Molecular Dynamics Simulation of Phase and Structural Transitions in Model Lung Surfactant Mixtures

by

Susan L. Duncan

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Engineering)
in The University of Michigan
2010

Doctoral Committee:

Professor Ronald G. Larson, Chair
Professor Ayyalusamy Ramamoorthy
Professor Peter J. Woolf
Professor Robert M. Ziff

To my two boys Aiden and Montana and my husband Aaron. Thank you for your love and support. You truly are a blessing and I will cherish you forever.

And to my family Janet, Ronnie, Seth, Sarah, Shannon, Shirley, Clarence, Michael, Jonathan, and Bita; who will always hold a special place in my heart.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Prof. Ronald G. Larson for always being supportive, giving me the freedom to explore my research interests, and providing me excellent advise along the way. I would like to thank my committee members for their commitment and for their valuable suggestions that have challenged me to think of my research in a new light. I would like to thank the University of Michigan for funding this research. I would like to thank all members of the Larson group for their support and companionship. I further thank Indranil Saha Dalal for his contribution to the work presented in Chapter 3, and Senthil Kandasamy, Zuowei Wang, and Hwankyu Lee for useful discussions. Finally, I would like to thank my family for their love and support, without which, this journey would have been a lot tougher.

TABLE OF CONTENTS

DEDICATION ii

ACKNOWLEDGEMENTS iii

LIST OF FIGURES vii

LIST OF TABLES ix

ABSTRACT x

CHAPTER

I. Introduction 1

II. A Comparison of Simulated and Experimental Pressure-Area Isotherms of DPPC. 8

2.1 Introduction 8
2.2 Simulation Methods 11
2.2.1 Coarse Grained Simulations 13
2.2.1.1 Surface Tension Coupling 15
Independent Runs 16
Cycling 16
2.2.1.2 Anisotropic and Semi-isotropic Pressure Coupling 17
2.2.1.3 NVT 17
2.2.2 Atomistic Simulations 17
2.3 Results 18
2.3.1 Simulated Isotherms 18
2.3.2 Effect of Ensemble 23
2.3.3 P-N Orientation 26
2.3.4 Radial Distribution Functions 28
2.3.5 Hole Formation 32

2.3.6 Effect of Bead Size . . . 33
2.4 Discussion . . . 35
2.4.1 Comparing Simulated and Experimental Isotherms . 35
2.4.2 Area Compressibility Modulus . . . 46
2.4.3 Effects of System Size, Time-Scale, and Hysteresis . 50
2.5 Summary . . . 53

III. Folding Transitions in Lipid and Lipid-Peptide Monolayers. . 57

3.1 Introduction . . . 57
3.2 Simulation Methods . . . 63
3.3 Results . . . 68
3.3.1 Folding . . . 68
3.3.2 Peptide Aggregation . . . 70
3.3.3 Mutants . . . 74
3.3.4 Order Parameters . . . 78
3.3.5 Temperature and Surface Tension . . . 84
3.3.6 Bicelle Formation, Fusion, and Re-expansion . . . 86
3.4 Discussion . . . 87
3.4.1 Folding Mechanism . . . 87
3.4.2 Component Effects . . . 92
3.4.2.1 Pure DPPC . . . 93
3.4.2.2 Unsaturated Lipids . . . 94
3.4.2.3 Surfactant Proteins . . . 94
3.4.2.4 Palmitic Acid . . . 96
3.4.3 Defects and Aggregation . . . 97
3.4.4 Mutants . . . 101
3.4.5 Squeeze-out and Supercompression . . . 103
3.4.6 Reversibility and Bicelle Formation . . . 106
3.4.7 Fusion and Surface Refining . . . 107
3.4.8 Effects of System Size and Time-Scale . . . 108
3.5 Summary . . . 110

IV. Preliminary Study of LC-LE Phase Transitions in Lipid and Lipid-Peptide Monolayers. . . . 112

4.1 Introduction . . . 112
4.2 Simulation Methods . . . 117
4.3 Results and Discussion . . . 122
4.3.1 LC-LE Phase Transitions in Pure DPPC Monolayers 122
4.3.2 LC-LE Phase Transitions in DPPC Monolayers Containing Peptides . . . 124
4.3.3 LC-LE Phase Transitions in Binary Lipid Monolayers Containing DPPC . . . 129
4.3.3.1 Monolayers Containing Cholesterol . . . 131

4.3.3.2 Monolayers Containing POPC 133
4.3.4 Hysteresis Loops 136
4.3.5 Nucleation and Growth and Fluctuations in Packing 138
4.3.6 System Size Effects and Simulation Time 143
4.4 Summary 144

V. Summary and Future Work 146

APPENDIX 154

BIBLIOGRAPHY 179

LIST OF FIGURES

Figure

2.1	Typical features of pressure area isotherms	10
2.2	Pressure-area isotherms at various temperatures	19
2.3	Coarse grained pressure-area isotherm obtained by cycling at 293.15K and corresponding images of the packing	20
2.4	Pressure-area isotherms at 323.15K	22
2.5	Coarse grained pressure-area isotherms obtained using three different ensembles	24
2.6	P-N tilt angle distribution	27
2.7	Radial distribution functions	30
2.8	Hole formation	34
2.9	Simulated and experimental pressure-area isotherms at various temperatures	36
3.1	The occurrence or absence of folding in the simulated monolayers	71
3.2	Folding or the absence of folding in monolayers containing PA	72
3.3	The occurrence of folding in two monolayers separated by water, containing 256 DPPC and 4 SP-B_{1-25} molecules each	73
3.4	A different initial orientation for SP-C	75
3.5	Two lipid-peptide simulations where folding does not occur within 500ns of simulation under small negative surface tension	77
3.6	Order parameter versus time	80
3.7	The carbon-deuterium order parameters (-S_{CD}) versus the number of hydrophobic residues per peptides present	81
3.8	The carbon-deuterium order parameters (-S_{CD}) are reported for various lipid and lipid peptide mixtures	82
3.9	Decreasing the temperature affects the occurrence of folding	85
3.10	Bicelle formation	86
3.11	An example of peptide-mediated surface refining	88
3.12	Folding in monolayers containing pre-aggregated peptides	98
4.1	Ordered and disordered initial configurations	123
4.2	LC-LE phase transitions in DPPC monolayers	124
4.3	LC-LE phase transitions in DPPC monolayers containing SP-B_{1-25}	125
4.4	Packing in larger monolayers composed of DPPC with SP-B_{1-25}	126

4.5 The influence of peptide orientation on the LC-LE phase coexistence 128
4.6 The phase behavior of monolayers containing SP-C 130
4.7 LC-LE phase transitions in DPPC monolayers containing cholesterol 133
4.8 LC-LE phase transitions in 1:1 DPPC:POPC monolayers 135
4.9 Hysteresis loops 137
4.10 Nucleation and growth of the LC phase 140
4.11 Packing fluctuations 142
4.12 Longer simulation times lead to decreased hysteresis in pure DPPC monolayers 144

LIST OF TABLES

Table

2.1	Experimental conditions used to obtain pressure-area isotherms . .	39
2.2	Experimental and simulated area compressibility moduli	48
3.1	Simulations performed at 323K and small negative surface tension .	69
3.2	Peptides used in folding simulations	76

ABSTRACT

Lung surfactant (LS) is a complex mixture of lipids and proteins that reduces and regulates the surface tension in the lungs, thereby decreasing the work of breathing. A thorough understanding of LS function is critical to the development and optimization of synthetic surfactants for the treatment of neonatal and adult respiratory distress syndrome. We have utilized coarse-grained (CG) molecular dynamics simulation to study the dynamic, hysteretic changes occurring in the structure and phase of model surfactant mixtures with varying temperature, pressure and composition. In particular, we have studied the effects of the LS components palmitoyloleoylphosphatidylglycerol (POPG), palmitoyloleoylphosphatidylcholine (POPC), palmitic acid (PA), cholesterol, and two surface-active proteins SP-B_{1-25} (the 25-residue N-terminal fragment of SP-B), and SP-C on model surfactant monolayers containing the primary lipid component dipalmitoylphosphatidylcholine (DPPC). The results indicate that POPG, POPC, SP-B_{1-25} and SP-C act as fluidizers and PA and cholesterol act as condensing agents, which change the phase-transition temperature, LC-LE phase distribution, and the extent of hysteresis. To explore the role of LS proteins SP-B and SP-C in storing and redelivering lipid from lipid monolayers during the compression and re-expansion occurring in lungs during breathing, we have simulated 2D-to-3D transitions at the interface. These simulations show that at near-zero surface tension the presence of a fluidizing agent, such as POPG, SP-C, or SP-B_{1-25} decreases the monolayers resistance to bending allowing the monolayers to form large undulations and ultimately folds. Another folding mechanism is also observed in monolayers

containing peptides, involving the lipid-mediated aggregation of the peptides into a defect, from which the fold can nucleate. The occurrence of folding depends on the hydrophobic character of the peptides; if the number of hydrophobic residues is decreased significantly, monolayer folding does not occur. In contrast, the addition of PA has a charge-dependent condensing affect, which can eliminate folding. Our results suggest that the peptides play a significant role in the folding process, and provide a larger driving force for folding than does POPG. In addition to promoting fold formation, the peptides also display fusogenic behavior, which can lead to surface refining.

CHAPTER I

Introduction

As computational resources improve, so do the possibilities for computer simulation. Increased computational power has allowed computer simulation to advance, to a level of detail, beyond what was thought possible just a decade ago. In addition, progress in simulation techniques, such as the development of well-parameterized coarse-grained models has also allowed the simulation of more complex systems. By grouping atoms into "pseudo-atoms" or "beads" the number of calculations is dramatically reduced leading to a large increase in computational efficiency. The development of coarse-grained models has allowed the simulation of longer length- and time-scale events that are not accessible by atomistic simulation. As computational resources and techniques have advanced, computer simulations are being applied to more complex systems; moving from small molecules to large proteins and bio-membranes. In recent years, simulation studies of bio-molecules have exploded in number, and have been applied to a number of interesting biophysical phenomena.

Computer simulations of bio-membranes are of great interest because they can yield molecular-level insight into the structure and dynamics of these systems at small spatial and temporal scales not accessible experimentally. The vast majority of these simulations have focused on bilayers, both lamellar and non-lamellar. These bilayer simulations have been use to simulate self-assembly, domain formation, poration and

ion permeation, lipid flip-flop, bilayer deformations including membrane undulations and buckling, vesicle formation, fusion, and lipid-protein interactions, as reviewed in references [1–4].

The simulation of domain formation in particular, has received much attention in the last decade. Lipid membrane phase equilibria and the formation of domains, such as cholesterol-rich lipid rafts, are thought to play a highly active role in membrane function [5, 6]. For example, protein-binding and lipase activity have been correlated with the formation of nano-domains [6]. Just within the last five years, several coarse grained (CG) and atomistic molecular dynamics (MD) simulation have been applied to the simulation of the domain formation in phospholipid bilayers including gel phase formation [7–13], formation of the rippled gel phase [10, 14–16], domain formation in lipid mixtures consisting of short and long tail lipids [17, 18], and raft-like mixtures [19–22]; for a detailed review of these simulations along with DPD and meso-scale simulations see reference [1]. In addition to these simulations there are numerous studies devoted to the simulation of other bilayer behavior.

There has been comparatively few MD simulation studies focused on the simulation of phospholipid monolayers [23–43]. However, monolayers also provide an intriguing area of study. Lipid monolayers can also act as model systems for biological membranes, which can be considered as two weakly coupled monolayers [44]. Moreover, their rich polymorphism make them attractive model systems [6, 44]. Slight changes in composition, temperature, or subphase pH can lead to dramatic changes in phase behavior. This rich phase behavior opens the possibility of altering the structure, dynamics, and lateral organization of the membrane in response to minor changes in environmental conditions [5]. Additionally the study of phospholipid monolayers is particularly relevant to the study of lung surfactant, which forms a monolayer at the air-water interface in the lungs.

We have applied coarse grained molecular dynamics to the study of monolayers,

with a particular focus on the study of important components of lung surfactant. Lung surfactant (LS) is a mixture of phospholipids, fatty acids, neutral lipids, and surfactant proteins that forms the surface-active lining in the lungs. Lung surfactant consists of approximately 90% lipids and 10% proteins by weight [45]. Of the surfactant lipids, about 80% are phosphatidylcholines, about half of which is dipalmitoylphosphatidylcholine (DPPC) [45]. There is also a significant concentration of anionic phosphatidylglycerol (PG) molecules. Other components include fatty acids such as cholesterol and palmitic acid (PA). Of the four lung surfactant proteins SP-B and SP-C are highly hydrophobic and readily associate with surfactant phospholipids, while SP-A and SP-D are hydrophilic proteins, which reside in the aqueous subphase.

Lung surfactant plays a very important physiological role. By lining the air-water interface in the alveoli, lung surfactant reduces and regulates the surface tension at the interface, decreasing the work of breathing and stabilizing the lungs against alveolar collapse. Infants born prematurely lack functional lung surfactant and develop respiratory distress syndrome (RDS). Surfactant replacements have greatly reduced the mortality rate of RDS, but are not optimal [46]. Furthermore, acute respiratory distress syndrome (ARDS), which develops in adults, is often more difficult to treat than RDS due to underlying complications such as lung injury. ARDS has a mortality rate of 30-40% [47] and can be attributed to many causes including inhibition by inflammatory processes or plasma protein leakage. In addition to RDS and ARDS, surfactant abnormalities have been observed for a variety of diseases that are characterized by airflow obstruction, including asthma, chronic bronchitis, cystic fibrosis, bronchiolitis, pneumonia, cardiogenic lung edema, meconium aspiration, and human immunodeficiency virus (HIV) [48–50].

Surfactant replacements come in two forms; synthetic and animal derived. Animal derived surfactants are costly, limited in quantity, can exhibit batch-to-batch variability and they pose a high risk of infection and a high risk of rejection by the

patient's immune system [51]. Treatment of ARDS using expensive and scarce animal-derived surfactants is particularly problematic because a large amount of surfactant is required to treat adults and exogenous surfactants are particularly susceptible to inhibition by plasma components [51]. For this reason much effort has been placed on the development of synthetic surfactants, which can potentially be produced in larger quantities, at a lower cost, and designed to be inherently resistant to inactivation. However, current synthetic surfactants exhibit poor efficacy compared to natural surfactants. This can be attributed largely to the lack of synthetic analogues of SP-B and SP-C, which play a critical role in lung surfactant function [52]. Protein-free surfactants such as Exosurf, which includes DPPC and spreading agents, exhibit inferior efficacy compared to surfactant preparations containing proteins [51]. As a result, recent investigation has turned toward the development of a new generation of synthetic, biomimetic surfactants that include synthetic hydrophobic surfactant protein mimics [51, 52]. Another issue plaguing the development of surfactant replacements is the current lack of understanding regarding the dynamic response of lung surfactant to demanding changes in surface tension occurring during the respiratory cycle, and how individual surfactant components modulate this dynamic response. For instance, although cholesterol is thought to be an important lung surfactant component and is know to have a substantial effect on the phase behavior and lateral organization of membranes, it is systematical removed from most surfactant replacements. Similarly, PA constitutes only a very small fraction of natural surfactant, but is routinely added in large proportions to some surfactant replacements. However, the additional of palmitic acid may substantially alter the response of the surface film to high pressures, as discussed in chapter 3. Many animal derived and synthetic surfactants have a lipid composition that differs greatly from that in natural surfactant, this difference in the lipid composition is often required to accommodate for a lack of surfactant protein or for the use of synthetic peptides that are less than optimal [51]. Mak-

ing significant changes in the concentration of surfactant components with respect to physiological concentrations, without fully appreciating the physiological role of these components leads to less than optimal surfactant replacements. Some of the most significant developments in lung surfactant therapy may be yet unsuspected due to the current state of knowledge regarding underlying molecular mechanisms [51].

In order to aid the design of effective synthetic and animal-derived surfactant replacements, further research is needed to understand the mechanisms that promote physiological surfactant function, and the roles of individual components in the respiratory process. To be effective, lung surfactant must display rapid adsorption, the ability to compress to near-zero surface tension upon end-expiration, and rapid respreading upon film expansion [53]. There are a few theories regarding lung surfactants ability to perform two seemingly conflicting tasks: being stable (solid) enough to avoid collapse and being fluid enough to adsorb and respread readily. Both phase and structural (reversible collapse) transitions have been implicated in this process. Despite much effort many questions remain, see references [46, 54, 55] for review.

Experiments can provide images of surfactant films on the micrometer scale but are limited in their ability to assess molecular-level interactions, which cannot be assessed through direct visualization. In order to obtain insight into the mechanisms involved, experiments can be supplemented with computer simulations to provide molecular level detail. This purpose of this thesis work is to apply molecular dynamics simulation to the simulation of monolayers composed of DPPC with other surfactant components, in order to obtain a better understanding of the mechanism by which lung surfactant is able to successfully regulate the surface tension in the lungs and prevent irreversible monolayer collapse. Particular emphasis will be placed on assessing the influence of individual components on this process. A secondary purpose for this work is to evaluate the application of the coarse grained MARTINI model [56–58] to lipid and lipid/peptide monolayers, undergoing phase and structural

transitions. The remainder of this dissertation is outlined as follows:

Before recent advances in spectroscopy and diffraction techniques, biophysicists relied heavily on the measurement of the pressure-area isotherms to provide thermodynamic data about the monolayer. The pressure-area isotherm is still the most common form of analysis today. Therefore, in chapter 2, we compare experimental and simulated pressure-area isotherms for DPPC (dipalmitoylphosphatidylcholine) at temperatures ranging between 293.15K and 323.15K, and explore and possible factors influencing the shape and position of the isotherms. From this analysis, it is evident that there is much more variation among experimental isotherms than between isotherms obtained from CG simulations and atomistic simulations. Although the measurement of pressure area isotherms is common place, it is not always appreciated how much they can vary depending on experimental conditions. Therefore, a detailed review of the factors affecting the shape of pressure area isotherms is provided in the appendix. These factors are also briefly discussed in chapter 2.

In chapter 3, the collapse transitions of model mixtures are examined. To explore the role of lung surfactant proteins SP-B and SP-C in storing and redelivering lipid from lipid monolayers during the compression and re-expansion occurring in lungs during breathing, we simulate the folding of lipid monolayers with and without these proteins. We also test the affect of altering the hydrophobicity of SP-B_{1-25}, by using several peptide mutants. Additionally re-expansion of the monolayer and the fusogenic properties of the peptides are also assessed.

In chapter 4, LC-LE phase transitions are examined by obtaining snapshots and hysteresis loops for DPPC, mixed lipid, and lipid-peptide monolayers at various temperatures. LC-LE transitions are of particular interest since they occur as the monolayer undergoes changes in surface pressure associated with the dynamic cycling occurring in the lungs with each breath. Also, phase transitions play an important role in the functionality of lung surfactant, since the phase of the monolayer determines

its mechanical properties.

Finally, in chapter 5, we conclude and provide suggestions for future work.

CHAPTER II

A Comparison of Simulated and Experimental Pressure-Area Isotherms of DPPC.

2.1 Introduction

Lung surfactant (LS), the surface-active lining of the alveoli, consists of approximately 90% lipids and 10% proteins by weight [45]. Of the surfactant lipids, about 80% are phosphatidylcholines, about half of which is dipalmitoylphosphatidylcholine (DPPC, phosphatidylcholine with two palmitic acid tails, also known as dipalmitoyl lecithin or DPL) [45]. Not only is DPPC the primary component of lung surfactant, but it is also thought to be primarily responsible for the reduction of surface tension in the lungs to near-zero. Thus, understanding the response of DPPC to changes in surface area is fundamental to determining the functionality of lung surfactant and how to better design lung surfactant replacements for respiratory distress syndrome, both neonatal and adult. Measurements of the surface behavior of surfactant films under dynamic compression have been among the most prevalent methods of study of pulmonary surfactant [59]. These measurements are typically reported in the form of a pressure-area isotherms.

The defining features of a typical pressure-area isotherm for DPPC, in the proximity of the main phase transition temperature, are shown in Figure 2.1(a). The surface pressure π is calculated as:

$$\pi = \gamma_0 - \gamma \tag{2.1}$$

where γ_0 is the surface tension of pure water and γ is the surface tension of the monolayer-coated air-water interface [60]. The monolayer area is typically given in terms of area/lipid. With increasing area and decreasing surface pressure, the phase transitions of the DPPC monolayer proceed in the following order: LC (liquid-condensed), LC-LE (coexistence between the liquid-condensed and liquid-expanded phases), LE (liquid-expanded), and LE-G (coexistence between the liquid-expanded and gaseous phases). The LC-LE phase transition is a first-order transition and is thus ideally represented by a perfectly horizontal plateau; however experimental coexistence plateaus are only roughly horizontal. Once the monolayer has been compressed into a condensed phase, it becomes relatively incompressible and very low surface tensions (high surface pressures) are achieved with little change in area; thus the LC portion of the isotherm has a steep slope. When the monolayer is compressed past its limiting area, monolayer collapse occurs. Collapse is signified by a decrease in area at constant surface pressure (a collapse plateau), resulting from the loss of lipids from the monolayer. In general, as the temperature is increased, DPPC isotherms shift to higher surface areas or equivalently higher surface pressures at a fixed area, and the coexistence region becomes less horizontal and is shifted to higher surface pressures [61]. As shown in Figure 2.1(b), this behavior is seen in the isotherms of Crane et al. [62], which were obtained at 298.15K, 303.15K, and 310.15K using the captive bubble apparatus. This behavior is attributed to an increase in the thermal motion of the chains at higher temperature, which leads to an increase in surface pressure [63]. Phillips and Chapman [64] found the static DPPC pressure-area isotherms obtained at various temperatures differed in the coexistence region, but converged at high (near-zero surface tension) and low (near-zero surface pressure) surface pressures. Similar observations can be seen in the isotherms obtained at various temperatures

by Crane et al. [62] using the captive bubble apparatus (Figure 2.1(b)), and in the film balance experiments of Baldyga and Dluhy [65].

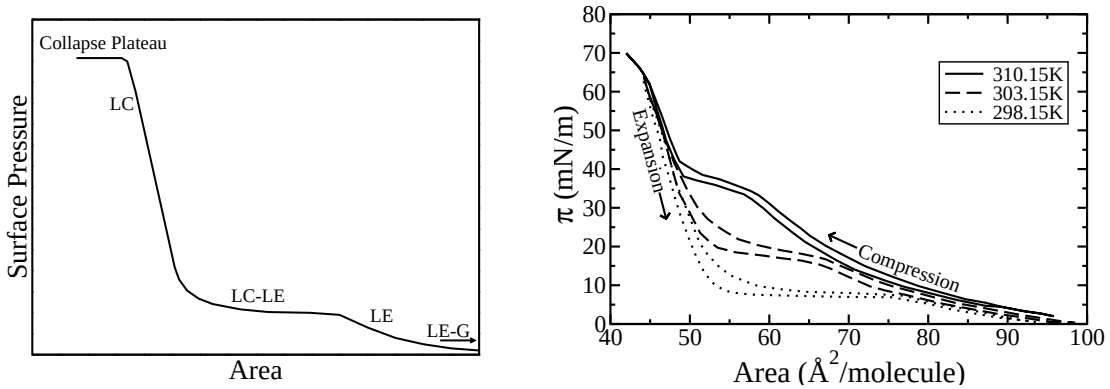


Figure 2.1: Typical features of pressure area isotherms. (a) The defining features of a typical pressure-area isotherm for DPPC near the main transition temperature. The phase regions include the LC (liquid-condensed), LE (liquid-expanded), and the LC-LE and LE-G transition regions. The LC-LE horizontal coexistence region and the horizontal collapse plateau are identified. (b) Experimental results showing the effect of temperature on the shape of compression and expansion pressure-area isotherms of DPPC. These isotherms are reproduced from those published by Crane et al. [62], at 298.15K (dotted line), 303.15K (dashed line), and 310.15K (solid line). The experimental results presented in this figure (right) and in subsequent figures were obtained using Data Thief III V.1 [66].

