
<tr>
<td>44</td>
<td>6269.57</td>
<td>452.47</td>
</tr>
<tr>
<td>58</td>
<td>6663.68</td>
<td>448.29</td>
</tr>
</tbody>
</table>

Table 4.5. Solvent accessible surface area averaged over last 100 ns of all atomistic MD simulations of equilibrated systems of 5, 8, 10, 13, 17, 22, 30, 44 and 58 single PCL chain lengths in POPC membrane, with standard deviations

Figure 4.24. a) Solvent accessible surface area (SASA) dependence of number of monomers N (including standard deviation) in all atomistic system equilibration for PCL chains placed in POPC membrane. b) Standard deviation in SASA versus number of monomers N computed for all atomistic non-accelerated MD simulations of equilibrated single PCL chains placed in POPC membrane

### 4.12 Metadynamics and free energy calculations

Next, given the limited sampling time of the previous simulations, we decide to perform advanced sampling simulations of three representative chain lengths interacting with the phospholipid membrane. We perform three sets of metadynamics simulations with three different
polymer chain lengths—short (5 monomers PCL), intermediate (17 monomers PCL), and long (58 monomers PCL)—interacting with a POPC phospholipid membrane. Two reaction coordinates are used: the radius of gyration of the polymer chain and the center of mass distance from the polymer chain to the center of mass of the phospholipid headgroups. The 2D free energy surfaces are shown for each polymer chain length in Figure 4.25.

Figure 4.25. All-atomistic metadynamics free energy distribution contour graphs. PCL placed in POPC membrane: a) 5PCL; b) 17PCL; c) 58PCL.
Metadynamics results indicate that short chains (5 PCL) have a localized minimum free energy distribution of 0-10 kcal/mol occurring at a distance between the center of mass of PCL and the POPC headgroups of 45 Å (Figure 4.25 (a), Table 4.6 (a)), decreasing with polymer length to 15 Å (17 monomers PCL) (Figure 4.25 (b), Table 4.6 (b)) and 5 Å (58 PCL) (Figure 4.25 (c), Table 4.6 (c)). These minimum all represent folded states of PCL chains, with a minimum radius of gyration, however we see that the minimum energy confirmation changes from the PCL chain sitting at the interface for short chains Figure 4.26 (a) to the globular state that is fully hydrated by the hydrophobic membrane for longer PCL chains as shown in Figure 4.26 (g). For short chains, there is one collapsed state at 0 kcal/mol, for which radius of gyration is 4.7 Å (Figure 4.26 (a)), with an additional similar state, with radius of gyration of 6.7 Å, at a higher energy of 14.66 kcal/mol (Figure 4.26 (b)). Intermediate chains (17 PCL) tend to extend the energy range of folded chain conformations in the same section of 0-10 kcal/mol. Three such states have been observed here (Figure 4.26 (c), (d), (e)) at 0 kcal/mol, 5.32 kcal/mol and 7.37 kcal/mol, with respective radii of gyration of 15.2 Å, 12.2 Å and 8.7 Å. However, the state with minimum radius of gyration of 7.7 Å occurs at a higher energy of 12.17 kcal/mol (Figure 4.26 (f)) and at a higher center of mass distance of 25 Å. Longer chains display a different pattern, with only one location for a local minimum energy configuration at 0 kcal/mol, for which the radius of gyration is 10.7 Å (Figure 4.26 (g)). This is a collapsed globular state for the PCL chain. Other low energies are in the range 10-20 kcal/mol (14-18 kcal/mol), with larger radii of gyration (14-21 Å). Such states are disordered (Figure 4.26 (h), (i)) and dispersed globules. To summarize, we find a transition of the minimum energy state for the polymer chain at the interface moving to the center of the hydrophobic core; we characterize
this shift in minimum energy configuration/location of the chain with 2D metadynamics calculations.

<table>
<thead>
<tr>
<th>5PCL</th>
<th>Free Energy</th>
<th>r&lt;sub&gt;min&lt;/sub&gt;</th>
<th>distance&lt;sub&gt;COM&lt;/sub&gt;</th>
<th>time</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Minimum FE, 0-10 kcal/mol (purple region)</td>
<td>(kcal/mol)</td>
<td>(Å)</td>
<td>(Å)</td>
<td>(ns)</td>
</tr>
<tr>
<td>0.00</td>
<td>4.7</td>
<td>45</td>
<td>134.9</td>
<td></td>
</tr>
<tr>
<td>b) Additional FE, 10-20 Kcal/mol (blue region)</td>
<td>14.66</td>
<td>6.7</td>
<td>45</td>
<td>7.1</td>
</tr>
</tbody>
</table>

*Table 4.6. a) Metadynamics: radius of gyration at minimum free energy for 5PCL. COM of PCL-POPC and time are also included.*

<table>
<thead>
<tr>
<th>17PCL</th>
<th>Free Energy</th>
<th>r&lt;sub&gt;min&lt;/sub&gt;</th>
<th>distance&lt;sub&gt;COM&lt;/sub&gt;</th>
<th>time</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Minimum FE, 0-10 kcal/mol (purple region)</td>
<td>(kcal/mol)</td>
<td>(Å)</td>
<td>(Å)</td>
<td>(ns)</td>
</tr>
<tr>
<td>0</td>
<td>15.2</td>
<td>15</td>
<td>60.8</td>
<td></td>
</tr>
<tr>
<td>b) Minimum FE, 0-10 kcal/mol (purple region)</td>
<td>5.32</td>
<td>12.2</td>
<td>15</td>
<td>103.9</td>
</tr>
<tr>
<td>c) Minimum FE, 0-10 kcal/mol (purple region)</td>
<td>7.48</td>
<td>8.7</td>
<td>15</td>
<td>213.9</td>
</tr>
<tr>
<td>d) Additional FE, 10-20 kcal/mol (blue region)</td>
<td>12.17</td>
<td>7.7</td>
<td>25</td>
<td>48.6</td>
</tr>
</tbody>
</table>