Computer simulations of phospholipid systems are of great interest because they can yield molecular-level insight into the structure and dynamics of these systems on a resolution and time-scale that may not be feasible experimentally. Coarse grained simulations have the further advantage of realizing increased simulation times and larger system sizes. Like their experimental counterparts, pressure-area isotherms obtained from simulations of lipid monolayers also vary from study to study. For comparison, simulations of DPPC monolayers using both coarse grained (CG) and atomistic models are included here, both from the work of other authors and from our own new simulations. To the best of our knowledge there has not yet been a comprehensive review of the factors that could effect the shape of the pressure-

area isotherm, nor a critical comparison of experimental and simulated pressure-area isotherms obtained from varying methods and experimental conditions. Therefore, here, in addition to presenting our new simulation work, we review a broad and diverse sample of the huge number of published isotherms for DPPC monolayers.

The remainder of this chapter is outlined as follows: First, we provide details of our simulations, then present the simulation results, and finally compare them with experimental results with a brief discussion of factors that might contribute to the observed large variation among experimental results. Although our discussion will focus on DPPC, many of the factors discussed here affect the isotherms of other phospholipids similarly.

2.2 Simulation Methods

Our simulations are divided into five categories: coarse grained (CG) pressure-area isotherm simulations using 1) surface tension coupling, 2) anisotropic pressure coupling, 3) semi-isotropic pressure coupling, and 4) the NVT ensemble, as well as 5) atomistic pressure-area isotherm simulations using surface-tension coupling. Simulation parameters are given for each type of simulation below. For all simulations, temperature was maintained by coupling to a Berendsen thermostat with a 1ps time constant [67]. All simulations were run with periodic boundary conditions. All simulations and analysis were performed using GROMACS simulations software [68, 69]. The GROMACS analysis tool g_energy was used to extract the surface tensions and box dimensions at each time step [70]. To obtain surface pressure from our surface tensions, pure water surface tensions of 72.8, 72.5, 72.0, 71.2, 69.6, and 67.9mN/m were used at temperatures of 293.15K, 295.15K, 298.15K, 303.15K, 313.15K, and 323.15K, which are roughly the surface-tension values given in the CRC Handbook of Chemistry and Physics [71]. It should be noted that the simulated surface tensions at the air-water interface actually differ considerably from the experimental

values, due to the peculiar nature of water [38, 72]. Vega and Miguel [73] calculated a surface tension of 54.7mN/m from their SPC water simulations at 300K, which underestimates the experimental value by 17mN/m. This could conceivably lead to an over-estimation of surface pressures in our isotherms, which are calculated from the experimental surface tension. If this were the case, the low-surface-pressure expansion observed in our simulations at surface pressures near 30mN/m would actually be occurring at significantly lower surface pressures. However, errors in simulated water/vapor surface tension are thought to have little effect on the measurement of monolayer surface tension, which is dominated by headgroup/water and chain/vapor interactions [33]. Thus, it is unlikely that our surface pressures are over-estimated significantly. Because sources of error in simulation of water surface tension are likely to be particular to water and not expected to similarly affect the simulation of monolayer surface tensions, we believe that it is more accurate to use the experimental values of water surface tension instead of the simulated ones, in our calculation of monolayer surface pressure.

Experimental results are typically performed under atmospheric pressure, corresponding to a normal pressure of 1bar. An applied normal pressure of 1bar is commonly used in bilayer studies [43, 74–76]. However, the simulation of monolayers requires the use of empty space placed above the monolayer to prevent the monolayer from interacting with the periodic image of the simulation box. Despite the presence of the lipid/vacuum interface, implying a normal pressure of 0bar, some monolayer studies have used an applied normal pressure of 1bar [39, 43, 77]. We have simulated several points along the CG isotherm at 298.15K using both a normal pressure of 0bar and of 1bar. Allowing the height of the box to fluctuate with an applied normal pressure of 1bar leads to shrinkage in the z-dimension, upon lateral expansion, requiring the box size to be manually adjusted by periodic addition of more vacuum space. However, the use of 1bar vs. 0bar led to no detectable difference in the isotherm.

Therefore all results presented here will be for simulations performed at 1bar. It has also been noted that due to large fluctuations in instantaneous pressure on the order of hundreds of atmospheres, in a simulation, 1bar is essentially equivalent to 0bar [43, 78].

2.2.1 Coarse Grained Simulations

For all of our coarse grained simulations, we utilize the peptide force field parameters developed by Marrink et al. [58]. The area/headgroup for DPPC bilayers using the coarse grained model of Marrink et al. was found to match the experimental value, and many other properties have been found to match experiment at a quantitative or semiquantitative level [58]. The CG model for DPPC has one bead representing the phosphate moiety, one bead representing the choline moiety, two tail beads representing the glycerol linkage, and four beads for each of the tails (each tail bead corresponds to four tail carbons). This model is used in conjunction with the coarse grained model of Marrink et al. for water, which merges four water molecules into a single coarse grain bead. The structure files for the CG DPPC monolayers were adapted from the CG structure files given on Marrinks website for DPPC bilayers in the fluid phase [79] and energy minimized. The resulting fluid phase monolayer files contained two monolayers (composed of 256 lipid each) placed so that their headgroups were initially separated by 7nm of CG water molecules (10,654 CG molecules) and their tail groups were separated by 10nm of empty space. The resulting disordered monolayers were contained in a box of size 12.6847nm x 12.8295nm x 23.2nm. However, in some of our CG simulations, spontaneous box shrinking became an issue, and intermittent addition of vacuum was necessary to prevent the two monolayers from merging into a single bilayer. For all simulations the following parameters were taken from Marrinks website [79] and have been optimized for the coarse grained model: short-range electrostatic and van der Waals cutoffs of 1.2nm, with van der Waals interaction shifting

smoothly to Lennard Jones interaction at 0.9nm, and with the Lennard Jones cutoff set to 1.2nm. The neighbor list was updated every 10 steps using a grid with a 1.2nm cutoff distance. In all coarse grained simulations, the energy parameters were saved every 0.4ps and used for analysis with the GROMACS analysis tool g_energy [70].

Most of our coarse grained simulations were 20ns in duration. Marrink and Mark [80] suggested that only a few nanoseconds of simulation time are needed to measure area/lipid for CG simulation. However, our results have shown that 10ns of equilibration time was necessary before areas settled down to steady values. Thus only the last 10ns of our 20ns simulations were used for the calculation of average surface tension and area. The radial distribution functions and angle distributions were also averaged over the last 10ns of the 20ns CG simulations. In some cases, near a phase transition, from mostly LE to mostly LC phase and vice versa, simulations appeared to be meta-stable, and longer simulation times up to 100ns were necessary. In each case, the last 10ns of simulation time were used for calculations. At large values of surface tension, the box size diverged and eventually exploded, making movement further down the isotherm to low surface pressures impossible. The divergence of box size is attributed to the onset of hole formation, followed by expansion and ultimately the rupture of the monolayer. A plot of lateral area versus simulation time is given in the supplementary material for a CG simulation displaying uncontrollable box expansion.

Because we are using the original CG model of Marrink et al., all liquid-condensed phases simulated will be untilted [7]. Marrink and co-workers have shown that tilted phases can be simulated using the CG model, if the model is altered to increase the size difference between the head and tail group beads. By decreasing the tail group bead size by 10%, Marrink et al. [7] succeeded in simulating the tilted phase in a DPPC bilayer. It should also be noted that due to the use of smoother potential functions for CG simulations the dynamics of CG simulations are significantly faster (of course

in computer time, but also even in physical time, as reported by the simulation) than for atomistic simulations. As a result the effective time, which has been determined from water and lipid lateral diffusion rates, is roughly four times longer than the physical time [58]. All times reported in this paper will be physical time as reported by the simulation not the effective times.

Three different pressure-coupling methods were employed: anisotropic, semi-isotropic and surface tension pressure coupling. Anisotropic pressure coupling allows the box to flex independently in six directions (xx, yy, zz, xy/yx, xz/zx and yz/zy) in response to a change in the pressure tensor. Semi-isotropic pressure coupling only allows the box to change dimension laterally (x/y) and vertically (z). Surface tension coupling is similar to semi-isotropic pressure coupling, however it uses normal pressure coupling for the z-direction, whereas the surface tension is coupled to the x/y dimensions of the box. The average surface tension γ(t) is calculated from the difference between the normal and the lateral pressure and the box is allowed to change dimension laterally (x/y) to adjust the surface tension back toward the set value. For more details on each coupling mechanism the reader is referred to the Gromacs User Manual [70] and relevant simulation papers [41, 43, 74, 76, 78].

2.2.1.1 Surface Tension Coupling

Simulations with surface tension coupling were run at 293.15K, 295.15K, 298.15K, 303.15K, and 323.15K. These simulations were run at several surface tensions varying between -50 and 62.5mN/m. For all simulations the z pressure component was set to 1 bar. Berendsen pressure coupling was used with a 1ps time constant and with all compressibilities set to 5E-6 bar^{-1}. A timestep of 0.04ps was used for most simulations. However, simulations undergoing a large change in box size (near the phase transition plateaus) required a smaller time step of 0.02ps and longer simulation times. Two types of initial configurations were used:

Independent Runs The simulations hereafter referred to as independent runs involved the independent quenching of each simulation from a state that was initially disordered. These simulations were run with the fluid phase monolayer files described above as the initial configurations. All independent runs lasted 20ns, except at 295.15K where runs were 100ns in length, because 20ns simulations had not fully converged. In addition, independent runs were also performed at 298.15K from an initial configuration containing 1024 lipids. This configuration was obtained from the disordered configuration containing 256 lipids/monolayer (described above) by patching four boxes together and performing energy minimization.

Cycling For each temperature, the fluid phase monolayer was used as a starting configuration for a 200ns simulation at a surface tension of -50mN/m. The large negative value of surface tension is physiologically meaningless, but was chosen to ensure that the resulting configuration was well ordered. This resulting configuration was then used as the starting configuration for a 20ns simulation at zero surface tension, and then the final configuration of this run was used as the starting configuration for the next run at higher surface tension. This process of using the previous run as the starting point for the next run was repeated, stepping down the isotherm to the largest surface tension attainable. When the surface tension reached the largest value possible without a diverging box size, the process was reversed, stepping back up the isotherm until a zero surface tension was reached. This process of cycling enables the simulation of a complete hysteresis loop. At 303.15K, the cycling simulations were also performed with a simulation time of 100ns for each run, to test the extent of equilibration of the 20ns cycling simulations.

2.2.1.2 Anisotropic and Semi-isotropic Pressure Coupling

Anisotropic and semi-isotropic pressure simulations were run at 298.15K and at lateral pressures of 0, -10, -20, -30, and -40bar. For these simulations the z pressure component was set to 1bar and the off-diagonal pressure components of the anisotropic pressure tensor were all set to 0bar. Berendsen pressure coupling was used with a 1ps time constant and with all compressibilities set to 5E-6 bar^{-1}. For all simulations, a timestep of 0.04ps was used. These simulations were run independently starting from the disorder configuration, containing 256 lipids/monolayer, described above.

2.2.1.3 NVT

Two NVT simulations were run at 298.15K. Both simulations were started form the disordered monolayer configuration, containing 256 lipids/monolayer, described above. The first simulation was run with the initial box size unchanged. The other simulation was run with the box size widened to 14nm x 14nm x 23.2nm and then energy minimized. For both simulations, a timestep of 0.04ps was used.

2.2.2 Atomistic Simulations

Atomistic simulations were performed using the GROMACS force field [68, 69]. An atomistic structure file containing a 128 lipid DPPC bilayer was taken from the Tieleman group website [81] and modified to create a system containing two DPPC monolayers composed of 64 lipids each. The monolayers were placed with their head-groups facing each other and initially separated by 7nm of SPC water molecules (9662 molecules) and their tail groups separated across a periodic boundary by 10nm of empty space. The resulting system was then energy minimized and used as the starting configuration for each simulation. A 2fs time step was used and each simulation was run for 10ns. The bond lengths were constrained using the LINCS algorithm [82]. A particle mesh Ewald summation [83] was used to calculate the electrostatic

interactions with a Fourier spacing of 0.12nm and a 4th order interpolation. The Coulomb cutoff was set to 0.9nm and the van der Waals cutoff was set to 1.2nm. The neighbor list was updated every 10 steps using a grid with a 0.9nm cutoff distance. Temperature was maintained at 323.15K with a Berendsen thermostat [67]. Surface-tension coupling was used with a Berendsen barostat and a time constant of 1.0ps with all compressibilities set to 4.5e-5 bar^{-1}. The z pressure component was set to 1bar. The simulations were run at several surface tensions varying between 0 and 60mN/m. Energies were output every 0.4ps for the calculation of pressure-area isotherms. Calculations were made over only the last 5ns of each simulation using the GROMACS analysis tool g_energy [70]. The radial distribution functions and angle distributions were also averaged over the last 5ns of the 10ns atomistic simulations.

2.3 Results

2.3.1 Simulated Isotherms

We performed 20ns cycling coarse grained simulations of DPPC monolayers, using surface tension coupling, as described in the section on simulation method, at 293.15K, 295.15K, 298.15K, 303.15K, and 323.15K. The resulting compression and expansion isotherms, for each temperature, are shown in Figure 2.2. An increase in temperature results in an upward shift to larger surface pressures, a shortening of the LC-LE coexistence region of both the compression and expansion isotherms, and an increasing slope in the coexistence region of the compression isotherms. With the exception of the isotherm at 323.15K, which is shifted slightly to the right, all of the isotherms overlap except in the coexistence region. Although some experimental isotherms exhibit large hysteresis loops, the hysteresis seen in our isotherms is much larger than usually seen experimentally (Figure 2.1(b)), our LC-LE coexistence regions occur at much larger pressures, and our isotherms are also shifted to larger

areas/lipid than those seen experimentally. Despite these differences, there are also some similarities. Experimental isotherms show, as seen in the simulations, that as the temperature is increased the coexistence region becomes less horizontal and is shifted to higher surface pressures, although the limiting high-pressure area of the isotherm remains invariant with temperature (Figure 2.1(b)). At 323.15K hysteresis can be seen between compression and expansion isotherms at near zero surface tension, suggesting meta-stability of the LC phase in the expansion isotherm at high surface pressure (Figure 2.2).

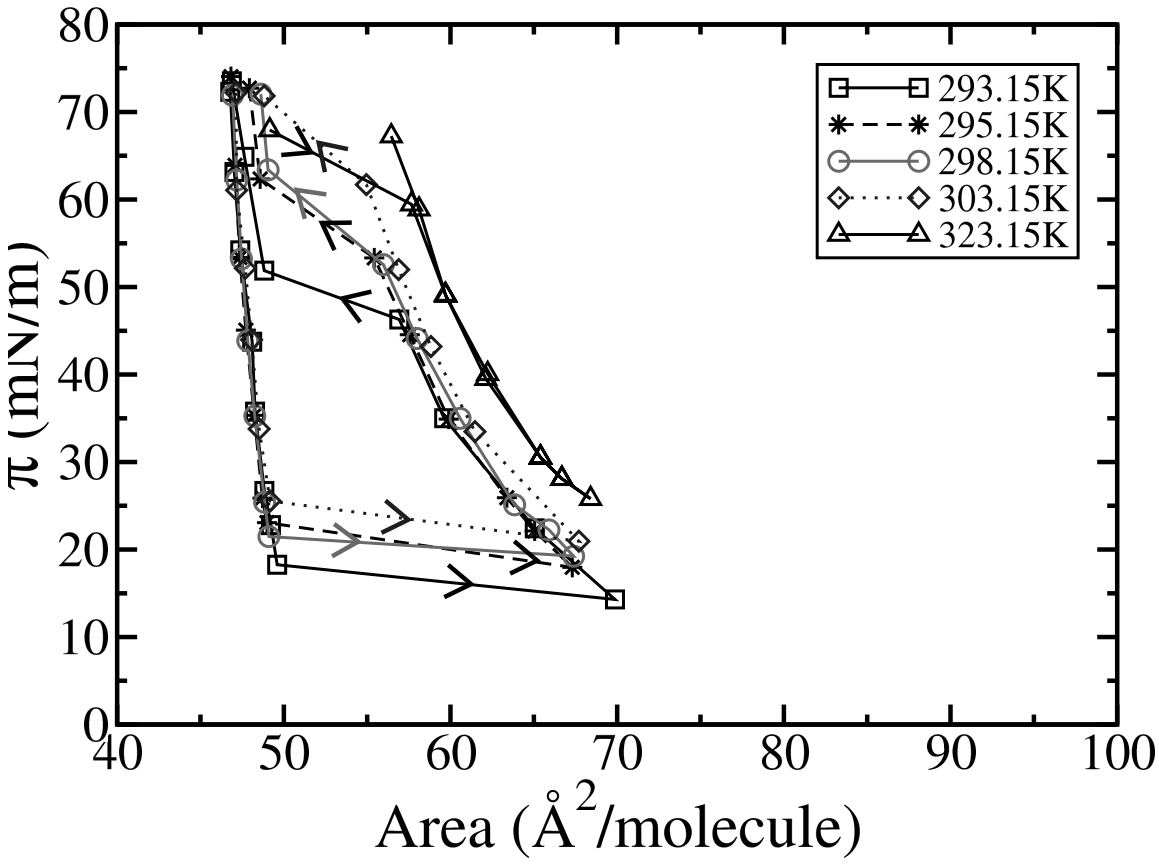


Figure 2.2: Pressure-area isotherms at various temperatures. Our pressure-area isotherms, obtained using cycling of coarse grained simulations at 293.15K (squares), 295.15K (asterisks), 298.15K (circles), 303.15K (diamonds), and 323.15K (triangles). The arrows indicate the direction of cycling. In this and subsequent figures, the error bars (standard error) on our simulated isotherms are roughly the same size as the symbols.

Figure 2.3 shows the coarse grained cycling isotherm at 293.15K and the corresponding changes in the packing of the C2 tail beads with movement along the isotherm. Hexagonal packing, which is characteristic of the LC phase, is clearly visible at low areas/lipid. Whereas at larger areas/lipid the tail beads display disordered packing typical of the LE phase. As expected, the phase transition region, or plateau region, is accompanied by a visible change in the degree of order of the chain packing.

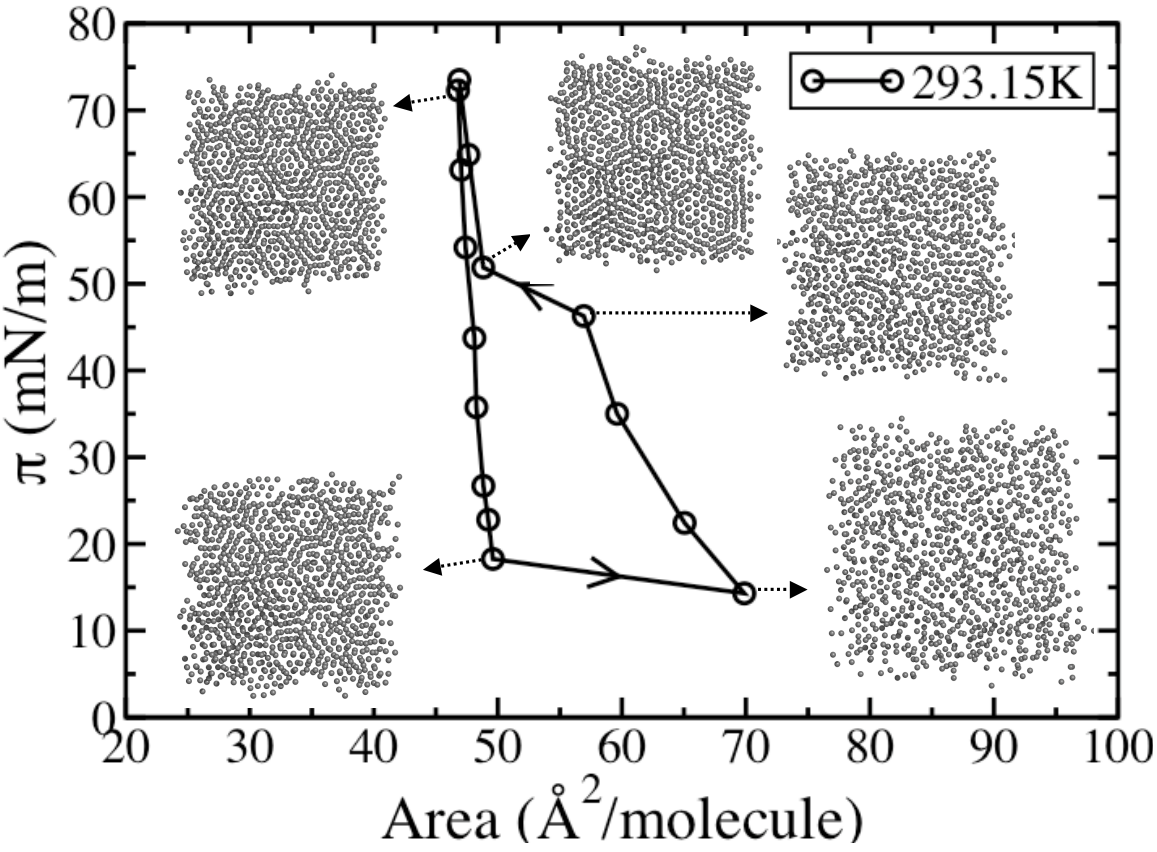


Figure 2.3: Coarse grained pressure-area isotherm obtained by cycling at 293.15K and corresponding images of the packing. In the packing images the C2 tail beads (from both monolayers) are shown at various points along the isotherm.

We therefore compare our coarse grained simulations to atomistic simulations, both our own and those obtained by others, as well as to the coarse grained results of Adhangale et al. [77] all at 323.15K. In Figure 2.4, our coarse grained results,

both from independent quenching and cycling, are compared to our atomistic results from independent quenching, as well as to the atomistic results of Kaznessis et al. [38], Skibinsky et al. [36], and Klauda et al. [33] and to the coarse grained results of Adhangale et al. [77], and to the experimental results of Crane et al. [62]. Kaznessis et al., Skibinsky et al. and Klauda et al. obtained their atomistic pressure-area isotherms using the NVT ensemble in CHARMM. Adhangale et al. used the coarse grained model developed by Marrink et al. [58], with the $NP_N\gamma T$ ensemble in the simulation package NAMD. The experimental pressure-area isotherm of Crane et al. [62] was obtained using a captive bubble apparatus. Our coarse grained results are very close to those obtained from our atomistic simulations. This indicates that the shift of the pressure-area isotherms to larger areas/lipid (relative to most experimental isotherms) is not an artifact of the coarse grained model, but occurs for coarse grained and atomistic simulations alike. Our simulations also resemble the atomistic results of Skibinsky et al. and Klauda et al. and the experimental results of Crane et al., differing slightly in magnitude and slope, whereas the results of Adhangale et al. are shifted to considerably lower area/lipid, and the results of Kaznessis et al. are shifted to considerably lower surface pressures.