*Table 4.6. b) Metadynamics: radius of gyration at minimum energy for 17PCL. COM of PCL-POPC and time are also included.*

<table>
<thead>
<tr>
<th>58PCL</th>
<th>Free Energy</th>
<th>r&lt;sub&gt;min&lt;/sub&gt;</th>
<th>distance&lt;sub&gt;COM&lt;/sub&gt;</th>
<th>time</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Minimum FE, 0-10 kcal/mol (purple region)</td>
<td>(kcal/mol)</td>
<td>(Å)</td>
<td>(Å)</td>
<td>(ns)</td>
</tr>
<tr>
<td>0</td>
<td>10.7</td>
<td>5</td>
<td>35.3</td>
<td></td>
</tr>
<tr>
<td>b) Additional FE, 10-20 kcal/mol (blue region)</td>
<td>14.79</td>
<td>14.2</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>c) Additional FE, 10-20 kcal/mol (blue region)</td>
<td>18.39</td>
<td>21.2</td>
<td>5</td>
<td>121.6</td>
</tr>
</tbody>
</table>

*Table 4.6. c) Metadynamics: radius of gyration at minimum energy for 58PCL. COM of PCL-POPC and time are also included.*

**4.13 Summary**

Herein we described and characterized the biophysical interactions between a model phospholipid membrane bilayer composed by POPC and single polycaprolactone chains of varying lengths. First, we characterized the equilibrium interfacial location of the polymer in the membrane after 150 ns all-atomistic MD simulations. We also characterized the shape of the polymer, using radius of gyration. We find that the polymer shifts its behavior from shorter to intermediate to longer chains, moving from the membrane-water interface to the hydrophobic core.
of the membrane. We also characterized the molecular order of the POPC bilayer, showing that longer polymers display greater disruption of the phospholipid tails.

The interaction between the polymer and phospholipid tails consists of van der Waals attraction between polymer segments, the polymer and the membrane, as well as steric and electrostatic interactions between the PCL chain and the membrane that tend to expand the polymer. It is found that intermediate chains cause a membrane thinning due to a lateral pressure of the chain, which sits near the interface, expanding the membrane. However, as the chain becomes longer the polymer prefers to sit inside the center of the hydrophobic core of the membrane, which does not lead to a membrane thinning effect. Likely, the elasticity of the membrane is also affected by the presence of these polymer chains at the interface or the center of the hydrophobic core of the membrane. However, this would need to be further characterized and would depend on the concentration of polymer interacting with the membrane.

In order to improve the sampling of the conformation of the polymer chain and characterize the interaction free energy in detail, we next utilized metadynamics, an accelerating sampling technique. The results confirm that the minimum in the free energy surface shifts from the membrane-water interface to the center of the hydrophobic core of the membrane as a function of chain length. Likely, these results would also be concentration dependent on the number of polymers interacting with the hydrophobic membrane. Indeed, multiple polymers interacting with the membrane may porate and disrupt the membrane entirely due to aggregation of the hydrophobic polymer in the center of the hydrophobic core of the polymer. However, this effect may depend on the chain length of the polymer, as suggested here in these large-scale atomistic simulations.
CHAPTER 5

Coarse-Grained Molecular Dynamics Simulations of Phospholipid Bilayers with Polycaprolactone

5.1. Coarse-grained models

All-atom molecular dynamics (MD) simulations are used to build a coarse grain (CG) model of polycaprolactone (PCL) in solvated phospholipid bilayer (POPC) to examine the morphological dependence of POPC with PCL chains, along with hydrophobic collapse of PCL chains, PCL is a biodegradable, hydrophobic polymer widely used in biomaterial applications (169) due to its gradual hydrolysis to monomers. (186).

CG models must maintain certain physical properties to effectively describe physical properties such as polymer hydrophobicity. To be chemically exact, the solubilities and flexibilities of each polymer chain are defined accurately, such that proper balance between them is achieved accordingly for self-assembly. (143,187). As PCL is semi-crystalline (171), its phase behavior can be complicated, depending on the method of POPC bilayer preparation. The methods of analysis quantify PCL chain dispersion in the membrane and water.

Figure 5.1. Coarse grain models of PCL. Based on all-atom computations, polymer is modeled three CG beads per caprolactone monomer (S, A, B) plus one interfacial CG bead (R). (Schematically drawn from Loverde SM, Klein ML, Discher DE. Nanoparticle Shape Improves Delivery: Rational Coarse Grain Molecular Dynamics (rCG-MD) of Taxol in Worm-Like PEG-PCL Micelles. Advanced materials. 2012 Jul 24;24(28):3823-30.)
CG flexibility parameters were obtained from AA simulations of PCL chains at 1Atm and 300K. In CG model, each molecule of PCL contains 2-5 beads (Figure 5.1, Table 5.1), according to chemistry and symmetry of its structure. As a result, intermolecular potentials are separated into chemical components (188) and AA intramolecular potentials are optimized (123,189)

<table>
<thead>
<tr>
<th>CG Monomer</th>
<th>Chemical Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>(H₂O)₃</td>
</tr>
<tr>
<td>R</td>
<td>-CH₂ O-</td>
</tr>
<tr>
<td>S</td>
<td>-CO CH₂-</td>
</tr>
<tr>
<td>A</td>
<td>-CH₂CH₂CH₂-</td>
</tr>
<tr>
<td>B</td>
<td>HOCH₂-</td>
</tr>
</tbody>
</table>

Table 5.1. The mapping for water and the mapping for PCL, consisting of 3 different CG beads (B, A, S) for every caprolactone (CL) monomer. (Schematically drawn from Loverde SM, Klein ML, Discher DE. Nanoparticle Shape Improves Delivery: Rational Coarse Grain Molecular Dynamics (rCG-MD) of Taxol in Worm-Like PEG-PCL Micelles. Advanced materials. 2012 Jul 24;24(28):3823-30.)