Skibinsky et al. obtained starting configurations for their NVT monolayer simulations at each area, from $NP_N\gamma T$ bilayer simulations. This provided a well-equilibrated starting point for the monolayer simulations, which is necessary to obtain an accurate surface pressure in constant volume simulation, which does not allow area to adjust to bring the system to equilibrium. The simulations of Klauda et al. were started from the final coordinates obtained by Skibinsky et al., and run under the same conditions as used by Skibinsky et al. but with the addition of the Isotropic Periodic Sum method to treat long-range Lennard-Jones interactions. This isotherm agrees very well with the Skibinsky isotherm, only shifted slightly, suggesting that the treatment of long-range LJ interactions has only a small effect on the isotherm. On the

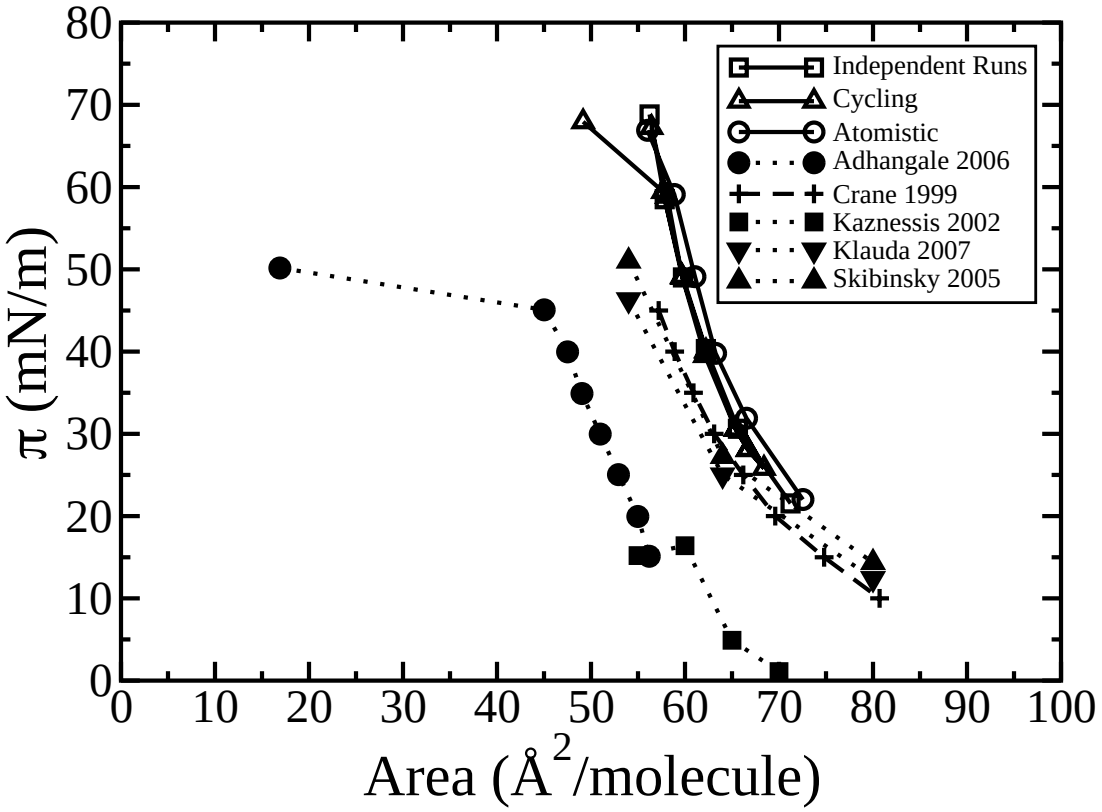


Figure 2.4: Pressure-area isotherms at 323.15K. Comparison of simulated and experimental pressure-area isotherms at 323.15K: our independent coarse grained simulations (open squares), our cycling coarse grained simulations (open triangles up), our atomistic simulations (open circles), the atomistic simulations of Kaznessis et al. 2002 [37] (filled squares), Klauda et al. 2007 [33] (filled triangles down), Skibinsky et al. 2005 [36] (filled triangles up), the coarse grained simulations of Adhangale et al. 2006 [77] (filled circles), and the experimental results obtained by Crane et al. 1999 [62] using the captive bubble apparatus (+ symbols). For simplicity our simulations are denoted by open symbols and solid lines, experiments are denoted by characters and dashed lines, and filled symbols and dotted lines denote simulations by other groups.

other hand, our results were obtained using the $NP_N\gamma T$ ensemble with two different starting conditions: independent quenching from a disordered state and cycling (stepping down and back up the isotherm point by point from an initially ordered state). The results of Adhangale were obtained using the same coarse grained model used in our simulations (the CG model of Marrink et al.), but with long-range electrostatics added in the form of a smooth particle mesh Ewald summation. The large difference between the results of the simulations Adhangale et al. [77] and our simulations may result from a problem with their periodic boundary conditions, which leads the monolayer to curve substantially at the edges, seemingly suggesting buckling, while maintaining disorder in the acyl chains even at increased surface pressure, where our simulations and experiments show highly ordered tails. The low surface pressures shown by the isotherm of Kaznessis et al. may result from the short simulation time of 1.3ns, which is not adequate for pressure convergence. Simulation of a DPPC monolayer has also been performed by Mauk et al. [42], using a united atom model and the CHARMM22.0 force field at 21°C; however in this very early paper, only two points of the isotherm were simulated, and the time scale simulated was only 120ps, too short to provide reliable results.

2.3.2 Effect of Ensemble

For comparison, we ran two NVT simulations at 323.15K (Figure 2.5 diamonds). The first simulation was run without making adjustments to the box size (63.6Å^2/molecule), and the second simulation with the box size increased (76.6Å^2/molecule). When the box size is increased, an unphysical increase in pressure is observed, suggesting that the NVT ensemble does not allow for sufficient pressure relaxation. Other authors have noted the inability of constant-area and constant-volume simulations to equilibrate to appropriate pressures. Simulations of DPPC bilayers performed by Feller et al. [43][41] also show that constant-area simulations tend to predict larger surface

pressures at a given surface area than those predicted by constant-surface-tension simulations. Mauk et al. [42] found that the NπT ensemble was more favorable than the NAT ensemble, the latter of which yielded inaccurate equilibrium pressures and chain order. Furthermore, Mauk et al. [42] have suggested that the inaccuracy of NAT simulations of phospholipids monolayers is due to the lack of fluctuations in the periodic cell, which restricts the phospholipids from assuming energetically favorable conformations.

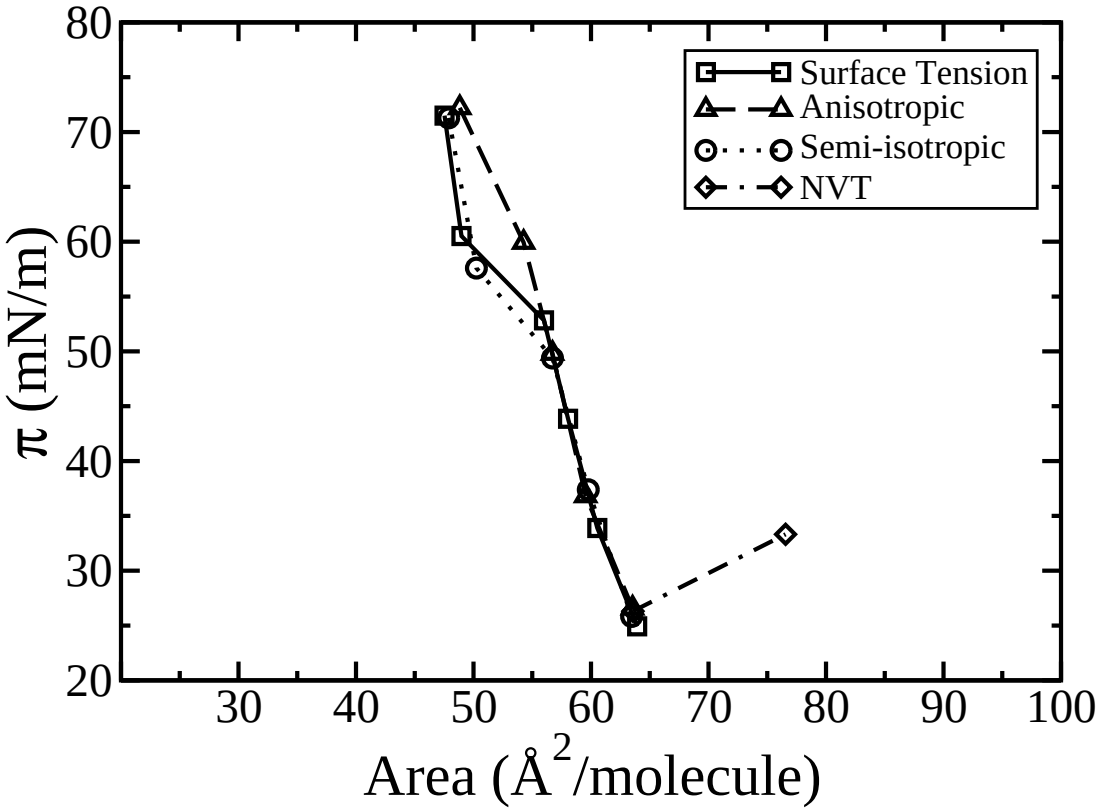


Figure 2.5: Coarse grained pressure-area isotherms obtained using three different ensembles. Coarse grained pressure-area isotherms obtained at 298.15K using the NVT ensemble (diamonds) and the NPT ensemble with three pressure coupling mechanisms: surface tension (squares), anisotropic (triangles) and semi-isotropic (circles).

Enforcing a constant surface area imposes a stronger restriction on the phase space available to the system then does enforcing a constant average pressure [84]. Area

is an extensive property that does not fluctuate when constrained. On the other hand, pressure is an intensive property, which is constrained as a time-averaged constant with fluctuations allowed. Also a change in pressure can be provoked by small intermolecular displacements, whereas a change in area requires large concerted motions of the lipids. Thus, the system is slow to equilibrate in response to imposed changes in area [84]. However, it should be noted that constant-area simulations give reasonable results if the starting conditions are well equilibrated. In their simulations of DPPC bilayers, Feller and Pastor [41] found that order parameters, lateral diffusivities, magnitudes of area fluctuations, area fluctuation decay rates, and bilayer area compressibility moduli did not depend significantly on choice of ensemble (NP_NAT versus $NP_N\gamma T$). In more recent studies, DPPC bilayer simulations showed that the pressure-area isotherms obtained using both ensembles were consistent with each other, suggesting the equivalence of the ensembles [36, 85].

In addition to surface-tension coupling and NVT simulations, we also performed coarse grained simulations using anisotropic and semi-isotropic pressure coupling methods, to test the accuracy of each method. The isotherms obtained with each coupling method at 298.15K are shown in Figure 2.5. At 298.15K, each coupling method gives nearly the same isotherm, differing only in LC-LE coexistence region, where they give different slopes. Although the choice of coupling method does not seem to have a big impact, the surface-tension coupling method yields the flattest plateau. Furthermore, in their simulations Feller et al. [43] set surface tension and allowed area to vary, regarding this as the most natural ensemble for simulating lipid/water interfaces. For these reasons surface-tension coupling was chosen as the preferred method and used for the majority of our simulations. Feller and Pastor [41] have suggested that simulation results depend much more on area than on ensemble used, which is consistent with our findings at 298.15K.

2.3.3 P-N Orientation

We calculated the distribution of P-N tilt with respect to the membrane normal from our atomistic simulations at 323.15K (Figure 2.6). For comparison, the P-N tilt in our CG simulations is taken as the tilt of the bond connecting the PO4 and NC3 CG beads with respect to the membrane normal, which is calculated at 298.15K from simulations on the larger system size (1024 lipids/monolayer), and at 323.15K for the smaller size of 256 lipids/monolayer. The tilt angle was compared at areas/lipid corresponding to the two endpoints of each isotherm. No change is observed in the coarse grained P-N tilt angle distribution as the area is changed, at either 298.15K or 323.15K. However, the atomistic simulations show a noticeable difference in the P-N tilt distribution as the monolayer is expanded from 56 to 73Å^2/molecule. At 73Å^2/molecule the distribution is narrower than for the distribution at 56Å^2/molecule and shifted so that although the probability of an angle below 60° is unchanged, the probability of an angle between 60° and 105° is increased, and the probability of an angle between 105° and 160° is decreased. The coarse grained distributions are similar to the atomistic distribution at 56Å^2/molecule. However, the CG PO4-NC3 tilt distribution does not exhibit the dependence on surface area seen in the atomistic simulations. The coarse grained distributions show a shift to lower angles as the temperature is increased, and the distribution narrows slightly, excluding angles above 160°. Our atomistic simulations at 323.15K give a single peak centered at 90° at 56Å^2/molecule and at 85° at 73Å^2/molecule. Our coarse grained simulations peak at 90° at 298.15K and 78° at 323.15K.

Numerous experimental studies, including surface-potential measurements, on phospholipid bilayer systems suggest that the P-N orientation is parallel to the bilayer surface [86, 87]. A recent sum frequency generation (SFG) spectroscopy study performed by Ma and Allen [88] suggests that the choline methyl groups are tilted from the surface normal and lie roughly parallel to the air-water interface. The SFG

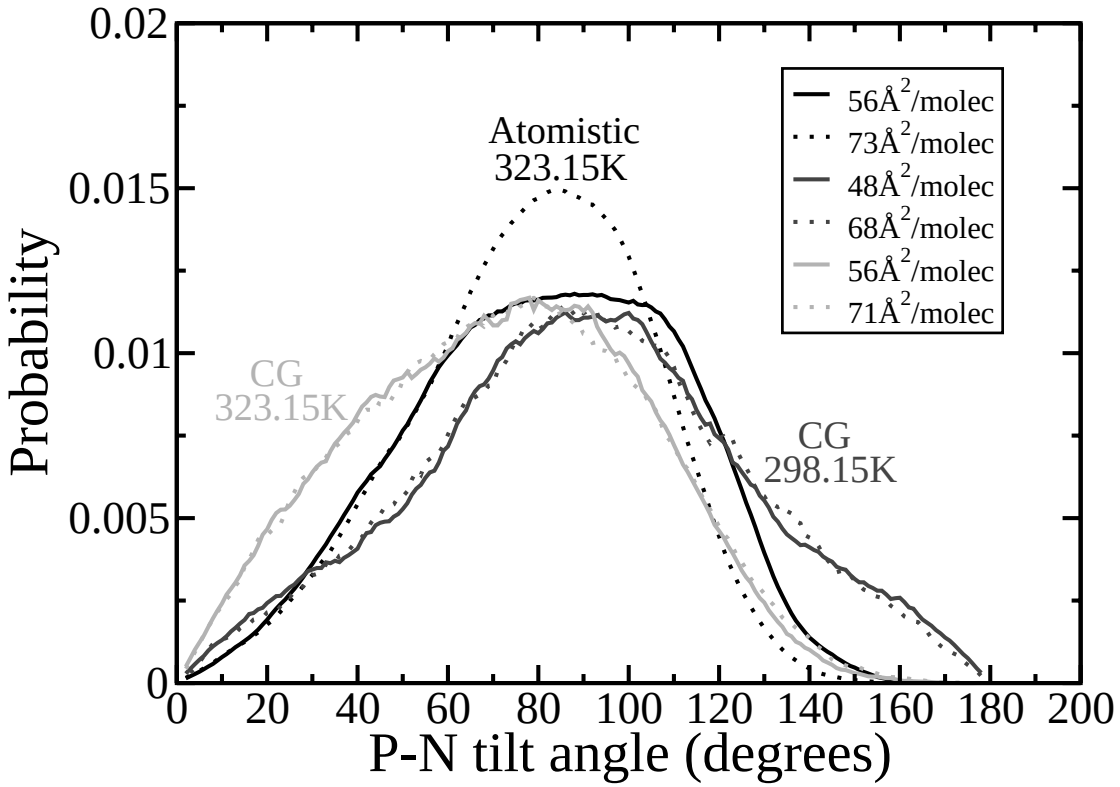


Figure 2.6: P-N tilt angle distribution. The P-N tilt angle distribution for atomistic simulations at 323.15K with areas 56Å^2/molecule and 73Å^2/molecule, for coarse grained (CG) simulations with 1028 lipids/monolayer at 298.15K with areas 48Å^2/molecule and 68Å^2/molecule, and for coarse grained simulations with 256 lipids/monolayer at 323.15K with areas 56Å^2/molecule and 71Å^2/molecule. The black, dark grey, and light grey lines represent the atomistic simulations, and CG simulations at 298.15K and 323.15K, respectively. For each shade, the solid and dotted lines represent the smaller and larger area per lipid, respectively. For clarity, the data shown here has been smoothed using time-averaged values.

spectra obtained by Ma and Allen [88] at 12mN/m (LE phase) and 42mN/m (LC phase) are similar. These results suggest that the choline headgroup orientation is not significantly different in the LE and LC phases, in accordance with the previously held hypothesis that the overall conformation of the headgroup is not as sensitive to the aggregation state and the nature of the environment as the tails [88, 89]. The P-N tilt angle distributions obtained from our atomistic and coarse grained simulations are also centered at or near 90°, in accord with experiments. Our results are also in agreement with previous atomistic simulations of a DPPC monolayer performed by Dominguez et al. [40], which showed that the average angle between the monolayer surface and the P-N vector was 5°. Although the shape of simulated P-N distributions vary, more recent atomistic [39] and coarse grained [77] simulations have also shown average P-N tilt angles in the proximity of 90° with respect to the membrane normal.

As the DPPC monolayer undergoes a transition from the liquid-expanded to the liquid-condensed phase, the methylene groups of the DPPC tails transform from predominantly gauche conformations to all-trans conformations [88]. The lipid tail dihedral distribution was calculated from the four CG tail beads for a system size of 1024lipids/monolayer. At 298.15K we found that at 48Å^2/molecule the trans tail configuration (180°) is highly preferred over the gauche configuration (60°) and at 68Å^2/molecule the trans configuration becomes less favorable and the distribution broadens such that all tail dihedrals are sampled almost equally, as is expected (data not shown).

2.3.4 Radial Distribution Functions

In Figure 2.7, the PO4-PO4, PO4-NC3, NC3-NC3, and C2-C2 radial distribution functions (RDFs) are shown, where PO4 is the phosphate moiety, NC3 is the choline moiety, and C2 is the second CG tail bead from the glycerol linkage (which corresponds to the 5th through 8th carbon atoms from the glycerol linkage); each of

these sites is represented by a single coarse-grained bead. Each radial distribution function is normalized so that the integral is equal to the total number of lipids (twice the number of lipids in the case of the C2-C2 distribution because there are two C2 sites/lipid). The atomistic results compared in Figure 2.7 were obtained using the following atoms: P, N, and the 6th tail carbon from the glycerol linkage. The two endpoints of each isotherm are selected to observe the effect of surface area on the shape of the radial distributions. Each isotherm used was obtained from independent runs rather then cycling. At 298.15K (Figure 2.7 left) the RDFs are compared at areas of 48 and 68Å^2/molecule for the larger CG system size (1024 lipids/monolayer). At 323.15K (Figure 2.7, center and right) the RDFs are compared at areas of 56 and 71Å^2/molecule for a CG system of size 256 lipids/monolayer and at areas of 56 and 73Å^2/molecule for an atomistic system size of 64 lipids/monolayer. The difference in the areas shown at 298.15K and 323.15K reflects the shift in the isotherms to larger areas/lipid as temperature is increased.

At 298.15K, the CG PO4-PO4 (Figure 2.7(a)), PO4-NC3 (not shown), and NC3-NC3 (not shown) RDFs show little difference as area is changed from 48 to 68Å^2/molecule; however, the C2-C2 (Figure 2.7(d)) RDF changes significantly. At 48Å^2/molecule the C2-C2 RDF reflects the highly ordered tails expected for a system in the LC phase, whereas at 68Å^2/molecule it reflects the disordering of the system. These CG results are in contrast to the atomistic results of Knecht et al. [34] at 293K, which show that decreasing the area/lipid causes lipids to bind closer together leading to an increase in the phosphate-phosphate correlation in addition to the increase in tail order observed here. Although our CG radial distribution functions show a clear increase in tail order as area is decreased, unlike the atomistic simulations of Knecht et al., we see only a small increase in the height of the first phosphate-phosphate correlation peak. These results suggest that the coarse grained model is better at capturing the effect of changing surface area on lipid tails than on lipid headgroups.

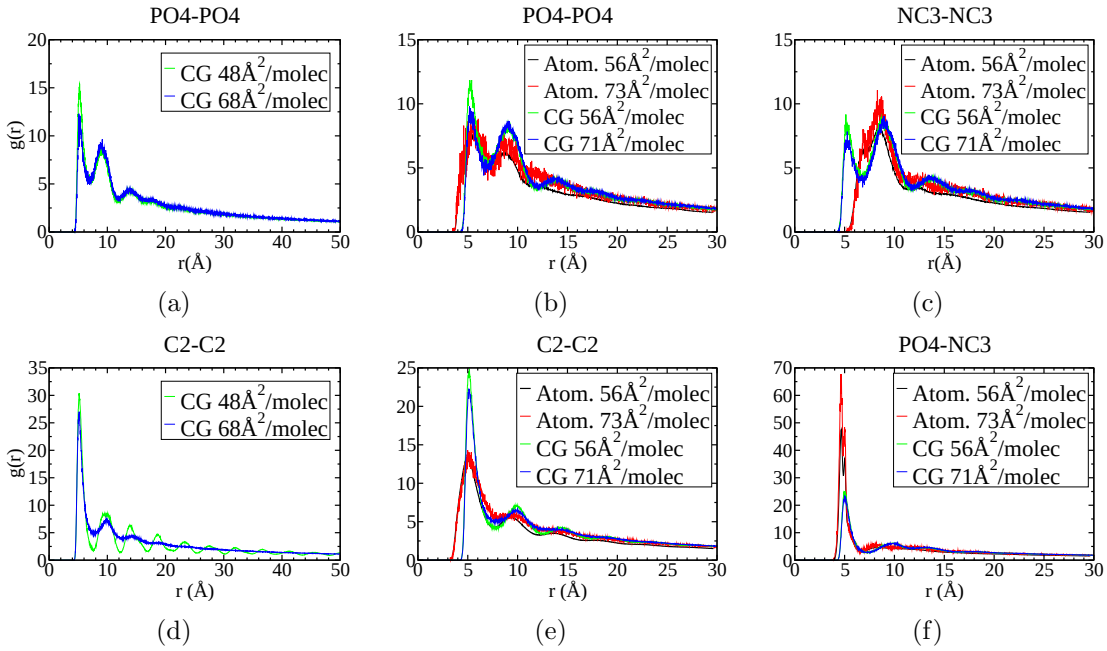


Figure 2.7: Radial distribution functions. Left: Independent coarse grained (CG) simulations at 298.15K for the larger system size (1024lipids/monolayer) at both 48Å^2/molecule (black) and 68Å^2/molecule (red). (a) PO4-PO4 distribution. (d) C2-C2 distribution. Center and Right: Atomistic (atom.) simulations at 323.15K with 64 lipids/monolayer at both 56Å^2/molecule (black) and 73Å^2/molecule (red) and independent CG simulations at 323.15K with 256 lipids/monolayer at both 56Å^2/molecule (green) and 71Å^2/molecule (blue). (b) PO4-PO4 distribution (c) NC3-NC3 distribution (e) C2-C2 distribution (f) PO4-NC3 distribution.

At 323.15K the simulated isotherms are in the expanded phase. The CG C2-C2 (Figure 2.7(e)) distribution indicates that the tails are slightly more ordered at 56Å^2/molecule than at 71Å^2/molecule. However, both areas/lipid give an RDF that reflects considerably less order than does the LC RDF at 298.15K and 48Å^2/molecule (Figure 2.7(d)), and is comparable to the less ordered distribution at 298.15K and 68Å^2/molecule (Figure 2.7(d)). At 323.15K, the CG PO4-PO4 (Figure 2.7(b)), PO4-NC3 (Figure 2.7(f)), and NC3-NC3 (Figure 2.7(c)) RDFs show little difference between the two areas/lipid and are almost identical to those at 298.15K (Figure 2.7(a), PO4-NC3 and NC3-NC3 distributions are not shown), suggesting that temperature has a larger effect on the RDF of lipid tails than that of lipid headgroups.

For the atomistic simulations at 323.15K (Figure 2.7, center and right), a change in surface area from 56 to 73Å^2/molecule does not strongly affect any of the RDFs; however the distributions do appear to fluctuate more at 73Å^2/molecule. Overall the CG and atomistic radial distribution functions match reasonably well at 323.15K. Despite some differences, the C2-C2 and C-C (Figure 2.7(e)), PO4-NC3 and P-N (Figure 2.7(f)), and PO4-PO4 and P-P (Figure 2.7(b)) RDFs correlate well. However, the NC3-NC3 and N-N (Figure 2.7(c)) RDFs differ from each other considerably, whereas the NC3-NC3 (Figure 2.7(c)) RDF is very similar to the PO4-PO4 RDF (Figure 2.7(b)), indicating that the coarse grained model is unable to capture the difference in the N-N and P-P interactions present in the atomistic simulations, which ultimately leads to inaccuracy in the NC3-NC3 RDF. The (inaccurate) similarity between the NC3-NC3 and PO4-PO4 distributions in the CG simulations is a direct result of an over-simplification contained in the CG model. The CG model uses bead types Qd (charged hydrogen-bond donor) and Qa (charged hydrogen-bond acceptor) to represent NC3 and PO4 sites, respectively. Qa-Qa and Qd-Qd Lennard-Jones interactions are both considered intermediate and have the same LJ parameters [73].

The shape and location of the peaks of our atomistic P-N and P-P RDFs correlate well with the atomistic results of Kaznessis et al. [38] for a DPPC monolayer and Sun [39] for a 1,2-dilignoceroylphosphatidylcholine (DLGPC) monolayer. Both our P04-NC3 (CG) and P-N (atomistic) RDFs show a strong attraction between choline and phosphate groups, in agreement with the atomistic results of Kaznessis et al. [38]. It has been proposed that electrostatic interactions between neighboring choline and phosphate groups are responsible for attraction between neighboring phospholipids [90].

2.3.5 Hole Formation

Our simulations show hole formation at areas in the proximity of 100Å^2/lipid, which could represent the onset of the liquid-gas phase transition. For the CG surface tension coupling simulations, at 323.15K, calculations were made for specified surface tensions between 0mN/m and 46.6mN/m, which yielded average surface pressures between 68.8mN/m and 21.5mN/m. When the specified surface tension was increased further to 47mN/m, a jump in area/lipid was observed from 71.4Å^2 to 129Å^2. As shown in Figure 2.8, this jump in area/lipid is accompanied by hole formation, which is not an artifact of the coarse grained method of simulation, because hole formation was also observed in our atomistic simulations (left). The holes are unstable and expanding, ultimately leading to the rupture of the monolayer. Knecht et al. [34] also saw hole formation in their united atom simulations of DPPC monolayers. They observed the transient formation of holes at 98Å^2/molecule and stable pore formation at 105Å^2/molecule. According to Knecht et al. the appearance of holes suggests the onset of the LE-G phase transition. Fluorescence microscopy has revealed that in the LE-G coexistence region the gas phase is present as holes in an interconnected liquid phase [91]. Due to limited spatial resolution of fluorescence images the LE-G coexistence region can not be directly determined using fluorescence microscopy [34]. However, the LE-G phase transition is thought to occur at areas of hundreds of Å^2/molecule [44]. Knecht et al. propose that the hole formation in their MD simulations corresponds to the sharp transition in the order of lipid chains recently detected by vibrational sum frequency generation spectra at 110Å^2/molecule, which they suggest could be associated with the onset of the gas-liquid coexistence region [34]. Knecht et al. also observed LC domain formation away from pore boundaries [34]. Whether LC domain formation can be seen in CG simulations at conditions beyond those needed to generate holes has not yet been tested. In contrast to our results and those of Knecht et al. [34], the results of Nielsen et al. [92] using a CG

model (which is structurally similar to the model of Marrink et al., but includes long-range electrostatics) showed that at large area/lipid, monolayer lipids become highly disordered and spread on the surface instead of forming holes. In the simulations of Nielsen et al. [92], the entropic benefit of spreading on the surface outweighs the van der Waals interaction energy, which suggests a possible problem with their energy parameterization, which they admit is exploratory and not yet validated. Hole formation has also been observed in atomistic simulations of DPPC bilayers. Leontiadou et al. [74] observed a critical surface tension (38mN/m) above which pores in the bilayer expand becoming unstable and ultimately leading to the rupture of the bilayer. Feller and Pastor [41] have also described large and sudden expansions at a surface tension of 50mN/m, which may suggest the disruption of the bilayer.

2.3.6 Effect of Bead Size

It is generally agreed that the packing of DPPC molecules is determined by the size difference between the head and tail groups, with the area required by the headgroup being substantially larger than that required for the tails, leading to packing adjustments such as lipid chain tilting and headgroup overlap [86, 89]. The coarse grained model of Marrink et al. utilizes a Lennard Jones bead size of $\sigma = 0.47$nm, for all bead types. Thus it does not capture the large difference in limiting area between the phosphatidylcholine headgroup and the acyl chains. To test the effect of the relative size difference between the headgroup and acyl chains on the packing of DPPC monolayers, we ran simulations (results not shown) with the bead size of the tails including the glycerols decreased, while the headgroup bead size remained at 0.47nm. Our simulations showed that decreasing the tail bead size by the proper amount allows the monolayer to achieve smaller minimum areas closer to the experimentally determined limiting area, while maintaining the correct packing arrangement. On the other hand, decreasing tail bead size too much impairs packing and the area is not

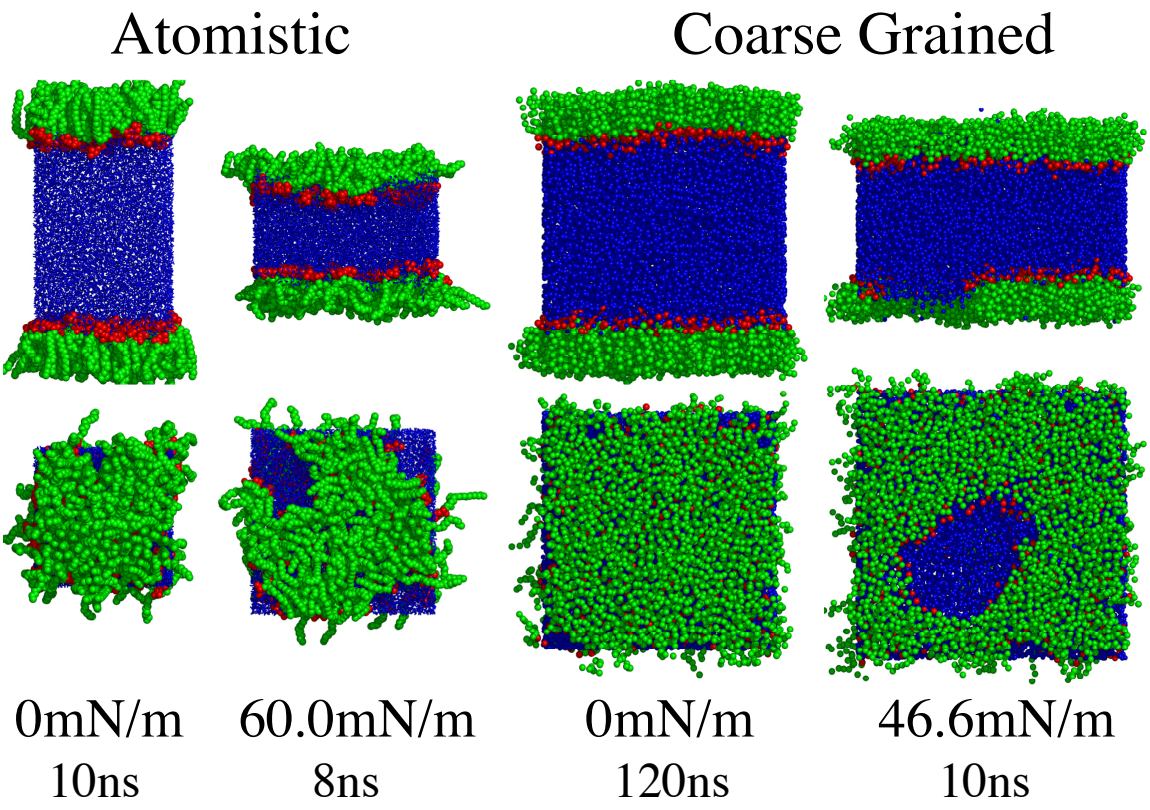


Figure 2.8: Hole formation. Hole formation in atomistic (left) and coarse grained (right) simulations at 323.15K, from the side (top) and corresponding top view (bottom). The lipid tails and glycerol groups are shown in green, the headgroups in red, and the waters in blue. The corresponding surface tensions and simulation times are given below the images.

minimized.

2.4 Discussion

2.4.1 Comparing Simulated and Experimental Isotherms

Many studies containing experimentally measured pressure-area isotherms for pure DPPC monolayers have been reported. However, very few studies compare their isotherms with those of others, and those that do tend to compare with only one or two selected isotherms that resemble their own. A major reason for this lack of comparison is due to the diverse conditions under which isotherms are obtained, making reproducibility problematic. Thus, even though the pressure-area isotherm of a monolayer is a thermodynamic relationship that, like pressure-volume isotherms for bulk substances, ought to be a universal function if measured accurately and under equilibrium conditions, in practice isotherms vary considerably, due to variability in compression rate, type and geometry of experimental apparatus, experimental artifacts (leakage, impurities, etc.), as well as pH, ionic strength, and spreading solvent [93]. The variation among selected experimental isotherms is illustrated in Figure 2.9, at 293.15K (top left), 295.15K (top right), 298.15K (bottom left), and 303.15K (bottom right) with our simulated isotherms included.

The complexity of phospholipid phase behavior and the many experimental factors involved can lead to results that are ambiguous and apparently conflicting. The difficulty in finding isotherms obtained under similar conditions has been noted before [64, 94]. Experimental artifacts can also lead to results that can be easily misinterpreted. Different authors may come to remarkably different, and often contradictory, interpretations of monolayer behavior, involving factors such as collapse mechanism, relaxation times, and the effect of the experimental conditions (spreading method, compression rate, etc.). These differences are not inconsequential; the shape of the

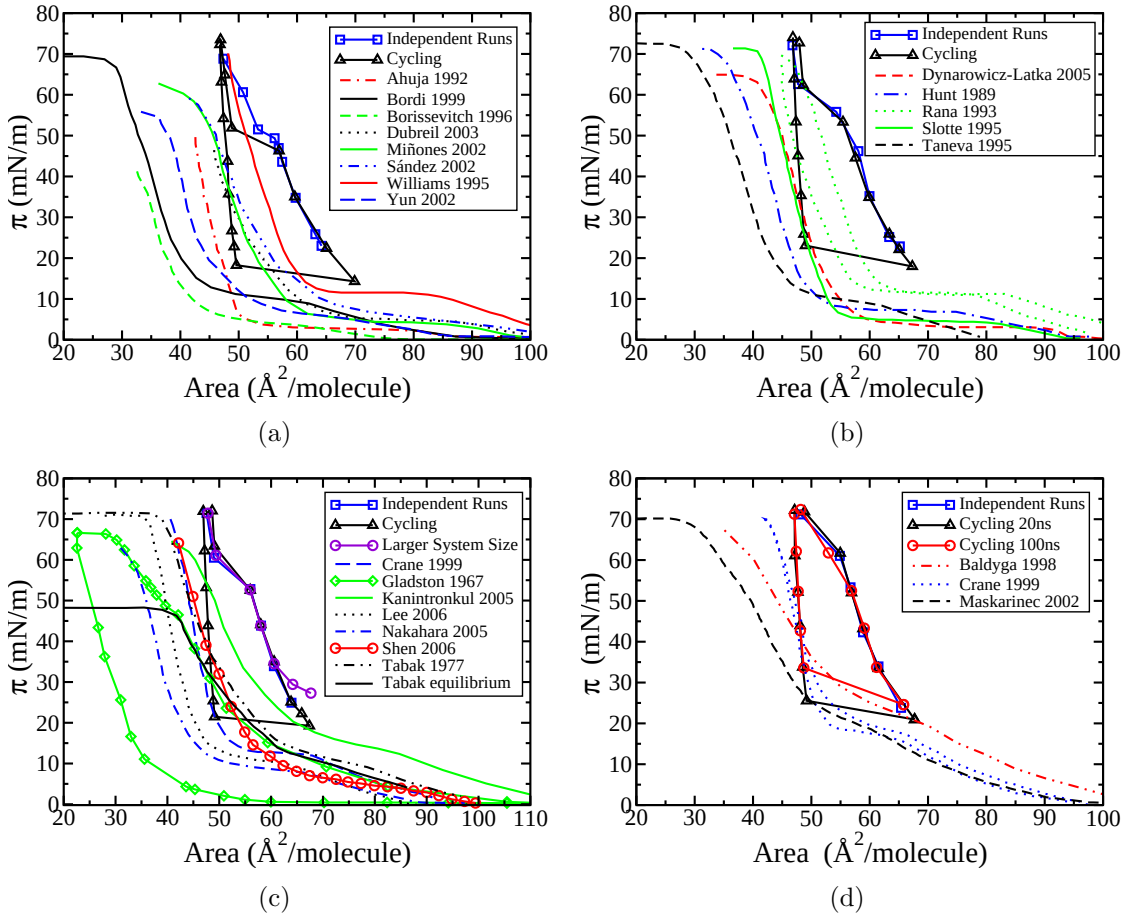


Figure 2.9: Simulated and experimental pressure-area isotherms at various temperatures. A comparison of simulated CG pressure-area isotherms with various experimental ones at (a) 293.15K, (b) 295.15K, (c) 298.15K, and (d) 303.15K.

isotherm is physiologically relevant, making accurate determination of it very important. For example, the very low surface tension when the film is compressed toward collapse is thought to be a mechanism for preventing alveolar closure at end-expiration [95], and the steep slope of DPPC postcollapse expansion isotherms is thought to be important for alveolar recruitment and stabilization of lung units during inspiration [96]. Furthermore, the shape of the isotherm is crucial to obtaining a proper understanding the behavior of the monolayer on the molecular level; for example the compressibility is determined from the slope of the isotherm [97].

When comparing experimental pressure-area isotherms, there are a few key ex-

perimental trends to keep in mind. Varying the dynamic compression rate is not expected to have a large effect [97–99] and in many cases the presence of relatively small concentrations of ions leads to little or no change in the isotherm of zwitterionic monolayers [100–104]. At moderate pH the isotherm shows little sensitivity to pH, but at low pH decreased hydrogen bonding leads to an increase in the maximum surface pressure and can cause a shift to smaller areas due to hindered solvation, and at high pH solvation is increased and equilibrium is shifted toward the fluid phase [104–106]. The type of experimental apparatus used is known to have an effect on the shape of pressure-area isotherms, and each type has a unique set of conditions and limitations to take into account [50, 59, 60, 93, 98, 107–117]. The geometry should be considered because of curvature effects, area available for creep and leakage, and disordering of lipids near walls that all effect the measurement of area/lipid. The potential for leakage is greatest at high temperatures and large dynamic pressures, and is greatest in a conventional Langmuir trough; however the use of devices such ribbon barriers help minimize or even eliminate leakage [98, 107, 111, 118]. The pulsating bubble surfactometer also suffers from leakage, whereas the captive bubble apparatus is free from the effects of leakage. Leakage leads to a shift in the isotherm to lower surface pressures and a decrease in its slope, which can be mistaken as premature collapse [107]. Even in the absence of leakage, creep along the walls can be an issue and problems with contact angle can give erroneously low surface tensions [112, 118]. Impurities may also arise from many sources including the experimental apparatus itself, and lead to isotherms that do not have a well defined phase transition region, are shifted, or do not reach near zero surface tensions upon end compression [64, 119, 120]. Careful consideration of the choice of spreading solvent is necessary, because it can have a large effect on the displacement of isotherms along the area/molecule axis and can impair film stability [94, 115, 121]. Polar components are surface active and may solubilize the lipids, causing a shift in the isotherm to very low areas/lipid due to the

loss of lipid from the interface. The effects of compression rate, pH, ionic strength, experimental apparatus, spreading agent and impurities are discussed in more detail in the appendix.

As noted by others [42], simulations of phospholipid monolayers are limited to the nanosecond timescale, which cant account for long-time adjustments that the monolayers undergo to reach equilibrium. Thus, the results of computer simulations of phospholipid monolayers must not be interpreted as equilibrium behavior, but rather as dynamic (i.e. metastable or quasi-equilibrium). This is important to take into account when comparing simulation results to experimental data. It is important to compare simulation results with dynamic isotherms (isotherms compressed relatively rapidly and thus allowed to reach near-zero surface tensions), rather than static isotherms, which have relaxed to equilibrium and reach substantially lower surface pressures.

In Figure 2.9, our simulated pressure-area isotherms are compared to experimental isotherms at 293.15K [63, 122–128] (top left), 295.15K [100, 129–132] (top right), 298.15K [62, 98, 133–137] (bottom left), and 303.15K [62, 65, 138] (bottom right), For each temperature our simulations were run both independently from an initially disordered state (black triangles) and cycled beginning from an initially ordered state (blue squares). At 303.15K, the results from cycling simulations are compared for run durations of 20ns and 100ns at each point (Figure 2.9(d) bottom right). The experimental conditions for each isotherm are given in Table 2.1, including compression rate, type of experimental apparatus, subphase composition, pH, and spreading solvent.

These experimental isotherms in Figure 2.9 vary greatly from one to the next in shape and magnitude. All of the isotherms presented here were obtained at moderate pH, except those of Kanintronkul et al. (pH 9) and Nakahara et al. (pH 2), both at 298.15K. pH is not expected to be a major factor affecting the isotherms that were

Table 2.1: Experimental conditions used to obtain pressure-area isotherms. Rate of compression, type of apparatus, subphase composition/pH, and spreading solvent used to obtain the isotherms reproduced in Figure 2.9. Abbreviations: EA= experimental apparatus, LT= Langmuir trough, MWB= modified Wilhelmy balance, LW= Langmuir-Wilhelmy balance, LWRB= Langmuir-Wilhelmy balance with a ribbon barrier, LB= Langmuir-Blodgett balance, LBW= Langmuir-Blodgett with a Wilhelmy plate, FRT = Fromherz-type round trough, TMT= Teflon-milled trough, JLFB= Joyce-Loebl film balance, CB= captive bubble method, Spread= equilibrium spreading in a beaker.

	Temp (°C)	Rate	EA	Subphase	Spreading Solvent
Ahuja[122]	20	discontinuous	FRT	pure water	Chloroform with 2% Ethanol
Bordi[123]	20	0.1cm/min	LBW	water+0.145M NaCl, pH 7.2	Chloroform/ Methanol (1:1)
Borissevitch[124]	20	2mN/m*min	LW	pure water, pH 5.9	Chloroform
Dubreil[125]	20	3cm/min	LW	phosphate buffer, pH 7.2	Chloroform
Miñones[126]	20	8.2Å^2/ molec*min	LBW	water, pH 6 (added HCl)	Chloroform/ Ethanol (4:1)
Sández[127]	20	27cm^2/min	LT	citrate, phosphate, sodium borate buffer, pH 7	Chloroform/ Ethanol (4:1)
Williams[128]	20	0.5cm^2/min	LBW	water+0.15M NaCl, pH 5.6	Chloroform/ Methanol(4:1)
Yun[63]	20	7.5cm^2/min	LBW	pure water	Chloroform
Dynrowicz-Latka[129]	22	30cm^2/min	LT	pure water	Chloroform/ Methanol (9:1)
Hunt[100]	22	5.1cm^2/min	LW	pure water	N-Hexane/ Ethanol (9:1)
Rana[130]	22	0.5cm^2/min	LBW	water+0.15M NaCl, pH 5.6	Chloroform/ Methanol (4:1)
Slotte[131]	22	<6Å^2/ molec*min	TMT	pure water	Hexane/ 2-Propanol (3:2)
Taneva[132]	22	40cm^2/min	LWRB	water+0.15M NaCl, pH 7	1-Propanol/0.5M Sodium Acetate(1:1)
Crane[62]	25,30,50	2.5-5Å^2 /molec*min	CB	10mM HEPES, 1.5mM $CaCl_2$, 0.15M NaCl, pH 7	Chloroform/ Methanol (1:2)
Gladston[133]	25	discontinuous	MWB	water+0.9% NaCl, pH 5.6	Chloroform/Methanol /Water (80:35:5)
Kanintronkul[134]	25	1cm/min	LW	carbonate buffer, pH 9	Chloroform
Lee[135]	25	4.6Å^2/ molec*min	LBW	pure water	Chloroform/ Methanol (9:1)
Nakahara[136]	25	10.3Å^2/ molec*min	LW	water+0.15M NaCl, pH 2	N-Hexane/ Ethanol (9:1)
Shen[137]	25	1.5cm/min	LB	pure water, pH 6.5	Chloroform
Tabak[98]	25	≤96Å^2/ molec*min	LWRB	pure water	Hexane/ Ethanol (9:1)
Tabak(eq.) [98]	25	n/a	Spread	pure water	Hexane/ Ethanol (9:1)
Baldyga[65]	30	not specified	JLFB	water+0.15M NaCl, pH 5.6	Chloroform
Maskarinec[138]	30	not specified	LW	pure water	Chloroform

obtained at moderate pH values, for which the monolayer is thought to be insensitive to pH. The isotherm of Kanintronkul et al. is shifted to a larger area/lipid relative to the other isotherms; it also displays elevated surface pressures at large areas/lipid, and does not display well defined phase transitions. This can likely be attributed to increased solvation and a shift in equilibrium toward the fluid phase, resulting from the high pH. In contrast, the isotherm of Nakahara et al. is shifted to lower areas/lipid reflecting hindered solvation attributed to the acidic medium.

No defining trends associated with the type of apparatus used are evident from the isotherms shown in Figure 2.9. All of the isotherms obtained at 293.15K and 295.15K were obtained in a trough (see Table 2.1), yet much variation among them remains. At 298.15K and 303.15K all pressure-area isotherms were obtained with a trough, except for the isotherms reported by Crane et al., which utilized the captive bubble apparatus. Despite this, the isotherms presented by Crane do not have any defining features that distinguish them from the other isotherms presented here. Leakage could be an issue in any of the experiments except those of Crane (because of the use of the captive bubble apparatus), the equilibrium isotherm of Tabak (because spreading inside a beaker was used), and the dynamic isotherms of Tabak et al and Taneva et al. (because of the use of a ribbon barrier). Furthermore, experiments performed without the use of a Wilhelmy plate or with discontinuous compression may be especially susceptible to leakage. Thus, leakage is a likely factor attributing to the large variation between the experimental isotherms shown here.