5.1.1 CG-model development for PCL

Short range interactions, mapped from small chemical groups (Table 5.1) are described by a 9-6 type Lennard-Jones potential (190), given that parameters σ and ε are obtained using surface tension and density for the chemical mapping of each bead type:

\[ V(r)_{9-6} = \frac{27}{4} \varepsilon \left( \frac{\sigma^9}{r^9} - \frac{\sigma^6}{r^6} \right) \]  \hspace{1cm} (5.1)

where combination rules Lorentz-Bethelot (191,192) provides numerical values for σ and ε:

\[ \sigma_{AB} = \frac{\sigma_{AA} + \sigma_{BB}}{2} \]  \hspace{1cm} (5.2)

\[ \varepsilon_{AB} = \varepsilon_A \varepsilon_B \]  \hspace{1cm} (5.3)

Bond and angular interactions between CG beads are obtained from AA simulations using a potential based on the CHARMM27 (143) force-field of a short PCL melt. Intramolecular interactions are modeled via harmonic potentials (3.1, 3.2), for which force constants are obtained using an inverse Boltzmann technique (3.3, 3.4).
5.1.2 CG-model development for POPC

In this approach of POPC coarse-graining, (Figure 5.2) all atom functional groups are assigned CG segment names as listed in Table 5.2. Water model consists of three water molecules packed into one CG site. Unrealistic water freezing is prevented by functional forms (193). CG resolution level is determined by the size of CG water particle. Thus, CG sampling efficiency is increased by three orders of magnitude than in the case of AA systems.

CG particle represents multiple molecules, which is a disadvantage to the present model in the sense that a reasonable structure factor cannot be expected. Moreover, using force-matching method to build CG model is not suitable for the situation (136). Most MD simulations in which water interaction are not primary interest are dominated by water-water interactions calculations. Thus it is preferred that CG water particle to be coarser, as the main interest is on lipids such that CG water model must have hydrodynamic interactions between surfactant aggregates; its liquid phase should be is in the range of 1-100 °C; its density and surface tension must be correct; must
reproduce a correct free energy transfer of solutes. Lennard-Jones 12-4 potential meets these conditions (136).

In this parametrization inter- and intra particle potentials have simple functional forms. Intramolecular potentials are characterized by harmonic bond stretching and angle bending potential specific to 1-2 and 1-2-3 bonded pairs. Pairs separated by more than two bond are subject to nonbonded potentials. Interactions of type 1-3 are corrected by shifting a nonbonded interaction to give a zero potential $r = r_s$:

$$U_{1-2} = \sum_{\text{bond}} K_b (r_{ij} - r_0)^2$$  \hspace{1cm} (5.4)

$$U_{1-2-3} = \sum_{\text{angle}} K_{\theta} (\theta_{ijk} - \theta_0)^2$$  \hspace{1cm} (5.5)

$$U_{1-3}^\text{correct} = \sum_{1-3} \{U_{nb}(r_{ij}) - U_{nb}(r_s)\}, r_{ij} < r_s$$  \hspace{1cm} (5.6)

The term in equation (5.6) serves to prevent bending in CG models, in which bending force constant is a quantity comparable with potential depth of nonbonded interactions. Bending force can cause angle collapse, in which 1-4 nonbonded potential is much greater than 1-2-3 bending potential. The correction term can be included in cases or artificial angle collapse only, taking the $r_s$ as the distance of LJ minimum potential ($r_{\text{min}}$) of nonbonded interactions, without affecting angle distribution.

Previously developed CG models (120,121) can be used here unmodified, except for the rare case in which 1-3 pair CG sites are closer than $r_{\text{min}}$. All atomistic simulations provided bond and angular distributions to parametrize force constants. Nonbonded interaction potential is represented by LJ functions of type (12-4) and (9-6):

$$U_{nb}(r_{ij}) = \frac{3}{2} \varepsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right], W \text{ pairs}$$  \hspace{1cm} (5.7)
\[ U_{nb}(r_{ij}) = \frac{27}{4} \varepsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^9 - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^2 \right] \text{, other pairs} \quad (5.8) \]

All atomistic simulations provided structural data of lipid assemblies to optimize lipid headgroup parameters. The current lipid model contains lipid headgroup CG sites (PH, NC, and NH) with partial charges. The electric potential between particles was calculated using Ewald sum and Coulomb functional form (122):

\[ U_c(r_{ij}) = \frac{q_i q_j}{4\pi\varepsilon_0 \varepsilon_r r_{ij}} \quad (5.9) \]

where \( q_i \) is the partial charge of ith particle, \( \varepsilon_0 \) is the permittivity in vacuum, and \( \varepsilon_r \) is relative permittivity due to screening effects of water. Coulomb potential was evaluated without truncation, to ensure that an adequate ionic distribution determines a plausible interaction between charged surfaces (lipid monolayer and water interface) at large separation.

<table>
<thead>
<tr>
<th>CG Segment</th>
<th>Atoms</th>
<th>Charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH</td>
<td>-CH_2CH_2-N-(CH_3)_3</td>
<td>+1</td>
</tr>
<tr>
<td>PH</td>
<td>-PO_4^-</td>
<td>-1</td>
</tr>
<tr>
<td>GL</td>
<td>-CH_2CH-CH_2-</td>
<td>0</td>
</tr>
<tr>
<td>EST1</td>
<td>-CH_2CO_2^-</td>
<td>0</td>
</tr>
<tr>
<td>EST2</td>
<td>-H_2CO_2^-</td>
<td>0</td>
</tr>
<tr>
<td>CM</td>
<td>-CH_3CH_2CH_2-</td>
<td>0</td>
</tr>
<tr>
<td>CMD2</td>
<td>-HC=CH-</td>
<td>0</td>
</tr>
<tr>
<td>CT2</td>
<td>CH_3CH_2-</td>
<td>0</td>
</tr>
<tr>
<td>W</td>
<td>(H_2O)_3</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 5.2. CG Segment Name and the Corresponding All-Atomic Functional Group

5.1.3 PCL in solvated membrane

Single CG PCL chains of sizes 5, 8, 10, 13, 17, 22, 30, 44, 58, 72, 86 and 100 monomers were first equilibrated in water for up to 20 ns. The POPC membrane, consisting of 236 CG lipids was
solvated in 7024 CG waters. Next PCL chains were placed in water above the membrane. All system sizes and set-ups are included in Table 5.3 Values for inter- and intramolecular potentials are shown in Tables 5.4-5.5.