Dynamic compression rate appears to play a role in the slope of the isotherms at high surface pressures (low areas/lipid). The slope tends to become steeper as compression rate is increased. Isotherms compressed the quickest, such as those obtained by Bordi et al., Williams et al., Rana et al., Slotte and Mattjus, and Crane et al. have the steepest slopes. This is made more evident by the magnitude of the area compressibility moduli calculated for these isotherms (discussed in detail in the

next section). Note that although slower compression leads to better equilibration, it does not necessarily produce more accurate isotherms. Additionally, isotherms compressed quickly better mimic physiological conditions.

The spreading solvent is typically not thought to have a large effect when used in a trough, which is open to air circulation and takes up a relatively large surface area. Nevertheless, in comparing these isotherms, spreading solvent does appear to have played a major role. At 293.15K, the isotherms obtained by Borissevitch et al., Dubreil et al., Ahuja and Möbius, and Yun et al. all reach relatively low surface pressures at end compressions of roughly 42, 48, 50, and 55mN/m, respectively. Collapse does not appear to have been reached before measurement was halted for the isotherms of Dubreil et al. and Ahuja and Möbius, and it remains uncertain what the actual collapse pressure would have been. For all of these isotherms the spreading solvent was pure or almost pure chloroform (98% in the case of Ahuja and Möbius). At 295.15K, the isotherm obtained by Dynarowicz-Latka et al. used the highest concentration of chloroform in the spreading solvent (90% by volume), and also has the lowest maximum surface pressure (highest minimum surface tension). At 298.15K, slightly low dynamic maximum surface pressures are obtained by Nakahara et al. (64mN/m), Kanintronkul et al. (65mN/m), Shen et al. (65mN/m), and Gladston and Shah (67mN/m). Of these, the isotherms obtained by Kanintronkul et al. and Shen et al. used pure chloroform as a spreading agent, and that of Gladston and Shah utilized 80% chloroform. Gladston and Shah claim that the onset of film collapse actually occurred at 44mN/m as indicated by an inflection point far before the plateau at 67mN/m [134]. This change in slope could be attributed to the squeeze-out of chloroform causing some of the DPPC molecules to be removed from the monolayer, leading to an underestimation of collapse pressure. At 303.15K, the isotherm obtained by Baldgya and Dluhy displays the lowest maximum pressure and uses pure chloroform as a spreading agent.

At 293.15K the isotherms of Bordi et al., Borissevitch et al., and Yun et al. are all shifted to lower areas/lipid relative to the other isotherms shown, with the isotherm of Bordi et al. and Borissevitch et al. reaching areas/lipid even smaller than the limiting area of 39Å^2/molecule [89]. The isotherms of Borissevitch et al., and Yun et al. were obtained with a spreading solvent that was pure chloroform, and that of Bordi was obtained with a 1:1 chloroform/methanol solution. At 298.15K, the isotherms obtained by Gladston and Shah, and of Lee et al. are shifted to low areas/lipids, with the former reaching areas/lipid that are smaller than the limiting area. These isotherms were obtained using chloroform-methanol spreading solutions containing 66.7% and 90% chloroform by volume, respectively. At 303.15K, the isotherms obtained by Baldgya and Dluhy and Maskarinec et al. are shifted to areas/lipid that are smaller than the limiting areas. Both of these isotherms were obtained using pure chloroform as a spreading agent. The isotherm of Maskarinec et al. displays a collapse plateau at a very low area/lipid (30Å^2/molecule), suggesting that DPPC has been lost from the monolayer before the collapse plateau is reached. The shift in these isotherms could result from the use of pure or almost pure chloroform as the spreading solvent. It should be noted that chloroform is known to be surface active due to its polarity [121].

Thus, spreading solvent effects may contribute significantly to the observed variation between isotherms [94]. The use of chloroform as a spreading solvent appears to shift the isotherms to lower areas/lipid and also decrease the surface pressure at collapse. This suggests the possible loss of lipid from the monolayer, perhaps through the removal of some DPPC molecules from the monolayer with the squeeze-out of chloroform, leading to a decrease of area and the appearance of premature collapse. Whatever the mechanism, the use of chloroform clearly impedes the ability of the monolayer to reach near-zero surface tensions. Others have noted that by increasing the amount of chloroform in the subphase, the amount of DPPC lost from the

film increases dramatically, indicating that the presence of chloroform can impair film stability [116]. According to Wüstneck et al. [121], the interaction between DPPC molecules is strongly depressed by the presence of chloroform, which is in incorporated into the monolayer and causes an apparent increase in the molecular area. At low pressures this causes a shift of the isotherm to larger areas and an increase in minimum surface pressure [121]. However, as surface pressure is increased the chloroform is squeezed-out taking DPPC molecules along and shifting the isotherm to low areas/lipid [121]. Wüstneck et al. [121] found that when chloroform is present at the surface in large enough concentrations, a plateau corresponding to the squeeze-out of chloroform can be visualized at 50mN/m. The increased molecular area at low pressures (due to the presence of chloroform in the monolayer) combined with the decreased molecular areas at high pressure (due to the loss of DPPC with chloroform from the monolayer) results in an isotherm that is broader, changing gradually with surface pressure. Other experimental spreading solvents may also cause a shift in the isotherm and the ratio of polar and nonpolar components in the spreading solvent is of critical importance [94]. At 295.15K, the isotherm of Taneva et al. is shifted to small areas/lipid, and reaches an area smaller than the limiting area. Taneva et al. note that the use of the propanol/sodium acetate solvent appears to have contributed to a shift in the isotherm of pure DPPC to low areas/lipid, perhaps due to partial dissolution of DPPC into the subphase or incomplete dissociation of aggregates formed in the spreading solution [132].

Although spreading solvent appears to play a role, with so many experimental factors involved, it is impossible to be sure what leads to a shift in area/lipid or shape of a given isotherm. At 295.15K, the isotherm obtained by Dynarowicz-Latka et al. was obtained without the use of a Wilhelmy plate, and is thus at a higher risk of film leakage, which could also explain the low surface tension at collapse. At 298.15K, the compression isotherm of Gladston and Shah displays an inflection point at 44mN/m,

which may be a result from the choice of spreading agent, but could also be attributed to leakage, which may be magnified by the use of discontinuous compression or by the experimental apparatus. Also at 298.15K, the isotherm obtained by Nakahara et al. is shifted to an area/lipid that is smaller than the limiting area. The isotherm of Nakahara et al. appears to exhibit material leakage from the monolayer, as indicated by a change in slope before the collapse plateau is reached. This isotherm is also distinct from the other isotherms shown in Figure 2.9, because it is obtained at low pH. At 303.15K, the isotherm of Baldgya and Dluhy appears to exhibit an inflection point (at 40Å^2/molecule), which could be caused by the onset of collapse at low collapse pressure (60mN/m), or by film leakage.

It is difficult to say which of the isotherms shown in Figure 2.9 are reliable, especially with so much variation among them and in the methods used to obtain them. It is more feasible to identify those that are most likely to exhibit experimental artifacts. Isotherms obtained at high and low pH (Kanintronkul et al. (pH 9) and Nakahara et al. (pH 2) at 298.15K), dynamic isotherms apparently exhibiting early collapse (Miñones et al., Sández et al., and Yun et al. at 293.15K, Dynarowicz-Latka et al. at 295.15K, Gladston and Shah and Kanintronkul et al. at 298.15K), isotherms that are suspected of exhibiting spreading-solvent artifacts (Taneva et al. 295.15K), and isotherms that are shifted to areas/lipid that are smaller than the limiting area of 39Å^2/molecule (Borissevitch et al. and Bordi et al. at 293.15K, Taneva et al. at 295.15K, Nakahara et al. and Gladston and Shah at 298.15K, Baldgya and Dluhy and Maskarinec et al. at 303.15K) are most likely to be misleading. This leaves the isotherms obtained by Ahuja and Möbius, Dubreil et al., and Williams et al. 293.15K; Hunt et al., Rana et al., and Slotte and Mattjus at 295.15K; Lee et al., Shen et al., and Tabak et al. at 298.15K; and Crane at 298.15K, 303.15K, and 323.15K as possibly the most trustworthy isotherms. These isotherms all exhibit the same general shape but vary in placement along the area/lipid axis and slope, with the latter likely

being affected by compression rate.

From Figure 2.9, it is clear that the simulated isotherms give areas that are too large, with limiting areas near 49Å^2/molecule, a phase transition plateau that is shifted upward to much higher surface pressures than those seen experimentally, and an overly large hysteresis loop. The steep slope upon expansion is typical of experimental expansion isotherms, which exhibit a sudden drop in pressure. The LE portion of the compression isotherm is steeper than the posttransition slopes seen in the experimental isotherms. Furthermore, the length of the coexistence region is much smaller for simulated compression isotherms than in experimental ones. Our LC-LE phase-coexistence plateaus occur at higher surface pressures than do those predicted experimentally and the surface pressures of our plateaus increase with increasing temperature, suggesting that the transition temperature for our simulated monolayers is likely too low. These factors indicate that our simulations do not accurately reproduce the behavior seen in experimental isotherms.

It should also be noted that some differences between simulations and experiments could be due to the absence of chain tilting in the simulations, because this absence leads to changes in area with increased pressure that are too small [111]. Also, experimental factors, particularly the choice of spreading solvent, not included in our simulations, may explain why our predicted isotherms exhibit abrupt changes and steep slopes that are uncharacteristic of experimental isotherms. Other atomistic [43, 72] and coarse grained models [92] also give pressure-area isotherms that are shifted to higher area/lipid relative to experimental values. Feller et al. attribute the differences between their results and experiments to difficulties in the evaluation of surface pressure, which depends on the accurate determination of long-range forces between atoms and has large instantaneous fluctuations [43]. They also suggest that the accuracy of simulated isotherm could be improved by better potential energy parameterization or by incorporation of long-range forces. Moreover, because the coarse

grained model lumps roughly four acyl tail carbon atoms into each tail bead, it is not able to capture the sensitivity of the phase transition temperature to chain length. Phillips and Chapman [64] showed that subtracting two methylene groups from each chain shifts the isotherm by an amount that is equivalent to raising the temperature by 20K. However, coarse grained lipids differing by only one or two methylene groups are represented by the same CG structure and thus cannot predict such effects. Also, as noted by Feller et al. [43], surface tension depends on the accurate determination of long-range forces, which are not considered by the coarse grained model. However, it is also important to consider inherent limitations associated with simulated isotherms due to system size and time scale limitations.

2.4.2 Area Compressibility Modulus

The compressibility (C_s) of the DPPC monolayer can be calculated from the slope of the pressure-area isotherm according to the following equation

$$C_s = -\frac{1}{A}\left(\frac{\partial A}{\partial \pi}\right)_T \tag{2.2}$$

where A and π are the area and surface pressure, respectively [139]. The area compressibility modulus (C_s^{-1}) is the reciprocal of the compressibility. Typical experimental values of the area compressibility modulus for DPPC monolayers are 10-50mN/m for LE films, 100-250mN/m for LC films, and >250mN/m for solid films [139, 140]. Here, the condensed and expanded phase moduli are approximated from the slopes of the experimental isotherms shown in Figure 2.9, using linear regression. Any moduli falling outside of the typical range are reported in Table 2.2. To avoid mislabeling a solid-phase modulus as a "high" value for the condensed phase modulus, in Table 2.2 we only report moduli for isotherms that did not have a kink, because a kink might indicate a transition to solid phase. Comparing the isotherms given in Table 2.2 with the experimental conditions listed in Table 2.1, we notice that

the isotherms giving LC moduli larger than typical condensed phase values were all compressed rapidly at rates of 0.5cm^2/min or at rates less than 6$\mathring{A}^2$/molecule*min, excluding the isotherm of Ahuja et al., which was compressed discontinuously. The rapid compression of these isotherms appears to be associated with their steep slopes and corresponding high compressibility moduli. Many studies have suggested that there is little or no variation in the shape of dynamic isotherms as compression rate is varied [97–99]. However, at high surface pressures, as the slope of the isotherm becomes nearly vertical, changes in the slope of the isotherm that may appear small can significantly increase the area compressibility modulus. Furthermore, faster compression rates are known to lead to the formation of smaller LC and LE domains, due to diffusion limited growth [99]. It is conceivable that this change in domain size could alter the compressibility of the monolayer.

The steep slopes of our simulated isotherms yield moduli that are larger than typical experimental values, and these values at 298.15K and 323.15K are given in Table 2.2 along with moduli approximated from the slopes from other simulated and experimental isotherms at 323.15K, from Figure 2.9. The corresponding areas at which the moduli were calculated are also given. For our simulation isotherms at 323.15K, a range of areas is given, because the moduli were evaluated at multiple points. The moduli were calculated by assembling results from independent runs at each pressure (as described in the simulation details section) and not from the cycling isotherms, because the slope of the expansion portion of the cycling isotherms is clearly too steep to give results that are comparable to experimental values. At 298.15K, the two end points of each isotherm, corresponding to smallest and largest area/lipid simulated, were selected to represent LC and LE phase moduli. At a temperature of 323.15K, the entire isotherm is in the expanded phase, and for this case the modulus was evaluated at each point along the isotherms. For comparison with our results, area compressibility moduli were approximated from the slopes of other

Table 2.2: Experimental and simulated area compressibility moduli. Moduli approximated from our atomistic, and coarse grained (CG) simulations, as well as from our CG simulations with the larger system size of 1024 lipids/monolayer, and approximated from the experimental and simulated isotherms of others. The following abbreviations are used: experimental (exp.), compression (comp.), and expansion (expan.). *Condensed phase values falling within this range have been reported at 293.15K [127, 141], 294.15K [142], 297.15K [140], 298.15K [105, 137, 139, 143, 144], and 310K [145]. Expanded phase values falling within the range have been reported at 293.15K [127], and 298.15K [105, 137, 139, 144].

	Temperature	Cs^{-1}	Area	Phase
Typical exp. values*	Varies	100-250mN/m	Varies	LC
Independent runs	298.15K	$\sim$ 363mN/m	47.5Å^2	LC
Larger system size	298.15K	$\sim$ 316mN/m	47.8Å^2	LC
Ahuja (exp.) [122]	293.15K	$\sim$ 326mN/m	44Å^2	LC
Williams (exp.) [128]	293.15K	$\sim$ 290mN/m	47.9Å^2	LC
Rana comp. (exp.) [130]	295.15K	$\sim$ 252mN/m	54.1Å^2	LC
Rana expan. (exp.) [130]	295.15K	$\sim$ 279mN/m	46.7Å^2	LC
Slotte (exp.) [131]	295.15K	$\sim$ 279mN/m	45.4Å^2	LC
Crane (exp.) [62]	298.15K	$\sim$ 293mN/m	44.6Å^2	LC
Crane comp. (exp.) [62]	303.15K	$\sim$ 313mN/m	45.9Å^2	LC
Crane expan. (exp.) [62]	303.15K	$\sim$ 265mN/m	47.1Å^2	LC
Typical exp. values	Varies	10-50mN/m	Varies	LE
Independent runs	298.15K	$\sim$ 169mN/m	63.9Å^2	LE
Independent runs	323.15K	$\sim$ 115-360mN/m	56.2-71.2Å^2	LE
Larger system size	298.15K	$\sim$ 41.5mN/m	67.7Å^2	LE
Atomistic	323.15K	$\sim$ 120-268mN/m	56.0-72.5Å^2	LE
Adhangale (CG) [77]	323.15K	$\sim$ 92-227mN/m	45.0-56.2Å^2	LE
Skibinsky (atom.) [36]	323.15K	$\sim$ 64.5-128mN/m	54-80Å^2	LE
Crane comp. (exp.) [62]	303.15K	$\sim$ 58mN/m	68.9Å^2	LE
Crane expan. (exp.) [62]	303.15K	$\sim$ 60mN/m	69Å^2	LE
Crane (exp.) [62]	323.15K	$\sim$ 67.3-168mN/m	57.2-80.7Å^2	LE

atomistic [36], coarse grained [77] and experimental [62] isotherms at 323.15K. When evaluating the isotherm of Adhangale et al. [77] the modulus was not approximated at the lowest area because there is a large jump in area between this and subsequent points.

As seen in Table 2.2, the compressibility moduli obtained for both coarse grained and atomistic simulations do not correlate well with those typically obtained from experiments. At 298.15K, the LE modulus obtained from our independent runs for monolayers composed of 256 lipids fell into a range expected for LC films, while our LC modulus was also too high, falling into the range of values expected for a solid film. At 323.15K, although the entire isotherm is considered to be expanded, the compressibility moduli obtained from our coarse grained (256 lipids/monolayer) and atomistic (64 lipids/monolayer) isotherms are again too large, again falling into the range expected for LC and even solid films. The values obtained from our atomistic simulations differ little from those obtained from our coarse grained simulations, although they give a narrower range of moduli which are slightly improved at low areas. The coarse grained and atomistic isotherms of Adhangale et al. [77] and Skibinsky et al. [36] also give moduli that are higher than those typically expected from experiments. However, it should be noted that the values obtained by Skibinsky et al. [36] correlate very well with those obtained from the experimental isotherm of Crane et al. [62], which also yield values of compressibility modulus that are larger than those typical of expanded films.

For our larger system size (1024 lipids/monolayer) we obtained an LC modulus that is lower by 15% than for the 256 lipids/monolayer isotherm, but still larger than the typical experimental values. However, the LE modulus is greatly improved in the larger system size, falling within the experimental range expected for LE isotherms. These results show that increasing system size decreases the area compressibility modulus, or conversely increases compressibility. This is to be expected because for a

larger system size, the surface can "wrinkle", adding to its ability to fluctuate in area, and thus increasing compressibility. Atomistic simulations performed on small bilayer patches also yield moduli that are significantly larger than the experimental estimates [146]. Marrink et al. [58] found moduli for a coarse grained DPPC bilayer at 323K, of 260 ± 40 mN/m for a bilayer composed of 6400 lipids and 400 ± 30 mN/m for a bilayer composed of 256 lipids. The difference in moduli for the two system sizes was attributed to the contribution of undulatory modes in the large system. Imposing a small box size is known to lead to artificial rigidity and suppressed undulations [7, 58, 80, 84, 147]. Monolayers and bilayers have different bending constants, and thus their undulations differ in magnitude, which should lead to different area compressibility moduli for monolayers and bilayers. However, they can be expected to react similarly to system size constraints. Applying a surface tension will decrease undulations and thereby reduce the undulatory contribution to the compressibility [147]. Thus finite size effects will decrease with increasing surface tension.

2.4.3 Effects of System Size, Time-Scale, and Hysteresis

There have been many studies of finite-size effects in lipid bilayers. De Vries et al. [84] found that for constant volume simulations of DPPC bilayers, the surface tension, electron density profile across the bilayer, and the carbon-deuterium order parameters, all converged to system-size-independent and time-independent values for a system size as small as 36 lipids/leaflet and a simulation time as short as 4ns. De Vries et al. [84] suggest that some finite size effects may be seen for systems larger than 36 lipids/leaflet. However, these are primarily due to the appearance of long-wavelength undulations. Klauda et al. [11] also found that a system size of 72 lipids (36/leaflet) was large enough to calculate accurately the structural properties (such as electron density profiles and deuterium order parameters) for a DPPC bilayer. In an earlier study, Lindahl and Edholm [147] found a slight system-size dependence in the

area/lipid when a cutoff method was used for evaluation of long range electrostatics, with a system of 64 lipids differing by 1.5Å^2 from a system 16 times as large. However, when a particle-mesh Ewald summation was used this difference was cut to less than 1Å^2 [148]. In a recent study of DOPC bilayers Castro-Román et al. [149] found that finite-size effects contributed very little to membrane structure, with virtually no differences observed between different system sizes in their neutron and x-ray scattering factors and scattering-length density profiles. Instead they suggest that force field inaccuracies account for large structural discrepancies between simulation and experiment.

Although finite size has little effect on the properties of leaflets composed of 36 or more lipids in the single-phase region, there are more serious finite-size effects when two phases co-exist. Experimentally, each co-existing region of liquid-condensed or liquid-expanded phase extends over distances of thousands of Angstroms. Simulating these biphasic systems in a box consisting of only hundreds of lipids lends concern over the magnitude of the line tension between such small domains and the correspondence of simulation results to experiment. The small size of simulated LC and LE domains raises uncertainty over whether such domains are stable. Experimentally, above the phase transition temperature, small nuclei can form known as hetero-phase fluctuations [150]. Due to the small size of these nuclei, there exists a large line tension, which opposes the thermodynamic driving force for the phase transition [7]. In their simulations of CG DPPC bilayers, Marrink et al. [7] observed fast fluctuations due to formation and disappearance of small clusters of the condensed gel phase. Marrink et al. [7] also observed long-lived fluid domains that remained trapped and metastable over a microsecond time scale with small defects persisting on even longer time scales. From their bilayer simulations, Marrink et al. [7] calculated a line tension between liquid crystalline and gel domains of 32pN, which matches within uncertainty the experimental value of 4pN estimated by the

kinetic model of Kharakoz and Shlyapnikova [150] for small gel clusters appearing in DPPC vesicles. For CG DPPC bilayers, Marrink et al. [7] found that regions smaller than a critical nucleus size of 10-40 lipids/monolayer, depending on the temperature, were unstable. The system size of our simulations is large enough to contain domains larger than the critical nucleus size reported by Marrink et al., but the nanosecond time-scale is shorter than that observed for meta-stable domains, suggesting that for our simulations the structures that form in the two-phase region are meta-stable. The meta-stable nature of two-phase structures in our simulations is also evident from the large hysteresis seen between our compression and expansion isotherms. Although the time and length scales are much different, it is important to keep in mind that dynamic experimental isotherms are also meta-stable. Experiments [151] have shown that the kinetics of the order-disorder transition are strongly dependent on heating and cooling rate, and under nonequilibrium conditions intermediate structures may form that differ from the equilibrium structure.

To study the effect of system size we ran simulations of monolayers composed of 256 lipids and 1024 lipids at 298.15K. Isotherms obtained from independent runs containing 256 lipids (black triangles) and 1024 lipids (purple circles) are compared in the bottom-left-hand side of Figure 2.9. Both system sizes gave the same isotherm, except at surface pressures below 30mN/m, where the larger system began to expand. The larger system also exhibited the onset of hole formation sooner (at a higher surface pressure) than the smaller system. These results correlate well with the finding of Knecht et al. [34] for an atomistic DPPC monolayer, which showed that in the LC-LE coexistence region increasing system size had little effect on the overall lipid order; however the rupture of the monolayer occurred at a smaller molecular area.

The effects of time-scale were also studied by comparing 20ns and 100ns cycling simulations at 303.15K. The isotherms obtained from these simulations are shown in the bottom-right-hand corner of Figure 2.9. Although the increased simulation time

yields little difference in the shape and position of the compression and expansion isotherms, there is a notable difference in the hysteresis. The 100ns cycling simulations undergo the LC-LE phase transition sooner (i.e. at higher pressure) than do the 20ns simulations, decreasing the size of the observed hysteresis loop. If the time scale of these simulations were increased arbitrarily, the hysteresis loops would be expected to narrow and eventually reach a true equilibrium value. However, even for the slower cycling, there is a marked difference between the hysteresis loops seen in our simulations and those of typical experiments. Although some experimental isotherms yield large hysteresis loops (Gladston and Shah; Figure 2.9 bottom left), most experimental hysteresis loops are much smaller (Rana et al. and Crane et al.; Figure 2.9, top and bottom right) than those seen in our simulations. Given the huge difference in time and length scales of our simulations compared to experiments, it would be computationally infeasible to carry out simulations that come significantly closer to attainment of the equilibrium isotherm [7].