<table>
<thead>
<tr>
<th>PCL Length</th>
<th>$&lt;R_g&gt;$ for last 100 ns (Å)</th>
<th>$\delta R_g$</th>
<th>Box Size</th>
<th>Water Beads</th>
</tr>
</thead>
<tbody>
<tr>
<td>no PCL</td>
<td>96.57x75.15x125.39</td>
<td>7024</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6.91</td>
<td>0.65</td>
<td>96.57x75.15x125.39</td>
<td>7024</td>
</tr>
<tr>
<td>8</td>
<td>9.92</td>
<td>1.05</td>
<td>96.57x75.15x125.39</td>
<td>7024</td>
</tr>
<tr>
<td>10</td>
<td>12.40</td>
<td>1.32</td>
<td>96.57x75.15x125.39</td>
<td>7024</td>
</tr>
<tr>
<td>13</td>
<td>15.67</td>
<td>1.41</td>
<td>96.57x75.15x125.39</td>
<td>7024</td>
</tr>
<tr>
<td>17</td>
<td>17.36</td>
<td>2.21</td>
<td>96.57x75.15x125.39</td>
<td>7024</td>
</tr>
<tr>
<td>22</td>
<td>23.46</td>
<td>2.53</td>
<td>96.57x75.15x125.39</td>
<td>7024</td>
</tr>
<tr>
<td>30</td>
<td>27.75</td>
<td>3.63</td>
<td>96.57x75.15x125.39</td>
<td>7024</td>
</tr>
<tr>
<td>44</td>
<td>22.54</td>
<td>3.83</td>
<td>96.57x75.15x125.39</td>
<td>7024</td>
</tr>
<tr>
<td>58</td>
<td>32.02</td>
<td>4.89</td>
<td>96.57x75.15x125.39</td>
<td>7024</td>
</tr>
<tr>
<td>72</td>
<td>26.23</td>
<td>1.63</td>
<td>96.57x75.15x125.39</td>
<td>7024</td>
</tr>
<tr>
<td>86</td>
<td>19.23</td>
<td>2.39</td>
<td>96.57x75.15x125.39</td>
<td>7024</td>
</tr>
<tr>
<td>100</td>
<td>31.46</td>
<td>2.33</td>
<td>96.57x75.15x125.39</td>
<td>7024</td>
</tr>
</tbody>
</table>

*Table 5.3 System sizes and setups.*
<table>
<thead>
<tr>
<th>Bond</th>
<th>Atoms</th>
<th>Spring Constant (kcal/mol/Å²)</th>
<th>Equilibrium Distance $r_0$ (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CH-PH</td>
<td>4.8</td>
<td>4.25</td>
</tr>
<tr>
<td>2</td>
<td>PH-GL</td>
<td>8.9</td>
<td>3.52</td>
</tr>
<tr>
<td>3</td>
<td>GL-EST1</td>
<td>30</td>
<td>2.88</td>
</tr>
<tr>
<td>4</td>
<td>EST1-CM</td>
<td>4.7</td>
<td>3.55</td>
</tr>
<tr>
<td>5</td>
<td>CM-CM</td>
<td>6.16</td>
<td>3.64</td>
</tr>
<tr>
<td>6</td>
<td>CM-CMD2</td>
<td>8</td>
<td>3.03</td>
</tr>
<tr>
<td>7</td>
<td>CMD2-CM</td>
<td>8</td>
<td>3.03</td>
</tr>
<tr>
<td>8</td>
<td>CM-CM</td>
<td>6.16</td>
<td>3.64</td>
</tr>
<tr>
<td>9</td>
<td>CM-CT2</td>
<td>9</td>
<td>3.13</td>
</tr>
<tr>
<td>10</td>
<td>GL-EST2</td>
<td>8.4</td>
<td>3.48</td>
</tr>
<tr>
<td>11</td>
<td>EST2-CM</td>
<td>5.1</td>
<td>3.61</td>
</tr>
<tr>
<td>12</td>
<td>CM-CM</td>
<td>6.16</td>
<td>3.64</td>
</tr>
<tr>
<td>13</td>
<td>CM-CM</td>
<td>6.16</td>
<td>3.65</td>
</tr>
<tr>
<td>14</td>
<td>CM-CM</td>
<td>6.16</td>
<td>3.65</td>
</tr>
<tr>
<td>15</td>
<td>CM-CT2</td>
<td>9</td>
<td>3.13</td>
</tr>
<tr>
<td>16</td>
<td>R-S</td>
<td>27.98</td>
<td>2.92</td>
</tr>
<tr>
<td>17</td>
<td>S-A</td>
<td>9.4</td>
<td>2.9</td>
</tr>
<tr>
<td>18</td>
<td>A-B</td>
<td>16.2</td>
<td>2.7</td>
</tr>
<tr>
<td>19</td>
<td>B-S</td>
<td>26.42</td>
<td>2.8</td>
</tr>
</tbody>
</table>

*Table 5.4. Bond stretching parameters*
<table>
<thead>
<tr>
<th>Angles</th>
<th>Atoms</th>
<th>Prefactor $K_a$ ($\text{kcal/mol/radian}^2$)</th>
<th>Equilibrium Angle $\theta_0$ (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CH-PH-GL</td>
<td>3.1</td>
<td>112</td>
</tr>
<tr>
<td>2</td>
<td>PH-GL-EST1</td>
<td>1.4</td>
<td>124</td>
</tr>
<tr>
<td>3</td>
<td>GL-EST1-CM</td>
<td>0.8</td>
<td>168</td>
</tr>
<tr>
<td>4</td>
<td>EST1-CM-CM</td>
<td>1</td>
<td>178</td>
</tr>
<tr>
<td>5</td>
<td>CM-CM-CMD2</td>
<td>1.9</td>
<td>161</td>
</tr>
<tr>
<td>6</td>
<td>CM-CMD2-CM</td>
<td>6</td>
<td>110</td>
</tr>
<tr>
<td>7</td>
<td>CMD2-CM-CM</td>
<td>1.9</td>
<td>161</td>
</tr>
<tr>
<td>8</td>
<td>CM-CM-CT2</td>
<td>1.6</td>
<td>172</td>
</tr>
<tr>
<td>9</td>
<td>PH-GL-EST2</td>
<td>2</td>
<td>138</td>
</tr>
<tr>
<td>10</td>
<td>GL-EST1-CM</td>
<td>0.8</td>
<td>172</td>
</tr>
<tr>
<td>11</td>
<td>EST2-CM-CM</td>
<td>1</td>
<td>178</td>
</tr>
<tr>
<td>12</td>
<td>CM-CM-CM</td>
<td>1.19</td>
<td>173</td>
</tr>
<tr>
<td>13</td>
<td>CM-CM-CM</td>
<td>1.19</td>
<td>173</td>
</tr>
<tr>
<td>14</td>
<td>CM-CM-CMT2</td>
<td>0.8</td>
<td>172</td>
</tr>
<tr>
<td>15</td>
<td>R-S-A</td>
<td>3.4</td>
<td>2.67</td>
</tr>
<tr>
<td>16</td>
<td>S-A-B</td>
<td>3.3</td>
<td>2.29</td>
</tr>
<tr>
<td>17</td>
<td>A-B-S</td>
<td>3.2</td>
<td>2.27</td>
</tr>
<tr>
<td>18</td>
<td>B-S-A</td>
<td>1.2</td>
<td>2.9</td>
</tr>
</tbody>
</table>

Table 5.5. Angle bending parameters.

5.1.4 Simulation parameters

All CG simulations were run using LAMMPS (216), a parallel molecular dynamics code developed by Sandia National Laboratory, with a timestep of 10 fs. The temperature and the pressure were controlled using the Nose-Hoover (154) algorithm at 300K and 1atm. For the electrostatic interactions we used and particle-particle-particle mesh (PPPM) (155). Lennard-Jones pair potential cut-of distance was set to 15 Å. The system was run initially in the canonical (NVT) ensemble for 1 ns at 300 K, then for 150 ns in isothermal-isobaric (NPT) ensemble also at 300 K and 1 bar pressure with a drag factor of 0.01 added to the barostat. The neighbor list for non-bonded interactions was updated at every step.
5.2. Coarse-grained PCL-POPC interactions

We describe results of coarse-grained molecular dynamics simulations performed for of single polycaprolactone (PCL) chains of varying lengths in solvated POPC lipid (Figure 5.3, Figure 5.4, Figure 5.5). We compare these results with all–atomistic simulations. We investigate the nature of PCL chains and POPC membrane conformation changes as a result of the repulsive interactions between PCL chains, POPC membrane and water.

![Figure 5.3 A Snapshots of coarse-grained PCL in POPC membrane before simulation (a, d, g), at 75ns (d, e, f) and at 150ns (g, h, i) for 5 PCL (a, b c), 17 PCL (d, e, f) and 58PCL (g, h, i). The polymer chain is displayed in a van der Waals representation while POPC membrane chains are displayed in a CPK representation. Water is indicated by transparent conformation background.](image_url)
Figure 5.4. Snapshots of coarse grained PCL in POPC membrane before simulation (a, d, g), at 75ns (d, e, f) and at 150ns (g, h, i) for 72 PCL (a, b, c), 86 PCL (d, e, f) and 100PCL (g, h, i). The polymer chain is displayed in a van der Waals representation while POPC membrane chains are displayed in a CPK representation. Water is indicated by transparent conformation background.

Figure 5.5. Snapshots of coarse grained PCL in solvated POPC membrane at the end of 150 ns simulation PCL: a) 5 PCL; b) 8 PCL; c) 10 PCL; d) 13 PCL; e) 17 PCL; f) 22 PCL; g) 30 PCL; h) 44 PCL; i) 58PCL; j) 72PCL; k) 86 PCL; l) 100PCL. The polymer chain, potassium and chloride ions are displayed in a van der Waals representation while POPC membrane chains are displayed in a CPK representation. Water is indicated by transparent conformation background.
5.3 PCL and POPC analysis tools

In greater part, analysis tools consist in a replication of all atomistic tools in coarse grain condition, which served as a method of comparing the results.

From the trajectories, we computed time evolution of radius of gyration, end-to-end-distance, membrane thickness, area per lipid and surface accessible surface area. For thickness of the membrane we used MEMBPLUG IN (181) run within the Visual Molecular Dynamics (VMD) (182). In addition, run in the same environment, we used Density Profile tool to compute density profiles for PCL chains, POPC and water. We calculated mass density profiles for PCL, membrane and water along z-axis (membrane normal is parallel with z-axis), with a resolution $\Delta z = 2\text{Å}$, which is the thickness of equally divided bins that along z-axis, where the z-size of the simulation cell is 125.39Å, x-size=96.57Å and y-size=2.0Å. The volume of the extended bin is then 96.57x75.15x2 Å³. Thus the mass density is calculated for each 3-D bin, as

$$\rho = \frac{m[\text{amu}]}{V[\text{Å}^3]} = \frac{m[\text{amu}]}{4,285.44 [\text{Å}^3]}$$

where amu masses for PCL, water and POPC are given in the input file of each system setup.

Next we computed averages for the last 100ns of simulation time and plotted these average versus numbers of monomers as power fit. The numerical results were able to confirm all atomistic results and provide additional information about the stability of collapsed states of PCL chains.