2.5 Summary

Although many experimental pressure-area isotherms for DPPC monolayers have been reported, there is a large variation among them to which many factors might contribute, making comparison difficult and misinterpretation easy. We can make educated guesses about what causes a given isotherm to display a shift or characteristic shape; however with so many complex factors involved, the cause of the variations among experimental isotherms remains somewhat ambiguous. A high concentration of chloroform in the spreading solvent appears to be associated with a shift in the isotherm to low areas/lipid and a decreased ability of the monolayer to reach near-zero surface tensions. This could be due to the removal of DPPC molecules from the monolayer with chloroform upon compression, leading to a decrease in area and the appearance of premature collapse. High pH appears to shift the isotherm to a

larger area/lipid due to increased solvation, while low pH shifts the isotherm to lower areas/lipid reflecting hindered solvation. No defining trends associated with the type of apparatus used are evident from the isotherms studied here. Isotherms obtained in a trough show much variation and those obtained with the captive bubble apparatus exhibit no clear defining features that distinguish them from those obtained in a trough. Dynamic compression rate appears to play a role in the slope of the isotherms at high surface pressures (low areas/lipid), with a steeper slope and correspondingly larger area compressibility modulus as compression rate is increased. However, the large variability in experimental isotherms remains largely unexplained. Thus, it is clear that there is a need for some standardization to make experimental isotherms more interpretable and to make comparisons, both to simulated isotherms and among experimental ones, feasible.

Values of area compressibility modulus obtained for both coarse grained and atomistic simulations (ours and those of others) overestimate those typically obtained from experiments, although the disagreement diminishes somewhat as simulation box size increases. Thus, it is conceivable that a simulation of a macroscopic system size could produce moduli within the range of typical experimental values. Furthermore, experimental isotherms tend to show higher moduli when obtained by more rapid compression, which might also help explain the relatively high moduli obtained from simulations, which of course are obtained under very rapid compressions compared to typical experiments.

PO4-PO4, PO4-NC3, and NC3-NC3 radial distribution functions (RDFs) show little difference between the LC and LE phases, while C2-C2 distributions show a significant decrease in tail order as the monolayer is expanded, indicating that the structure of the DPPC headgroups is affected much less by the phase transition than is the structure of the DPPC tails. Furthermore, P-N tilt angle distributions obtained from our atomistic and coarse grained simulations give an average P-N orientation

that is parallel to the interface and is not significantly affected by the LC-LE phase transition. In accord with experimental observations, this provides further evidence that the DPPC headgroup region is not strongly affected by the transition from LC to LE phase.

The coarse-grained NC3-NC3 and atomistic N-N RDFs differ considerably from each other and the coarse-grained NC3-NC3 RDF matches closely that of PO4-PO4, although the corresponding atomistic N-N and P-P RDFs differ considerably. Thus, the coarse grained model is unable to capture the difference between N-N and P-P interactions present in the atomistic simulations, which ultimately leads to inaccuracy in the coarse-grained NC3-NC3 RDF. Furthermore, the first correlation peak of the atomistic P-P distribution broadens as the monolayer undergoes a transition from the LC to LE [34]. This distinction is not seen in our coarse-grained PO4-PO4 radial distribution function at 298.15K. These results suggest that the coarse grained model is better at capturing the effect of changing surface pressure on lipid tails than on lipid headgroups.

Despite some limitations, molecular simulation could be a key to obtaining a more detailed understanding of the complex mechanisms involved in the phase transitions of DPPC, of other physiologically relevant lipids, and of mixtures of lipids and proteins. Simple coarse grained models, such as that developed by Marrink et al., are powerful tools for studying such systems, on length and time scales that are difficult or impossible to obtain using atomistic simulation. Using the coarse grained model of Marrink et al., we were able to quickly obtain compression and expansion isotherms for DPPC at five different temperatures and visualize the changes in packing from hexagonal to disordered as the DPPC monolayer underwent a phase change from the liquid-condensed (LC) to the liquid-expanded (LE) state. These fast and simple simulations provide a tool for comparison to experiment and clarification of the possible mechanisms involved in the rich phase behavior of DPPC. However, there

are limitations, shown by the fact that even the atomistic simulated isotherms tend to be shifted to higher areas/lipid than experimental ones and do not exhibit the correct shape. To obtain more accurate simulated isotherms, more work is needed on either potential energy parameterization or the evaluation of long-range forces, for both coarse grained and atomistic models.

CHAPTER III

Folding Transitions in Lipid and Lipid-Peptide Monolayers.

3.1 Introduction

Lung surfactant (LS) is a mixture of phospholipids, fatty acids, neutral lipids, and surfactant proteins that forms the surface-active lining in the lungs and decreases the work of breathing by reducing and regulating the surface tension in the alveoli. Lung surfactant consists of approximately 90% lipids and 10% proteins by weight [45]. Of the surfactant lipids, about 80% are phosphatidylcholines, about half of which is dipalmitoylphosphatidylcholine (DPPC, phosphatidylcholine with two palmitic acid tails, also known as dipalmitoyl lecithin or DPL [45]. Infants born prematurely lack functional lung surfactant and develop respiratory distress syndrome (RDS). Surfactant replacements have greatly reduced the mortality rate of RDS, but are not optimal [46]. In particular, acute respiratory distress syndrome (ARDS), which can develop in adults, has proven difficult to treat using surfactant replacement therapy [51]. Further research is needed to understand the mechanism by which lung surfactant is able to reduce and regulate the surface tension in the lungs, in order to develop more efficient surfactant replacements.

To be effective, lung surfactant must display rapid adsorption, the ability to compress to near-zero surface tension upon end-expiration, and rapid respreading upon film expansion [53]. The primary component of lung surfactant, DPPC, is able to

reduce the air-water surface tension to nearly zero; however under physiological conditions it is in a liquid condensed (LC) phase, and thus is rigid and exhibits poor respreadibility [152]. Other surfactant components enhance the surface activity of LS. Unsaturated phospholipids, such as palmitoyloleoylphosphatidylglycerol (POPG), fluidize lung surfactant surface films increasing adsorption to the interface and are thought to enhance respreading [45]. Neutral lipids and fatty acids are also present. Cholesterol constitutes (5-10wt%) of native surfactant. It has been suggested that cholesterol modulates the phase behavior of surfactant membranes by disordering phospholipids in the LC phase and ordering phospholipids in the LE phase, resulting in the formation of the liquid-ordered (Lo) and liquid-disordered (Ld) phases, respectively [54]. Cholesterol is an important component of lung surfactant, but is systematically removed from most surfactant replacements [54]. In contrast, palmitic acid (PA) constitutes only a small fraction of native surfactants extracted by lavage (0-3wt%) [153], but is used as an additive in surfactant replacements such as Surfacten and Survanta to enhance film stability [46].

Lung surfactant contains surface-associated surfactant proteins SP-B and SP-C, which are highly hydrophobic and amphipathic. SP-B contains seven cysteines, which form three intramolecular disulfide bridges and an intermolecular one leading to the formation of a homodimer [154]. SP-B contains more polar and hydrophilic residues than SP-C and is therefore less hydrophobic. SP-C contains palmitoyl chains on cysteine residues 5 and 6 [155]. The presence of these hydrophobic surfactant proteins SP-B and SP-C is essential to modulate the physical properties of the surface film and to promote the rapid formation of surface films capable of reaching near-zero surface tensions under repetitive cycling [156]. SP-B and SP-C significantly promote adsorption and are thought to promote respreading of phospholipids to the interface; see Refs [46, 54] and references therein. Electron micrographs of rabbit lungs [157] have confirmed the existence of a surface associated surfactant reservoir, which is thought

to be essential for rapid respreading upon expansion. The surfactant reservoir can form by either adsorption or collapse [46]. SP-B and SP-C have been shown to facilitate the formation of reversible surface associated 3-dimesional collapse structures [30, 53, 158–175]. These collapse structures act as reservoirs, into which material is inserted upon compression and from which material can be reversible reincorporated into the monolayer upon re-expansion. However, the exact roles of SP-B and SP-C are still a matter of speculation. The tertiary structure of SP-B is unknown; however the structure is known for the 25-residue N-terminal fragment (SP-B_{1-25}), which retains most of the functionality of the full-length protein, including the ability to form 3D collapse structures [158–162].

A common feature of almost all lung surfactants and model mixtures is the coexistence of a semi-crystalline solid phase known as the liquid condensed (LC) phase and a disordered fluid phase called the liquid expanded (LE) phase [176]. In the LC/LE phase coexistence region, the surface film becomes a mesh of finely divided LC/LE domains, which may impart film strength and flexibility [177]. Liquid expanded regions provided fluidity to normally rigid, condensed DPPC films, while condensed regions may act as "splints" in the alveolar wall, preventing collapse with decreasing alveolar volume [178]. It should be noted that lipids in the liquid condensed phase (LC) may be tilted or untilted. Lipids that are tilted are sometimes referred to as tilted condensed (TC). Since all of the lipids in our simulations are untilted, for simplicity the general term (LC) is used through out this dissertation to represent the condensed phase.

A prevalent theory known as the "squeeze-out" theory holds that to reach near-zero surface tensions the monolayer surface must be refined through the "squeeze-out" of fluidizing non-DPPC components; leading to a dense condensed monolayer of the saturated lipid DPPC [119, 179, 180]. These reports on classical "squeeze-out" were published before the first reports on the existence of hydrophobic surfactant proteins

[181, 182]. Later studies by Schürch et al. [157] and Post et al. [183], first suggested that partial squeeze-out can result in the formation of a reservoir where non-DPPC components such as surfactant peptides and unsaturated phospholipids remain readily available for re-insertion into the interface upon re-expansion. In addition, there have been a number of subsequent studies referring to "squeeze-out", however these studies do not necessarily imply that the monolayer is refined to become highly enriched in DPPC as held by classical "squeeze-out" theory.

Some recent experimental observations undermine classical squeeze-out theory; see Ref [46] for a review. Experimental studies involving spread films suggest no significant differences in lipid composition between the multilayer, the adsorbed interfacial monolayer, and the surfactant in the bulk phase [184]. Additional studies have shown that near-zero surface tensions can be obtained by films consisting of coexisting LC and LE phases [185]. Even monolayers composed only of the unsaturated phospholipid palmitoyloleoylphosphatidylcholine (POPC) have been shown to reach near-zero surface tension, if compressed quickly enough [186]. Therefore the widely held idea that to obtain low surface tension the surface has to be compressed into a tightly packed liquid-condensed or solid phase, and therefore must be enriched in DPPC, is under substantial scrutiny. A few theories have been proposed to explain the ability of films containing a significant amount of LE phase to reach near-zero surface tensions. In one such theory, the LE domains form a matrix in which LC nanodomains are uniformly distributed, producing an alloy or composite material that is both flexible and stable [53]. This alloy structure is thought to allow the partial collapse of the monolayer into multilayers, without preventing the attainment of near-zero surface tensions [53]. It has also been suggested that the 3D architecture of the collapse structures may provide additional stability against further collapse by acting as a skeleton or scaffold, thereby allowing the attainment of lower surface tensions [30, 53, 187]. The proposal that SP-B and SP-C may bridge the multilayer

collapse structures to the interfacial monolayer is consistent with this idea [46]. Hall and co-workers [186, 188, 189] offer a different interpretation, that monolayers compressed rapidly undergo a "supercompression" that is analogous to the supercooling of 3D liquids toward a glass transition. Monolayers undergoing "supercompression" do not have enough time to collapse and so form an amorphous structure that retains some disorder, but exhibits the same stability against collapse as the LC phase.

A mechanistic understanding of the ability of lung surfactant to reach near-zero surface tension remains elusive. Therefore, the structure and organization of lung surfactant in the low-surface-tension regime is of particular interest. The collapse dynamics of Langmuir monolayers, and specifically the role of proteins in the formation and respreading of 3-dimensional collapse structures, remain somewhat mysterious, due in part to the richness of collapse behavior observed under various conditions. This behavior includes both reversible and irreversible collapse, collapse toward the air or water subphase, the formation of crystallites, folds, vesicles, multilayers, and bilayer stacks. For more information on lung surfactant structure and function, readers are referred to detailed reviews on the phase behavior, collapse transitions, and biophysical properties of lung surfactant [46, 54, 55, 156, 190–192].

Despite intensive research, the complex action of natural lung surfactant remains poorly understood[193]. Computer simulations of phospholipid systems are of great interest because they can yield molecular-level insight into the structure and dynamics of these systems at small spatial and temporal scales not accessible experimentally. There are a few atomistic simulations that provide insight into the orientation and interactions of SP-B_{1-25} in DPPC [35, 37] and PA [194, 194, 195, 195] monolayers. Atomistic simulations of monolayer collapse have also been performed for phospholipid [27] and fatty acid [196] monolayers. Atomistic simulation of arachidic acid monolayers (up to 196 arachidic acid molecules per monolayer) compressed by two walls under a very large surface pressure (4426mN/m) collapsed either into the sub-

phase or into the air side of the interface leading to the formation of multilayers [196]. However the displacement of some of the arachidic acid below the monolayer leading to the formation of a multilayer on the air side of the interface is likely a consequence of the extreme surface pressure. Atomistic simulations of small DPPC, DPPC/POPG, DPPC/POPC, and DPPC/POPA monolayers (up to 64 lipids per monolayer) compressed to low lateral areas display buckling and the loss of a few lipids from the monolayer [27]. However these monolayers are limited in their ability to produce collapse structures by their small size, and are likely prone to system-size artifacts. In general, the length and time scales covered by atomistic simulations are too small to overlap with experimental time and distance scales, leaving a "no mans land" of intermediate spatial and temporal scales that are not accessible either experimentally or by atomistic simulations.

To help bridge part of this gap, much faster (around 1000-fold) coarse grained (CG) molecular simulations have been utilized in which three or four heavy atoms (such as carbons) are lumped into a single "bead", whose interactions are tuned to capture those of the group of atoms represented by the bead. Such simulations can achieve longer time scales for larger system sizes than are readily attained with atomistic simulations. In simulations performed by Nielsen et al. [92] monolayers of CG short and long-tail phospholipids ($DC_{14}PC$ and $DC_{29}PC$) exhibited collapsed into the air side of the interface facilitated by a lipidic bridge transport mechanism. The collapse mechanism observed was initiated under large negative surface tensions (-300mN/m) and displayed system size limitations. In larger system size simulations (250 lipids per monolayer) extreme curvature of the interface was evident and thought to be the onset of a collapse transition into the subphase [92]. Baoukina et al. [29, 31, 32] have applied the MARTINI CG model to study the collapse behavior of DPPC [31], DPPC/POPG [29], and DPPC/POPG/cholesterol/SP-C [32] monolayers. These simulations attained longer length and time scales then the aforementioned atomistic

and CG simulations and showed that in response to relatively small negative surface tensions the monolayers collapse into the water subphase by forming attached bilayer folds. Baoukina et al. demonstrated the ability of the MARTINI CG model to capture 2D to 3D transitions occurring at the monolayer interface. However, the role (if any) that the proteins played in the observed collapse was not clear. The role of proteins in reservoir formation and maintenance still remains to be addressed through molecular simulation. We therefore begin to address this role here. The remainder of this chapter is outlined as follows: First, we provide details of our simulations, then present the simulation results, and finally compare them with experimental results.

3.2 Simulation Methods

The MARTINI CG model [56, 57] was used in all simulations reported here. The CG mapping of DPPC, sodium ions, chloride ions, and water molecules is the same as that provided in the topology files on the MARTINI website [79]. The Perl script seq2cgtop_martini_v2.1tryout.pl [79] was used to generate the topology for the CG SP-B_{1-25} and SP-C. The structure files for SP-B_{1-25} and SP-C were created by coarse-graining structure files for SP-B_{1-25} [197] and SP-C [198] that were obtained by FTIR and NMR spectroscopy. Because all of the lipid head group and tail beads are the same size, the lipids are always untilted [7]. Marrink and co-workers have shown that tilted DPPC phases can be simulated using the CG model, if the tail group bead size is decreased by 10%, Marrink et al. [7]. The MARTINI model is parameterized based on thermodynamic data, and has successfully reproduced membrane properties such as area per lipid [56, 58], pressure-area isotherms [28], phase transitions between lamellar and non-lamellar phases [199, 200], self assembly of bilayers, and structural and dynamic features of protein-lipid interactions [57]. Furthermore, the MARTINI model yields profiles of lateral pressure versus vertical position in the monolayer that are qualitatively similar to those obtained from atomistic simulations, suggesting that

the MARTINI model captures the essential lipid/solvent properties [24].

We use three system sizes; by lateral replicating the smallest system size described below, we also built systems 4x and 9x larger in area. For the smallest system size, each configuration was constructed from two disordered lipid monolayers (each composed of 256 DPPC molecules) placed with heads facing each other across a layer of water, and tails separated by vacuum, in a periodic box, as described previously [28]. The z-dimension, normal to the layers, was adjusted to 100nm, which allows enough space for folding to occur without the tail regions of the two monolayers interacting. The SP-B_{1-25} molecules were then placed in each monolayer, oriented normal to the interface with the insertion sequence, which is the last eight residues on the N-terminus side, placed close to the head group region. Four peptides were inserted into each monolayer in three initial configurations: 1) with the peptides clustered together (not in contact, but with each peptide 1.7nm from the center of the box), 2) in a line (separated by 3.3nm), and 3) in a square (with each peptide placed 4.5nm from the center of the box). CG chloride ions were then added to make the system electroneutral. The system was then energy minimized. Larger system size simulations (4x or 9x) were started from the square initial configuration. Simulations of the smallest system size were run using the all three initial configurations; those containing SP-B_{1-25} were run from all three configurations, those containing SP-C were run from square and cluster configurations, and those containing mutants were run from line and cluster configurations. The occurrence or absence of folding in each peptide-containing monolayer was found to be consistent for all initial configurations tested. All figures and folding times reported, for the smallest system size, were obtained using the cluster initial configuration, unless otherwise noted.

All simulations with mutants were adapted from the cluster and line configurations, by making substitutions, insertions, and deletions of CG beads where necessary, and then performing energy minimization. The amino acid substitutions were chosen

to adjust the hydrophobicity of the selected residues while changing only minimally the residue structure (e.g., the number of CG beads per residue, or the presence of a ring structure). The parameters for the substitute residues were taken from the amino acid topology file [79], without altering any parameters associated with the secondary structure of the peptide. The insertion sequence was removed from some of the mutant peptides. Because the mutant peptides retained the same positioning as SP-B_{1-25}, these mutants were initially in contact with the monolayer, but not embedded into the monolayer as much as the other peptides. SP-C was placed in the monolayer in the same manner as SP-B_{1-25}, but with the α-helix placed in the subphase initially tilted with respect to the interface and with the palmitoyl chains inserted into the monolayer. To test the effect of initial peptide conformation, a monolayer containing SP-C with the α-helix initially embedded in the lipid tail region was also simulated. The CG topology and structure files for POPG were adapted from the lipid topology file and a palmitoyloleoylphosphatidylethanolamine (POPE) bilayer CG structure file taken from the MARTINI website [79], by replacing the ethanolamine head group bead (Qd) with a glycerol head group bead (P4). A system containing two POPG monolayers was constructed in the same way as for the DPPC monolayers. To create 1:1 DPPC:POPG monolayers, POPG molecules were randomly replaced with DPPC molecules to obtain a 1:1 mixture and the system was then energy minimized. In mixtures containing PA, 42 DPPC molecules in each monolayer were randomly replaced by PA molecules and the system was then energy minimized. The topology and structure files for PA were obtained by taking only one tail of DPPC and replacing the glycerol bead (Na) with a (Qa) bead to reflect the anionic charge of PA. For systems containing neutral PA the (Na) glycerol bead was used instead.

Although the simulated peptide concentrations are larger that the average physiological concentrations, as noted previously [201], local concentrations of physiological

lung surfactant components could be much higher than the physiological average, particularly if interactions between components result in non-uniform distributions.

For all simulations, temperature was maintained by coupling to a Berendsen thermostat [67] with a 1ps time constant. Berendsen pressure coupling was used with a 1ps time constant and all compressibilities set to 5E-6bar^{-1}. For all simulations, a timestep of 0.02ps was used, and periodic boundary conditions employed. All simulations and analyses were performed using GROMACS simulation software [68, 69, 202]. The trajectories were saved every 0.1ns and used for analysis. The following parameters were taken from the MARTINI website [79] and have been optimized for the coarse grained model: short-range electrostatic and van der Waals cutoffs of 1.2nm, with van der Waals interaction shifting smoothly to Lennard Jones interaction at 0.9nm, and with the Lennard Jones cutoff set to 1.2nm. The neighbor list was updated every 10 steps using a grid with a 1.2nm cutoff distance. In addition to the large computational speed-up, the molecular diffusivities of CG water and lipid molecules are around four times higher than for atomistic ones. As a result, the "effective time" of a simulation is roughly four times longer than the "physical" time [56]. However, all times reported here are physical times, as reported by the simulation.

Anisotropic pressure coupling was used for most of the simulations. Surface tension (γ) is calculated according to the following equation [31]:

$$\gamma(t) = \frac{L_z}{2}\left(P_{zz}(t) - \frac{P_{xx}(t) + P_{yy}(t)}{2}\right) \tag{3.1}$$

where t is time, L_z is the length of the box in the z-direction, and P_{xx}, P_{yy}, and P_{zz} are the pressures in the x, y, and z directions. Most of the simulations were run at small negative surface tensions, with the pressure in the x and y directions set to 1bar, the off-diagonal pressures to 0bar, and the normal pressure (P_N; equal to P_{zz}) to 0bar due to the vacuum between the monolayers. This corresponds to a surface tension set to -5mN/m; however due to pressure fluctuations the actual time-average surface

tension differs slightly from this set point. A limited number of simulations were also run with surface tension set to zero, where pressure was coupled anisotropically to 0bar in all directions. The average surface tension tends to be slightly higher than the set point; for instance with surface tension set to -5mN/m, the average surface tension is just above 0mN/m, and when surface tension set 0mN/m, the average surface tension a few mN/m. The standard error is 1-2mN/m. Our pressure fluctuations are similar to those reported previously [32]. Throughout the rest of this chapter all reported pressures are set-point values rather than actual pressures. To test the robustness of our results, a limited number of isotropic (P=1bar) and semi-isotropic (with a lateral pressure P_L=1bar and normal pressure P_N=0bar) simulations were also performed. For more details on each coupling mechanism the reader is referred to the GROMACS User Manual [70] and relevant simulation papers [41, 43, 75, 76, 78].

Monolayer re-expansion was simulated using semi-isotropic pressure coupling with P_N=0bar and P_L set to -5bar and 10bar. As reported previously, at large values of surface tension, the box size diverges and eventually explodes due to the onset of hole formation followed by expansion and ultimately the rupture of the monolayer [28, 31, 34, 41, 74].