5.3.1 Radius of gyration

Time evolution (Figure 5.6) of radius of gyration shows various patterns. According to PCL chain length. Small chains (5, 8, 10, 13PCL) show small fluctuation around a constant value, indicating that the chain does not undergo significant conformational change during simulation. Intermediate chains (17, 22, 30, 44 PCL) show larger fluctuations between minimum and
minimum R_g values than larger chains (58, 72, 86, 100PCL). This indicate that larger chains attain a relative compact shape stability after descending into membrane from water, while intermediate chains continue their motion in a disperse state. AA comparison (Figure 5.7) shows larger R_g values for larger chains (22, 30, 44, 58PCL), indicating that the state of disordered coil for larger chains in achieved faster in CG simulations.

Figure 5.6. Time evolution of radius of gyration for single CG PCL chains in solvated POPC membrane. a) 5PCL; b) 8PCL; c) 10PCL; d) 13PCL; e) 17PCL; f) 22 PCL; g) 30PCL; h) 44PCL; i) 58PCL; j) 72PCL; k) 86PCL; l) 100PCL.
Figure 5.7. AA-CG comparison for \( R_g \) time evolution single CG PCL chains in solvated POPC membrane. a) 5PCL; b) 8PCL; c) 10PCL; d) 13PCL; e) 17PCL; f) 22 PCL; g) 30PCL; h) 44PCL; i) 58PCL.

Last 100 ns \( R_g \) averages versus number of monomers (Figure 5.8) indicate that shorter chains (5, 8, 10, 13, 17 PCL) follow closely similar patterns in both AA and CG simulations. Intermediate and large chains follow irregular patterns of increase in \( R_g \) with number of monomers, where CG values for \( R_g \) are much larger than AA analog. A power fit for \( R_g \) as a function of \( N \), \( R_g (N) = 4.35N^{0.42} \), give a scaling coefficient \( \nu \) in the range \( 1/2 < 0.42 < 1/3 \), a state between disordered coil and collapsed globule. A satisfactory power fit could not be obtained for AA simulations.
Figure 5.8. AA-CG comparison of last 100 ns average PCL radius of gyration versus number of monomers for single PCL chains in POPC membrane.

5.3.2 End-to-end distance

Time evolution (Figure 5.9) of $R_e$ for shorter chains (5, 8, 10, 13PCL) show nearly constant values with light fluctuation, indicating that small chains tend to maintain their conformations throughout the simulation. Intermediate and long chains, however, display large changes in end-to-end distance, characteristic for disordered states, where chains are partially extended or collapsed. Compared with AA simulations (Figure 5.10), $R_e$ is greater for longer chains.
Figure 5.9. Time evolution of end-to-end distance for single CG PCL chains in solvated POPC membrane.

a) 5PCL; b) 8PCL; c) 10PCL; d) 13PCL; e) 17PCL; f) 22 PCL; g) 30PCL; h) 44PCL; i) 58PCL; j) 72PCL; k) 86PCL; l) 100PCL.
Figure 5.10. AA-CG comparison for Re time evolution of single CG PCL chains in solvated POPC membrane. 

- a) 5PCL;
- b) 8PCL;
- c) 10PCL;
- d) 13PCL;
- e) 17PCL;
- f) 22 PCL;
- g) 30PCL;
- h) 44PCL;
- i) 58PCL.

Last 100 ns Re averages (Figure 5.11) AA-CG comparison shows that shorter chains (5, 8, 10, 13, 17PCL) are quite similar, but starting with 22PCL, in AA simulation, Re is decreasing, while increasing. This account for previously explained behavior of AA intermediate chains, which approach POPC-water interface in a state of partial collapse, with a smaller Re than shorter and larger chains. CG simulations show only one abrupt decrease in average Re (86PCL). The CG power law fit, $R_e (N) = 13.87N^{0.35}$ gave a scaling coefficient close to 1/3, as in collapsed states, which could not be obtained in the case of AA simulations.
5.3.3 Membrane thickness

Membrane thickness remains practically unchanged during CG simulations (Figure 5.12), showing that the lipid membrane is not affected by the presence of PCL chain, regardless of its size, as opposed to membrane thickness in AA simulations (Figure 5.13), where there are significant changes in the thickness for the membrane.
Figure 5.12. Time evolution of membrane thickness for single CG PCL chains in solvated POPC membrane.
a) 5PCL; b) 8PCL; c) 10PCL; d) 13PCL; e) 17PCL; f) 22 PCL; g) 30PCL; h) 44PCL; i) 58PCL; j) 72PCL; k) 86PCL; l) 100PCL.
Figure 5.13. AA-CG comparison for membrane thickness time evolution of single CG PCL chains in solvated POPC membrane. a) 5PCL; b) 8PCL; c) 10PCL; d) 13PCL; e) 17PCL; f) 22 PCL; g) 30PCL; h) 44PCL; i) 58PCL.

Last 100ns averages show a very slight increase in membrane thickness with numbers of monomers, while in AA simulation membrane thickness varies in irregular patterns.

Figure 5.14. AA-CG comparison of last 100 ns average PCL membrane thickness versus number of monomers for single PCL chains in POPC membrane.
5.3.4 Area per lipid

Area per lipid increases slightly with the number of monomer during 150ns simulation time (Figure 5.15), unlike during AA simulations (Figure 5.16).

Figure 5.15. Time evolution of area per lipid for single CG PCL chains in solvated POPC membrane. a) 5PCL; b) 8PCL; c) 10PCL; d) 13PCL; e) 17PCL; f) 22 PCL; g) 30PCL; h) 44PCL; i) 58PCL; j) 72PCL; k) 86PCL; l) 100PCL.
Figure 5.16. AA-CG comparison for area per lipid time evolution of single CG PCL chains in solvated POPC membrane. a) 5PCL; b) 8PCL; c) 10PCL; d) 13PCL; e) 17PCL; f) 22 PCL; g) 30PCL; h) 44PCL; i) 58PCL.

Last 100ns average for area per lipid (Figure 5.17) are far below the AA values. Area per lipid is not affected by the presence of polymers.