Our simulations were 500ns in duration for the smallest system size and 300ns for the 4x and 9x larger systems. For the larger system sizes, some simulations which displayed folding sooner than 300ns where stopped early. Unless stated otherwise, all simulations were performed at 323K and a small negative surface tension (-5mN/m), as described above. This temperature was chosen to allow the development of thermal undulations in all monolayers including DPPC monolayers, which are in the LE phase at this temperature.

The MARTINI model for DPPC consists of two headgroup beads, two beads that represent the glycerol linkage, and eight tail group beads (four per tail). Lipid tail beads of are numbered in increasing order as the distance from the headgroup

increases (C1, C2, C3, etc.). For phospholipids, a further designation is included to indicate whether the tail bead is on chain A or chain B. Chain A is the chain attached to the glycerol bead, which is directly bonded to the phosphate headgroup bead, and chain B is attached to the other glycerol. Lipid tail order parameters were calculated every 0.2ns between 10 and 20ns of simulation time. Our results have shown that 10ns of equilibration is necessary to allow relaxation of the monolayer area. Unless stated otherwise, the reported order parameters are averages over all C2 tail sites in the mixture. The average order parameter of the C2A beads is slightly higher than that of the C2B beads. We found that averaging over all C2 sites yielded similar trends as averaging over C2A sites only, with only a very slight shift in magnitude. Also, order parameters for C3 tail sites show similar trends to those for C2 tail sites, but C3 sites are more disordered, and therefore give lower order parameters than the C2 sites.

3.3 Results

3.3.1 Folding

We performed CG simulations of lipid and lipid/protein monolayers of varying composition at 323K and small negative surface tension (γ set to -5mN/m), as listed in Table 3.1. Three system sizes were used. The smallest system contained 256 lipids with or without 4 peptides per monolayer. Systems containing 4 or 9 times as many molecules per monolayer were also simulated. For the smallest system, folding was not observed within 500ns of simulation time for the pure lipid monolayers (DPPC and 1:1 DPPC:POPG). Once the system size was increased to the intermediate system size, the 1:1 DPPC:POPG monolayer did fold (Figure 3.1(a)). In contrast, the pure DPPC monolayer did not fold even for the largest system size (Figure 3.1(b)). The addition of SP-B_{1-25} or SP-C to DPPC monolayers induced folding, even for the smallest system

size. Figure 3.1 (c and d) illustrates folds formed in DPPC monolayers containing SP-B_{1-25} at an intermediate system size. The folds are flat bilayers in one direction (Figure 3.1(d)) and teardrop shaped in the other (Figure 3.1(c)).

Table 3.1: Simulations performed at 323K and small negative surface tension. The surface tension is set to -5mN/m. The compositions are given in mol:mol ratios.

Composition	**System-size (per monolayer)**	**Folding**
DPPC	256 lipids	NO
DPPC with SP-B_{1-25}	256 lipids and 4 peptides	YES
DPPC with SP-C	256 lipids and 4 peptides	YES
1:1 DPPC:POPG	256 lipids	NO
1:1 DPPC:POPG with SP-C	256 lipids and 4 peptides	YES
1:1 DPPC:POPG with SP-B_{1-25}	256 lipids and 4 peptides	YES
~5:1 DPPC:PA with SP-C	256 lipids and 4 peptides	NO
~5:1 DPPC:PA with SP-B_{1-25}	256 lipids and 4 peptides	YES
~5:1 DPPC:PA (neutral) with SP-C	256 lipids and 4 peptides	YES
~5:1 DPPC:PA (neutral) with SP-B_{1-25}	256 lipids and 4 peptides	YES
~5:1 DPPC:PA with depalmitoylated SP-C	256 lipids and 4 peptides	YES
DPPC	1024 lipids	NO
1:1 DPPC:POPG	1024 lipids	YES
DPPC with SP-B_{1-25}	1024 lipids and 16 peptides	YES
DPPC	2304 lipids	NO
1:1 DPPC:POPG	2304 lipids	YES
DPPC with SP-B_{1-25}	2304 lipids and 36 peptides	YES
DPPC with SP-C	2304 lipids and 36 peptides	YES

For the smallest system, we observed folding for all of the mixed protein/lipid monolayers except for the 5:1 DPPC:PA monolayers containing SP-C (Figure 3.2(a)). In this case, the addition of the palmitic acid eliminates the folding that is seen in the DPPC monolayers containing only SP-C. In contrast, the addition of palmitic acid does not restrict folding in the DPPC monolayers containing SP-B_{1-25} (Figure 3.1(e)) as it does for those containing SP-C. Unlike SP-B_{1-25}, SP-C contains two palmitoyl chains. In order to test the effect of the palmitoyl chains, we ran a simulation containing 5:1 DPPC:PA monolayers with depalmitoylated SP-C and found that folding did in fact occur once the palmitoyl chains were removed. Also, if PA is neutralized, by changing the bead type of the glycerol bead from Qa to Na, folding occurs in all sys-

tems studied including the DPPC/PA monolayers containing SP-C (Figure 3.2(b)). Interestingly, once 5:1 DPPC:PA(neutralized) with SP-C have run for as little 20ns of simulation time (resulting in small undulations and loose aggregation of the peptides), if the palmitic acid is suddenly re-charged, then folding still occurs. Thus, charged palmitic acid and SP-C act together to suppress the initial fluctuations needed for folding. Even brief neutralization of PA suffices to produce fluctuations of sufficient size that folding will occur even if the PA is thereafter recharged.

Monolayers containing 256 DPPC and 4 SP-B_{1-25} molecules were started from configurations where the peptides were place clustered together, in a line, and in a square, as described in the simulation method section. Different pressure coupling methods (anisotropic, semi-isotropic, isotropic) were also applied to the monolayers containing 256 DPPC and 4 SP-B_{1-25} molecules. Monolayer folding was found to be robust, occurring irrespective of pressure coupling method and initial configuration, and differing only in time scale. For instance, when the peptides are placed further apart from one another, aggregation requires more time, resulting in folding at later times. Decreasing the concentration of SP-B_{1-25}, however, decreases the perturbation of the monolayer and slows or eliminates the folding. In DPPC monolayers containing only 3 SP-B_{1-25} molecules each, folding occurred more than 200ns later than the same system containing 4 SP-B_{1-25} molecules per monolayer. If the concentration was further reduced to 2 SP-B_{1-25} per monolayer, folding did not occur within the 500ns of simulation time.

3.3.2 Peptide Aggregation

For monolayers of the smallest system size, we observed peptide aggregation, and found that a defect is required to nucleate a fold. The aggregation of the peptides can provide such a defect. Once an aggregate has formed, undulations centered on the aggregate begin to grow (Figure 3.3(a)) leading to fold formation. The aggregate typ-

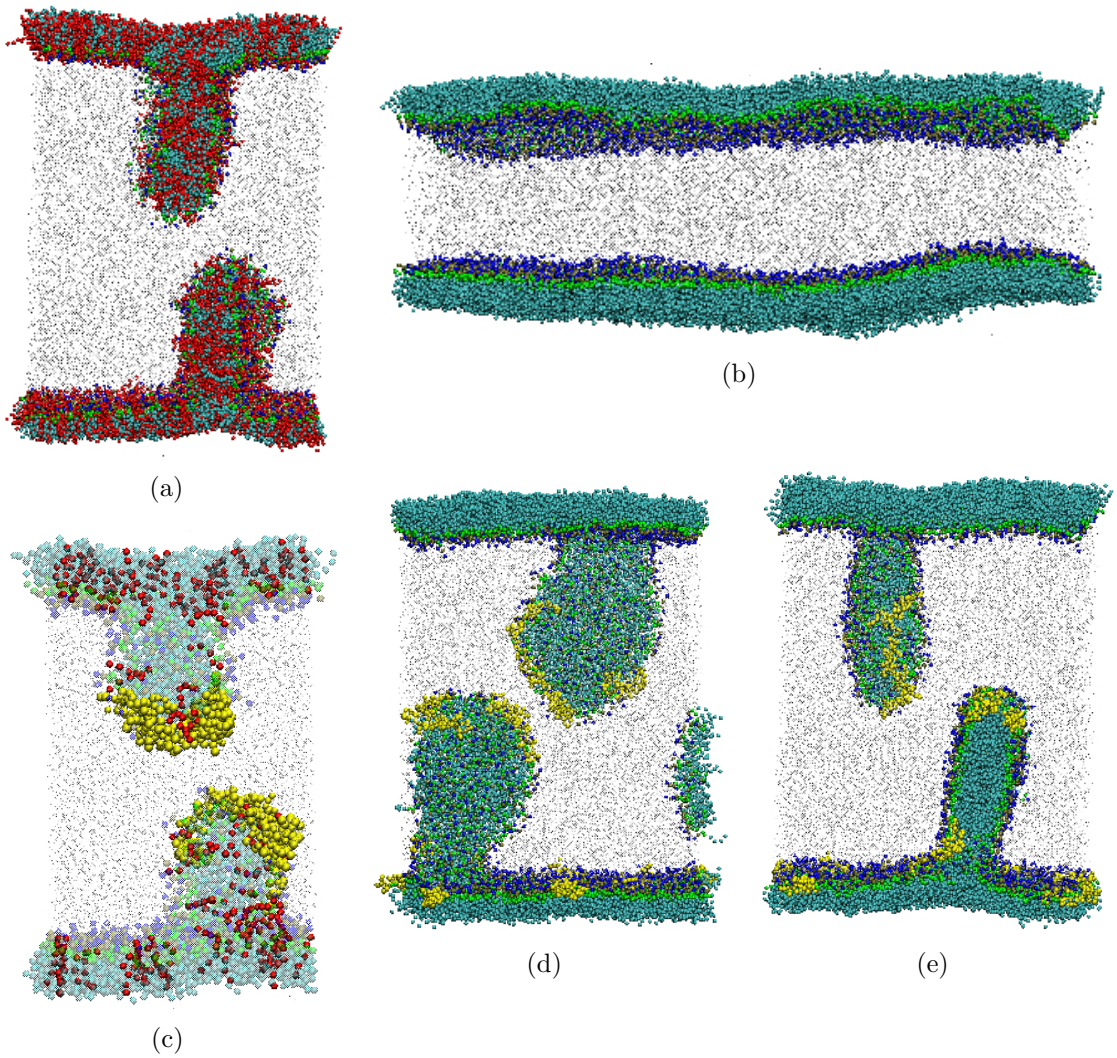
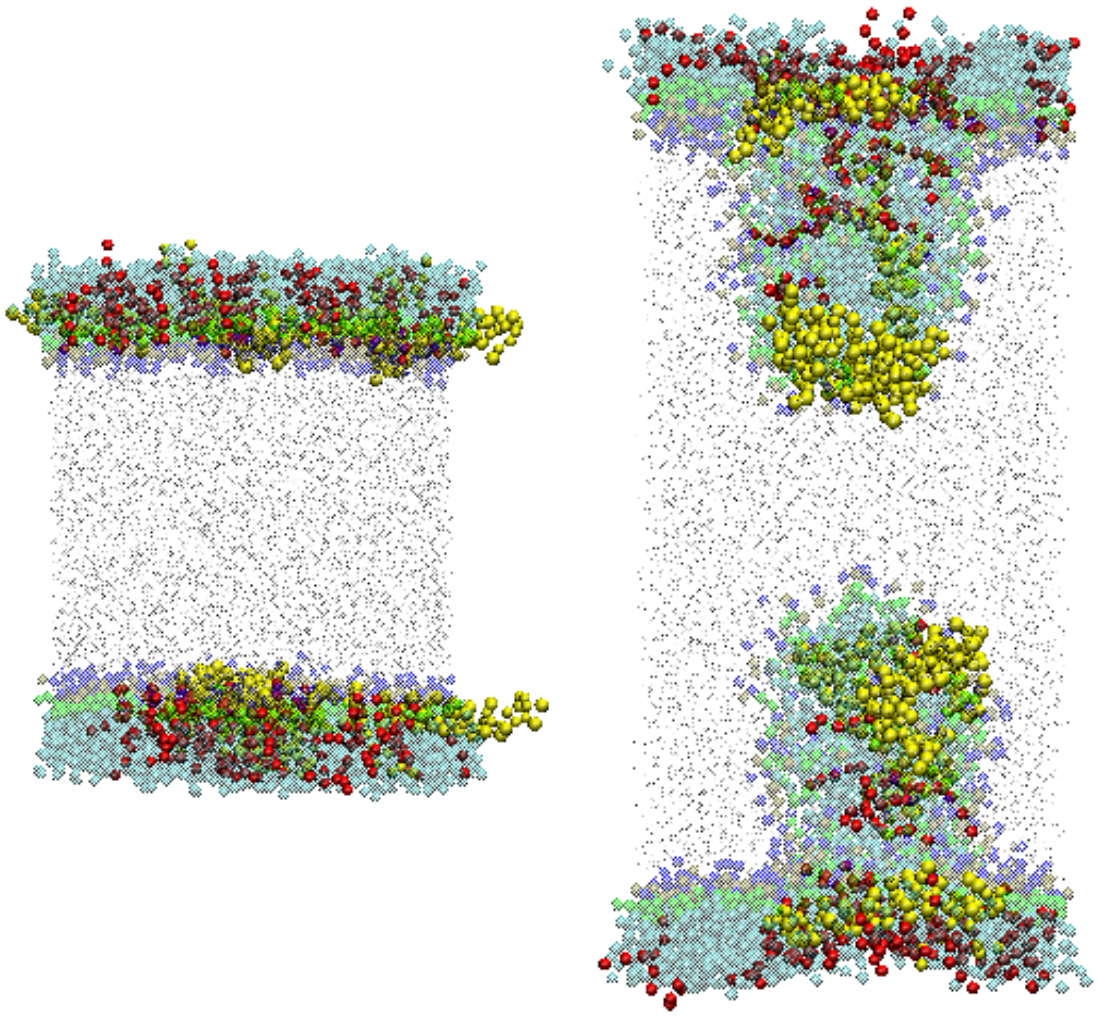


Figure 3.1: The occurrence or absence of folding in the simulated monolayers. (a) At an intermediate system size of 1024 lipids per monolayer 1:1 DPPC:POPG monolayers exhibit folding. (b) In contrast, even at the largest system size simulated (2304 lipids per monolayer) pure DPPC monolayers do not fold within 300ns of simulation time. Folds also form in monolayers composed of (c) 214 DPPC, 42 PA and 4 SP-B_{1-25} molecules per monolayer and (d-e)1024 DPPC and 16 SP-B_{1-25} molecules per monolayer. The folds formed in the DPPC monolayers containing SP-B_{1-25} are shown from (d) the side and (e) the front. In this and subsequent figures DPPC headgroups are blue, DPPC glycerol linkages are green, DPPC tails are turquoise, peptides are yellow, POPG and PA are red and small dots represent the water sub-phase. DPPC is transparent in some figures for clarity.

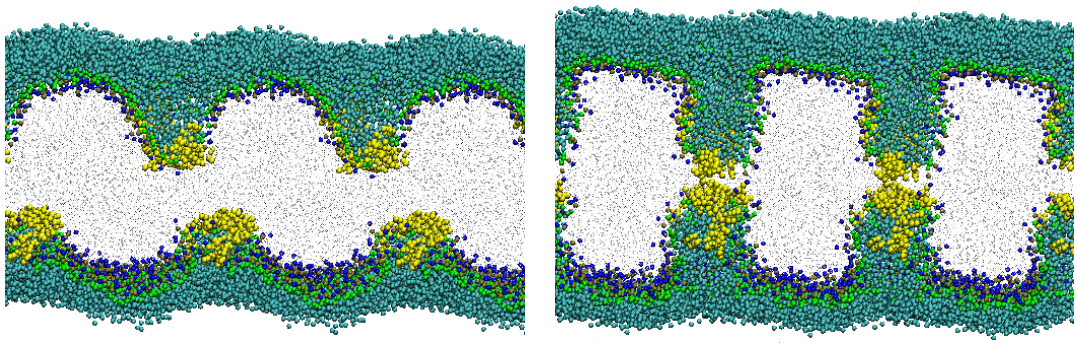


(a) (b)

Figure 3.2: Folding or the absence of folding in monolayers containing PA.(a) Monolayers composed of 214 DPPC, 42 PA and 4 SP-C molecules per monolayer do not exhibit folding after 500ns of simulation under small negative surface tension. (b) However, these monolayers do exhibit folding if the PA is neutralized. To test the effect of the palmitoyl chains, we also simulated monolayers composed of 214 DPPC, 42 PA and 4 depalmitoylated SP-C molecules. Folding did in fact occur once the palmitoyl chains were removed from SP-C (not shown).

ically forms a cap on the monolayer fold and acts like a zipper pulling the surrounding lipids into a bilayer fold. We find that while some peptides are more prone to aggregate (i.e., they display faster aggregation and more compact aggregates) than others, even peptides that do not appear to have a strong preference for aggregation exhibit

some aggregation as the monolayer buckles. This suggests that peptide aggregation is induced by the curvature of the monolayer.



(a) (b)

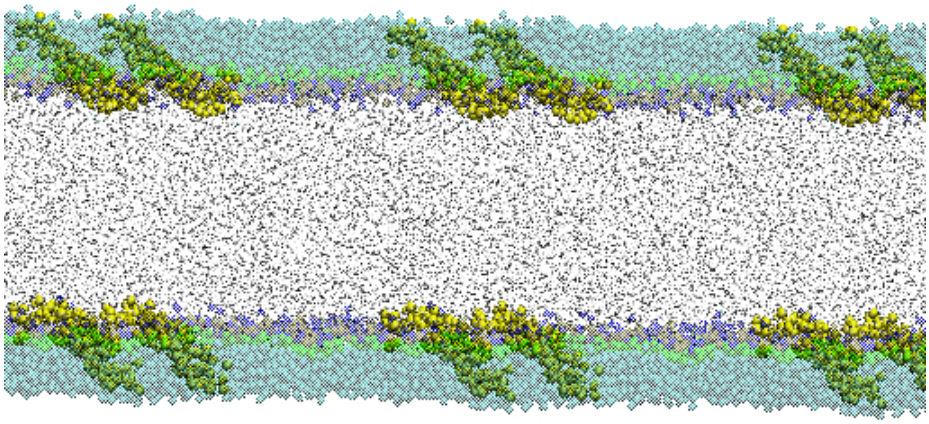
Figure 3.3: The occurrence of folding in two monolayers separated by water, containing 256 DPPC and 4 SP-B_{1-25} molecules each. Periodic images are shown for clarity. Undulations centered around peptide aggregates form after 140ns (not shown) and folds begin to grow. The peptides form a cap on the end of each fold. (a) After 190ns, fold growth is evident. (b) After 206ns of simulation, peptides in folds from opposing monolayers interact with one another eventually leading to fusion of the two folds into a lipid bridge. This simulation was started with peptides initially positioned at corners of a square.

In the simulations described in Tables 3.1 and 3.2, all peptides were initially partially inserted at the interface but mostly in the water subphase. The SP-B_{1-25} molecules (and mutants) were oriented normal to the interface with the insertion sequence placed close to the head group region. SP-C was placed with the palmitoyl chains inserted into the monolayer and the rest of the protein protruding out into the subphase and tilted with respect to the interface. The peptides quickly move into the interface, and adopt a final orientation that is parallel to the interface. However, this interfacial orientation could reflect a local energy minimum and not necessarily the equilibrium orientation of the peptides. SP-B is thought to reside near the headgroup region with its αhelix parallel to the interface [203]. However, SP-C is highly hydrophobic and infrared reflection-adsorption spectroscopy has revealed that SP-C adopts a tilted orientation [204] embedded within DPPC monolayers. To test whether

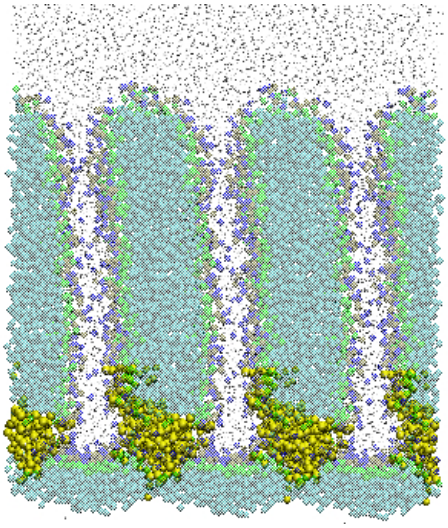
the observed folding mechanism is dependent on the initial peptide orientation, we placed four SP-C molecules within monolayers containing 256 DPPC molecules with an initial SP-C orientation that was tilted and embedded within the tail region (Figure 3.4). Once again the peptides aggregated and the aggregate acted as a site around which fold nucleation occurred. In this case, some of the peptides did not adopt a parallel orientation, but remained embedded within the monolayer. Although the peptide aggregate always acted as a nucleation site, it remained embedded in one monolayer rather than forming a cap on the fold as typically seen for peptides that are oriented parallel to the interface.

3.3.3 Mutants

DPPC monolayers containing 256 lipids per monolayer exhibit fold nucleation about a defect composed of either SP-B_{1-25} or SP-C. To determine the peptide characteristics required for fold nucleation, several SP-B_{1-25} mutants were created. The mutants, listed in Table 3.2, were altered by replacing hydrophobic residues of the original SP-B_{1-25} with hydrophilic ones or vice-versa, and/or by removing the insertion sequence. In each simulation, four mutant peptides were placed in a monolayer composed of 256 DPPC molecules, and simulated under a small negative surface tension. The results indicate that folding is a generic behavior that is not limited solely to the native LS peptides, but is observed for many of the mutant peptides as well. However, for folding to occur via nucleation about a peptide aggregate, the peptides must be sufficiently hydrophobic. It is clear from Table 3.2 that there is a strong correlation between the hydrophobicity of the peptide and folding. If the number of hydrophobic residues is decreased significantly (to less than 10 hydrophobic residues, with the exception of MUT10 that contains nine residues) the monolayer does not fold. If the insertion sequence is removed (MUT15), this decreases the number of hydrophobic residues to 8, and folding does not occur, at least not over the time



(a)



(b)

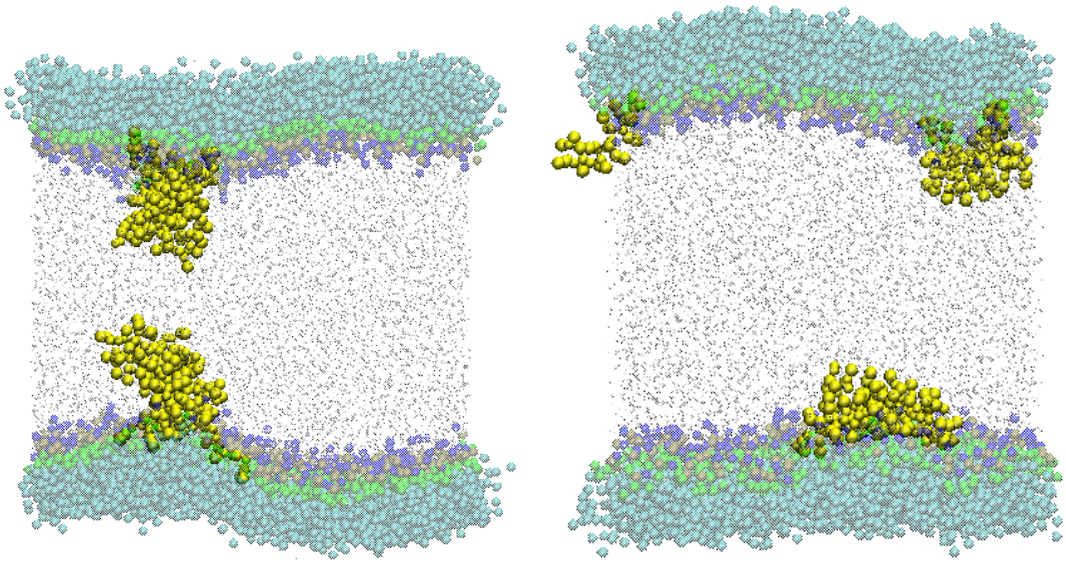
Figure 3.4: A different initial orientation for SP-C. (a) To test the affect of initial orientation, 4 SP-C molecules were placed initially tilted and embedded in the tail region of monolayers containing 256 DPPC molecules each. Under small negative surface tension both monolayers fold. (b) After 500ns the bicelle folds formed from the bottom monolayer are anchored to the monolayer by the peptide aggregate, which remains partially embedded in the monolayer interface. Periodic images are shown for clarity.

scale of our simulation. The results suggest that folding is a generic behavior that is not limited solely to the native LS peptides. However, there are likely overlapping, but distinctive, physiological roles for each hydrophobic surfactant protein (SP-B and SP-C) at the interface, which the present simulations are unable to detect.