Figure 5.17. AA-CG comparison of last 100 ns average area per lipid versus number of monomers for single PCL chains in POPC membrane.
5.3.5 Density profiles

As the chains descend from water and move through the membrane, theory assume certain position in the membrane according to their length. Last 10 ns density profiles indicate that small chains (5, 8, 10PCL) remain in the upper part of membrane (Figure 5.18). Intermediate chains length (13-30PCL) density has peaks in the lower part of the membrane. Larger chains densities (44-100PCL) had multiple peaks in upper and lower part of the membrane, indicating a tendency to spread during simulation (Figure 5.5). Compared with AA simulations, PCL densities (Figure 5.19) are lower than AA densities. PCL is distributed in all membrane space due to spreading and extending, which has not been observed for all chains during AA simulations. POPC (Figure 5.20) and water (Figure 5.21) densities are nearly the same for AA and CG simulations.
Figure 5.18. Average CG mass density distributions for of the PCL (green), water (cyan) and POPC (magenta) along the Z axis of the simulation box, for the last 10 ns of simulations. a) 3PCL; b) 8PCL; c) 10PCL; d) 13PCL; e) 17PCL; f) 22 PCL; g) 30PCL; h) 44PCL; i) 58PCL; j) 72PCL; k) 86PCL; l) 100PCL.
Figure 5.19. AA-CG average PCL mass density distributions for the PCL (green), water (cyan) and POPC (magenta) along the Z axis of the simulation box, for the last 10 ns of simulations. a) 5PCL; b) 8PCL; c) 10PCL; d) 13PCL; e) 17PCL; f) 22PCL; g) 30PCL; h) 44PCL; i) 58PCL.
Figure 5.20. AA-CG average POPC mass density distributions for of the PCL (green), water (cyan) and POPC (magenta) along the Z axis of the simulation box, for the last 10 ns of simulations. a) 5PCL; b) 8PCL; c) 10PCL; d) 13PCL; e) 17PCL; f) 22 PCL; g) 30PCL; h) 44PCL; i) 58PCL.
Figure 5.21. AA-CG average water mass density distributions for of the PCL (green), water (cyan) and POPC (magenta) along the Z axis of the simulation box, for the last 10 ns of simulations. a) 5PCL; b) 8PCL; c) 10PCL; d) 13PCL; e) 17PCL; f) 22 PCL; g) 30PCL; h) 44PCL; i) 58PCL.

5.3.6 Solvent accessible surface area (SASA)

As explained in chapter 4 for AA simulations, for a single hydrophobic polymer chain collapse of N monomers, $\text{SASA} \propto N^{2\nu}$, in which $\nu = 1/3$ for an ideal polymer in bad solvent, and so $\text{SASA} \propto N^{2/3}$. We will determine SASA scaling coefficient for coarse-grained PCL in POPC to assess the type of shape the polymer takes as a result of interaction with the POPC membrane and water and how it is different from AA simulations.

SASA is constant in time for most of the chains (Figure 5.22), except for 44, 86 and 100PCL, in which case SASA increases for the last part of simulation. Compared to AA simulations (Figure 2.23), CG SASA is greater for all PCL chains, but even more for larger chains.
Figure 5.22. Time evolution of SASA for single CG PCL chains in solvated POPC membrane.

a) 5PCL; b) 8PCL; c) 10PCL; d) 13PCL e); 17PCL; f) 22 PCL; g) 30PCL; h) 44PCL; i) 58PCL; j) 72PCL;

k) 86PCL; l) 100PCL.
Figure 5.23. AA-CG comparison for SASA time evolution of single CG PCL chains in solvated POPC membrane. a) 5PCL; b) 8PCL; c) 10PCL; d) 13PCL; e) 17PCL; f) 22PCL; g) 30PCL; h) 44PCL; i) 58PCL.

The last 100ns averages versus numbers of monomer power fit, \( \text{SASA}(N) = 366.53N^{0.86} \) (Figure 5.24(a)) gives a larger scaling coefficient than AA simulations, \( \mu > 2/3 \), indicating greater disorder or extending of the chain, more than in the case of AA simulations. SASA’s standard deviation (Figure 5.24 (b)) increases with the number of monomers, much more than AA case. SASA data points for large chains are greatly spread apart over a wider range of values; with a large PCL, surface area is accessible to POPC, due to great dispersion of chain sin the membrane.
5.4 Conclusions

The AA-CG 150 ns simulations comparison was effective in demonstrating the extending and spreading the behaviour of PCL in a lipid membrane. Qualitatively, as illustrated in snapshots in Figure 5.1 and 5.2 and quantitatively, by the numerical values of $R_g$, $R_e$, and SASA scaling coefficients and density distributions, CG simulation analysis described the conformations of PCL chains as a result of motion through a solvated POPC membrane.

CG analysis was able to show details about the behaviour of PCL in the membrane, but not explicitly about the membrane itself. Membrane thickness changes very little with PCL insertions (Figure 5.14 (b)), and area per lipid is not affected by PCL chains, irrespective of its size (Figure 5.17), according to the isotropy condition applied to all systems. By contrast, AA simulations, in anisotropic conditions, show distinctive changes in the membrane.
CHAPTER 6

Summary and Prospects

6.1. Summary

During the course of this thesis, we examine the collapse of hydrophobic polymers in two different environments: poor solvent (water) and the hydrophobic core of model cell membranes. The understanding of the varying morphologies of polymer in different solvent environments, as well as their interaction with model membranes, is essential to the design of new polymer chemistries, both single component and multicomponent. Furthermore, understanding the behaviour of single polymers in their unique environment is the first step in understanding the self-assembly behaviour of more concentrated systems. Such behaviours are better understood within the framework of computer simulations acting as a complement to theory, as described in this thesis.

Initially, molecular dynamics techniques are used to describe hydrophobic collapse of polystyrene and polycaprolactone in water. The polymer size/shape in its collapsed state is compared to simple scaling theories. Two different simulation models are used and compared: an all-atomistic model, in which the collective behaviour of the atoms allows us to understand in detail the behaviour of the solvent (water) close to the polymer chain. Next, coarse-grained models are used, which are computational models that replicate the behaviour of a complex system by breaking it down into simpler sub-components, with a specified degree of granularity for the system in question. It is found that both models provide accurate description of the behaviour of hydrophobic polymers, but in terms of computational efficiency, the coarse-grained model is more advantageous than all-atomistic models. Moreover, the lower resolution of the coarse-grain polymer and water models allows for observing additional conformations otherwise not accessible.
in all-atomistic simulations.

In the preceding chapters, a model of polymer in a solvated phospholipid bilayer is also explored by these same two methods, atomistic and coarse-grain. For AA simulations of short polycaprolatone (PCL) chains in water, the chains collapse to globular conformations. For PCL chains in POPC, assuming extended conformations, the membrane remains virtually unchanged, while intermediate length chains, in their partially collapsed/partially extended conformations approaching membrane-water interface, cause thinning of the membrane. Longer chains possess a different behaviour, progressing toward states of disordered globules, without approaching the membrane-water interface. As a consequence, they do not cause as much membrane deformation as intermediate chains. The membrane thickness in those cases is comparable to values as for short chains insertions. Compared to PCL chains globular collapse behaviour in water, POPC membrane environment causes the chain to assume extended and/or disordered globular conformations. Coarse-grained simulations provide additional insight into the dependence of membrane deformations on polymer chain length and the subsequent chain conformations inside the membrane during simulations. Metadynamics simulations for all atomistic simulations show that short and large chains display different behaviour in the center of the hydrophobic membrane. Short chains have a minimum energy configuration closest to the membrane-water interface while longer chains have a minimum energy configuration in the center of the hydrophobic membrane.

6.2. Prospects

There are many directions to continue with this work as a precursor. For example, elastic properties of the POPC bilayer can be derived from CG simulations (119) using a trajectory. Area expansion module, \( K_A \) can be calculated from linear response theory (194):
where \( A \) is the averaged cross-sectional area per lipid and \( N \) is the number of lipid molecules. Bending modulus also can be calculated for a sufficiently large simulated system in which membrane fluctuations are analysed in terms of undulatory and peristaltic modes.

There are several applications that can be explored for POPC, using CG molecular model and MD simulations (119), such as the self-assembled POPC bilayer. Simulations can reveal that random aggregates of lipids organize rapidly and orient their headgroups toward water. Another example is the Langmuir monolayer (119), in which addition of small amounts of amphiphiles (surfactants) can change air-water properties due to their strong adsorbent properties. Parametrization of CG model insures correct adsorption of molecules, given a well-fitted free energy transfer and a correct surface tension, employed to determine how the CG lipid monolayer behaves under compression. The surface pressure is defined as:

\[
\pi = \gamma_0 - \gamma
\]  

(6.2)

where \( \gamma \) is the surface tension of the monolayer and \( \gamma_0 \) is the surface tension of water. For the correct temperature, CG model gives the correct surface pressure, which measures the reduction of surface tension when lipids are adsorbed at air/water interface. Another application is vesicle formation, for which simulations of tens of nanoseconds are required to obtain a single expanding aggregate. The stability of the vesicle obtained can be tested by carrying out an MD simulation of two vesicles. External forces are required for interaction between the two vesicles. Adhesion, deformation and fusion are the effects of external forces. Reshaping occurs as a result of external forces.
forces removal.

Internal structure of polymer chain (195) can be analysed by determining the coherent structure function \( S(q) \) of an individual chain,

\[
S(q) = \frac{1}{N} \left( \sum_{i=1}^{N} \exp(i \mathbf{q} \cdot \mathbf{r}_i) \right)^2
\]

The coherent structure function \( S(q) \) can be averaged over several randomly chosen \( q \) vectors. It is expected that fractal scattering, \( S(q) \sim q^{-1/\nu} = q^{-2} \), where \( \nu = 1/2 \), independent of chain length.

Another measure of internal structure of the chain is the study of Rouse modes of the chain, given by (196,197):

\[
\mathbf{X}_p(t) = \frac{1}{N} \left[ \sum_{i=1}^{N} \mathbf{r}_i(t) \cos \left( \frac{p\pi(i-1)}{N-1} \right) \right] - \frac{1}{2N} \left[ \mathbf{r}_1(t) - \mathbf{r}_N(t) \right]
\]

These are eigenmodes for ideal chains, for which:

\[
\langle \mathbf{X}_p(t) \cdot \mathbf{X}_p'(t) \rangle = 0, \quad p \neq p'
\]

The amplitudes obey scaling relation:

\[
\langle \mathbf{X}_p(t) \cdot \mathbf{X}_p'(t) \rangle \propto \frac{1}{2\pi^2} \frac{\mathbf{R}^2(N)}{p^2} \alpha \frac{N}{p^2}
\]

The Rouse modes \( \mathbf{X}_p(t) \) defined by equation (6.4) are the eigenmodes of the chain. According to Rouse model (198,199), the relaxation function of the modes is given by:

\[
g_p(t) = \frac{\langle \mathbf{X}_p(t) \cdot \mathbf{X}_p(0) \rangle}{\langle \mathbf{X}_p(0) \cdot \mathbf{X}_p(0) \rangle} = \exp(-t/\tau_p)
\]

where \( p \) is the index of the mode and
The relaxation time of $p$th mode is the longest relation time of a chain made of $N/p$ units. Relaxation times are given by the analysis of relaxation functions.
REFERENCES


(44) Jessee JA. Composition and methods for transfecting eukaryotic cells 1996.


(64) YashRoy RC. Destabilisation of lamellar dispersion of thylakoid membrane lipids by sucrose. Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism 1994;1212(1):129-133.


(93) Branson J. Phase space and Liouville’s theorem. 2012.


(110) Shell MS. Coarse graining and multiscale techniques. 2012.