Table 3.2: Peptides used in folding simulations. Residues with a polar or charged side chain bead are shown in red, those with non-polar (or apolar) side chains in blue, and those without side chain beads in black. The termini have charged backbone beads and are shown in red. *The cysteines on SP-C are palmitoylated.

Peptide	Amino Acid Sequence	Folding	Hydrophobic Side Chains
SP-B_{1-25}	FPIPLPYCWL CRALIKRIQA MIPKG	Yes (60ns)	14
MUT1	FPIPLPFCFL CLALILLILA MIPLG	Yes (40ns)	21
MUT2	FPIPLPYCFL CLALILLILA MIPLG	Yes (60ns)	20
MUT3	FPIPLPFCFL CLALILLILA MIPKG	Yes (30ns)	20
MUT4	FPIPLPYCWL CRALILLILA MIPLG	Yes (100ns)	19
MUT5	FPIPLPYCFL CLALILLILA MIPKG	Yes (50ns)	19
MUT6	FPIPLPFCFL CRALILLIQA MIPLG	Yes (60ns)	19
MUT7	FPIPLPFCFL CRALIKRILA MIPKG	Yes (80ns)	17
MUT8	FPIPLPFCFL CLALIKRIQA MIPKG	Yes (60ns)	17
MUT9	FPIPLPYQWL CRAQQKRQQA MIPKG	Yes (100ns)	10
MUT10	FPIPLPYQWL QRAQQKRQQA MIPKG	Yes (100ns)	9
MUT11	FPIPLPYQWQ CRAQQKRQQA MIPKG	No	9
MUT12	FL CRALIKRILA MIPKG	No	9
MUT13	WL CRALIKRILA MIPKG	No	9
MUT14	FPIPLPYQWQ QRAQQKRQQA MIPKG	No	8
MUT15	WL CRALIKRIQA MIPKG	No	8
MUT16	FL CRALIKRIQA MIPKG	No	8
MUT17	FPIPLPYQWQ QRAQQKRQQA QQQKG	No	5
SP-C*	FGIPCCPVHL KRLLIVVVVV VLIVV VIVGALLMGL	Yes (90ns)	26

Figure 3.5 illustrates two cases in which folding does not occur. The formation of a peptide aggregate is observed in all cases. In the case of MUT11 (Figure 3.5(a)) the insertion sequence (the last eight residues on the N-terminus) is present but there are only nine hydrophobic residues. Because there are few hydrophobic residues outside of the insertion sequence, interactions between the hydrophilic (polar or charged) residues and the water subphase are dominant. The MUT11 aggregates are anchored to the monolayer by the insertion sequence, but reside mostly in the water subphase and perturb the monolayer very little. In the case of MUT16 (Figure 3.5(b)) there are only eight hydrophobic residues and no insertion sequence. The hydrophobic residues are distributed throughout the peptide; however the total number of hydrophobic residues is still small. The aggregates of MUT16 have a clear perturbing effect on the monolayer, which is evident within 170ns of simulation time; however, folding does not ensue within the 500ns of the simulation.



(a) (b)

Figure 3.5: Two lipid-peptide simulations where folding does not occur within 500ns of simulation under small negative surface tension. Monolayers containing (a) MUT11 and (b) MUT16 are shown at 500ns.

3.3.4 Order Parameters

The carbon-deuterium order parameter (-S_{CD}) is defined by Equation 3.2 [205], where the z-axis is normal to the monolayer, and the order with respect to the x-direction (S_x) is defined according to Equation 3.3 [70]. In Equation 3.3, θ_x is defined as the angle between the lipid tail and the x-axis, and the brackets represent an average over time and molecules. The order with respect to the y-direction (S_y) is similarly defined by replacing θ_x with θ_y.

$$S_{CD} = \frac{2}{3}S_x + \frac{1}{3}S_y \tag{3.2}$$

$$S_x = \frac{3}{2}\left\langle cos^2\theta_x \right\rangle - \frac{1}{2} \tag{3.3}$$

The carbon-deuterium order parameter (-S_{CD}) indicates the order of the lipid tails, which correlates strongly with monolayer fluidity. A value of 0.5 corresponds to a tail that is perfectly aligned parallel to the monolayer normal and a value of 0 corresponds to an isotropic (random) orientation.

To compare lipid tail order for the pure lipid and lipid-protein monolayers the average order parameter (-S_{CD}) of the lipid C2 tail beads is obtained between 10 to 20ns of simulation time, for each system. This relatively short time window is chosen to allow the system to relax from its starting state, but to avoid encompassing the change in order parameter at later times associated with folding itself. Figure 3.6 shows the order parameter versus time for three representative systems. In mixtures that do not fold (thick solid line), the order parameter equilibrates within the first 10ns and remains relatively flat for the entire simulation. Mixtures that do fold exhibit a steep drop in the order parameter (-S_{CD}) resulting from the folding process. As the fold grows, the number of lipids oriented parallel to the monolayer interface (positive S_{CD}) increases and the number of lipids oriented perpendicular to the monolayer interface (negative S_{CD}) simultaneously decreases. Because the time for initiation of

folding varies, some mixtures experience this steep drop in order parameter very early (thin solid line), while others dont drop steeply until more than 100ns have passed (dotted line). Even before this steep decline, the order parameter often declines gradually as undulations form and increase in magnitude. Note that the fluidization of the monolayer results in increased undulations and a decrease in order parameter (which would include undulatory effects). Furthermore, factors that promote larger undulations will favor folding.

In Figure 3.7, the average order parameter (-S_{CD}) of the lipid C2 tail beads is plotted versus the number of hydrophobic residues per peptide present for all of the simulations containing 256 lipids (squares) reported in Table 3.1 and the mutant-containing monolayers (circles) reported in Table 3.2. From this figure, a transitional order parameter -S_{CD}* $\sim$0.331 can be identified above which the monolayer resists folding and below which all monolayers fold. There is one exception; 5:1 DPPC:PA monolayers containing SP-C do not fold and yet for this monolayer -S_{CD} $<$ -S_{CD}*. However, a plot of order parameter versus time for this system (not shown) reveals that the order parameter increases gradually with time and when averaged over the last 10ns of the simulation (triangle) gives -S_{CD} $>$ -S_{CD}*. The simulations containing mutant peptides (circles) display a decrease in order parameter with increasing peptide hydrophobicity, in support of the hypothesis that a decrease in the number of hydrophobic residues leads to less perturbation of the monolayer and eventually the elimination of folding.

In Figure 3.8, the data from Figure 3.7 is reproduced with each data point labeled according to the lipid components of the corresponding simulation and with the data points corresponding to the mutant simulations removed for clarity. The order parameter for the pure DPPC monolayers decreases on addition of SP-B_{1-25} (14 hydrophobic side chains) and decreases even more on addition of SP-C (26 hydrophobic side chains) instead. The order parameter also drops on addition of POPG for all

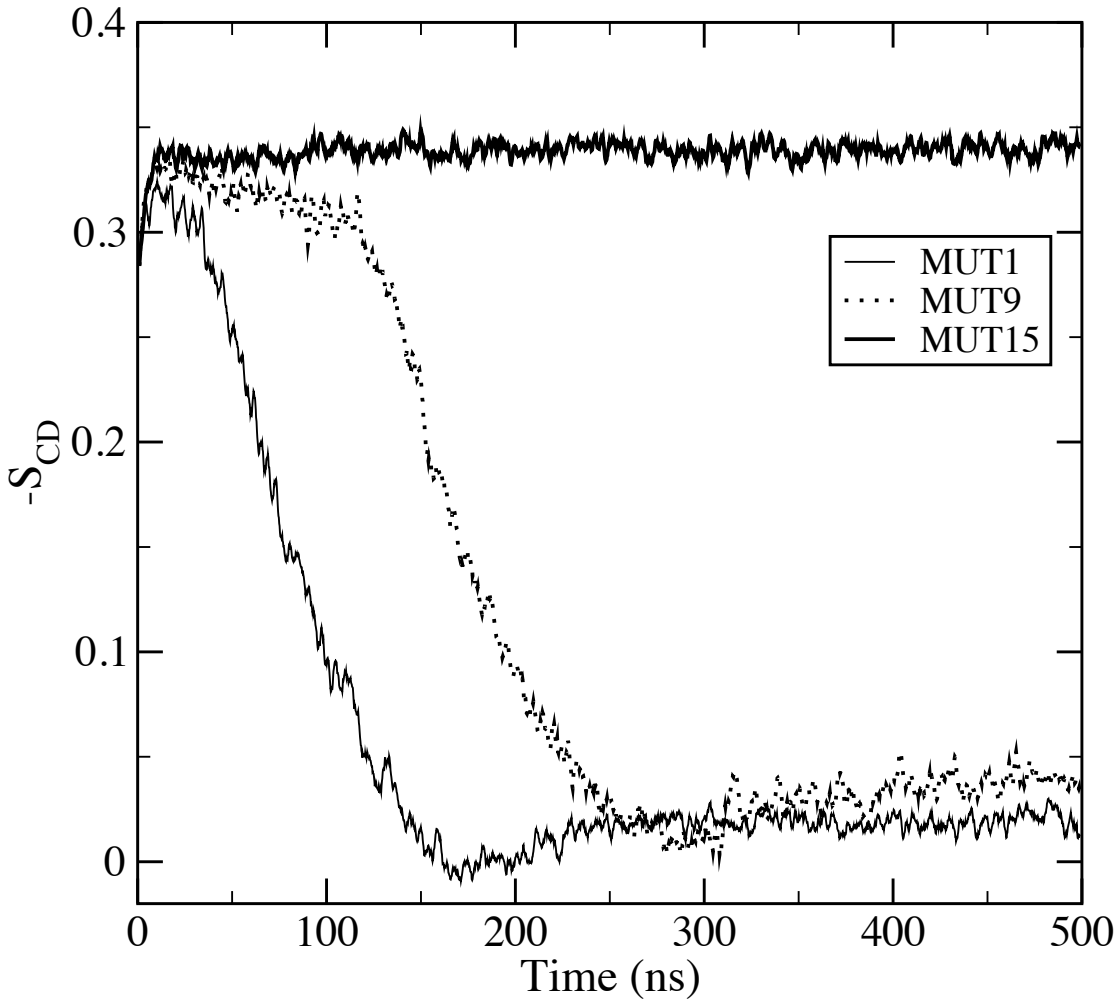


Figure 3.6: Order parameter versus time. Carbon-deuterium order parameter (-S_{CD}) versus time for three representative systems. In mixtures that do not fold (thick solid line), -S_{CD} equilibrates within the first 10ns and remains relatively flat for the entire simulation. Mixtures that fold exhibit a steep drop in the order parameter either very quickly (thin solid line), or more slowly (dotted line). The data are running averages over 10 data points with data points taken every 0.2ns.

pure lipid and lipid-peptide monolayers. The addition of PA leads to an increase in order parameter for DPPC monolayers containing SP-C, which is lessened if the SP-C is depalmitoylated, and to a decrease in order parameter for monolayers containing SP-B_{1-25}. In contrast, if PA is neutralized, a substantial drop in order parameter is observed both in monolayers containing SP-B_{1-25} and in those containing SP-C.

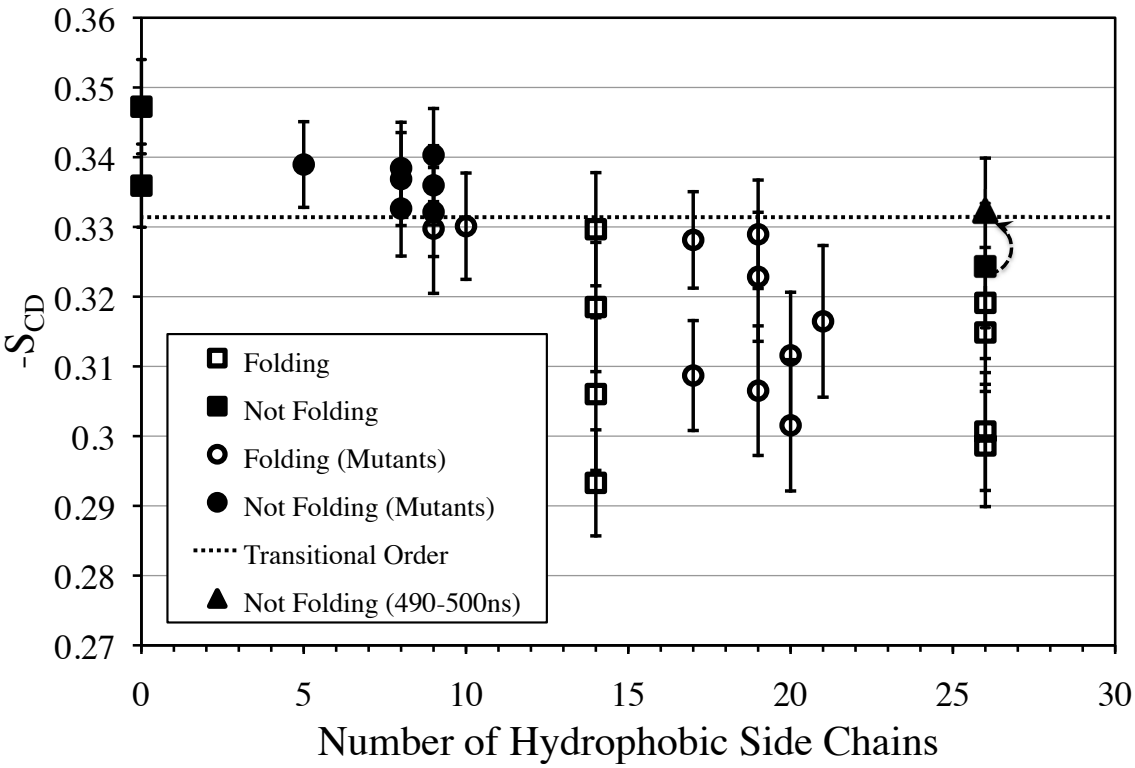


Figure 3.7: The carbon-deuterium order parameters (-S_{CD}) versus the number of hydrophobic residues per peptides present. (a) The carbon-deuterium order parameter (-S_{CD}) of the lipid C2 tail beads for pure lipid and lipid-peptide monolayers that do (open symbols) and do not (closed symbols) fold. Simulations containing mutants (circles) are distinguished from the other mixtures (squares) for clarity. The data are averages over the time from 10 to 20ns. The error bars are standard deviations in the average order parameter with respect to time. From these results, the transitional order parameter (dotted line) above which the monolayer resists folding is identified. For DPPC/PA monolayers containing SP-C the order parameter shifts (arrow) above the transitional order parameter at later times (triangle).

In addition to these order parameters, which are averaged for all of the C2 tail

sites in the mixture, the average order parameters of the C2 tails sites of each individual component were also calculated (not shown). In mixtures of DPPC and POPG, the order parameters for POPG are slightly lower than those for DPPC. This is to be expected because POPG contains an unsaturated chain (and a charged headgroup). In mixtures containing DPPC and PA, the PA tails exhibit a slightly higher order parameter than do the DPPC molecules. However, a substantial drop in order parameter is observed for PA if it is neutralized.

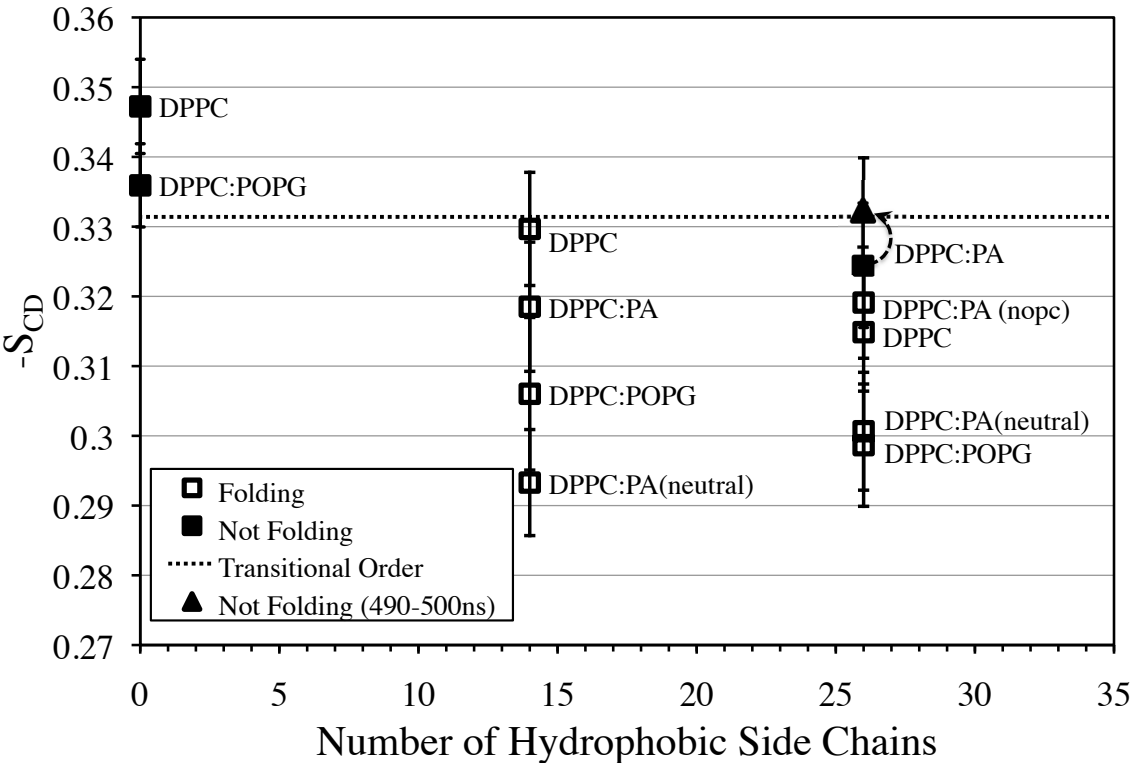


Figure 3.8: The carbon-deuterium order parameters (-S_{CD}) are reported for various lipid and lipid peptide mixtures. This figure is reproduced from Figure 3.7, with each data point labeled with the lipid components of the corresponding simulation and with the data points corresponding to the mutant simulations removed for clarity. The label "nopc" signifies that the palmitoyl chains have been removed from SP-C, and the label "neutral" signifies that PA has been neutralized.

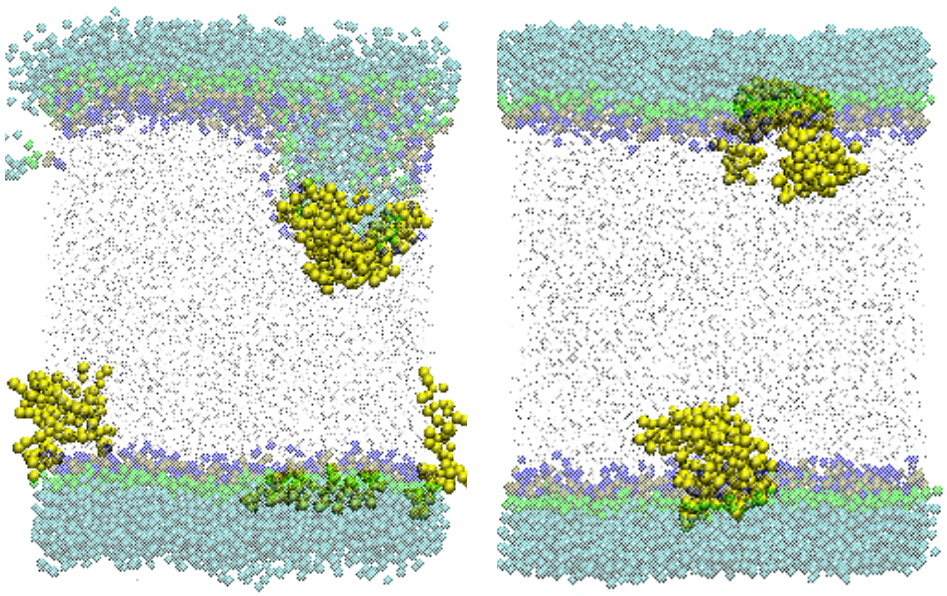
The larger systems, containing 1024 and 2304 lipids per monolayer, also show folding when their order parameters are below the same transitional value. For instance,

256 lipid 1:1 DPPC:DPPG monolayers do not fold; however once the monolayer size is increased to 1024 lipids folding occurs. This increase in system size is accompanied by a shift in order parameter from above to below the transitional order parameter. When 256-lipid monolayers of pure DPPC are increased in size by factors of four and nine, no folding is observed. In these cases, although the order parameter shifts slightly downward with increasing system size, it remains well above the transitional order parameter.

Both POPG and the surfactant peptides SP-B and SP-C fluidize the monolayer. However, unlike POPG, these peptides are not distributed uniformly throughout the monolayer and therefore the order parameter can be resolved locally. For monolayers containing 256 DPPC and 4 SP-C molecules per monolayer, the distance between the C2A beads of DPPC and the closest SP-C bead was calculated for each tail at 10ns. The lipids were then binned into two categories based on their distance from a peptide and the average order parameter for the C2A tail sites was calculated for each category. Lipids with C2A beads $\leq$1nm away are considered "neighboring" lipids, and those with C2A sites >1nm away are considered "distant" lipids. The order parameter for the "neighboring" lipids (0.284 $\pm$ 0.029) is smaller than the average order parameter over all C2A beads for the entire monolayer between 10-20ns (0.321). In contrast, "distant" lipids have an average order parameter (0.345 $\pm$ 0.008) that is essentially the same as the average order parameter, for the C2A tail sites, obtained for the pure DPPC monolayer (0.350). The standard errors given here were approximated by using a binwidth of 0.10nm to group molecules into multiple data points within each category (neighboring and distant). The lipid tail order of DPPC monolayers containing SP-B_{1-25} displays a similar distance dependence. These results suggest that peptides increase the monolayer fluidity locally.

3.3.5 Temperature and Surface Tension

The folding exhibited at 323K for pure lipid and lipid-protein mixtures occurs in LE phase monolayers. As the temperature is decreased the monolayers undergo a phase transition from LE to LC phase that affects the folding transition. In Figure 3.9, monolayers containing 256 DPPC and 4 SP-B_{1-25} molecules at 308K (Figure 3.9(a)) and 303K (Figure 3.9(b)), are shown after 500ns of simulation under a small negative surface tension. The transition from LE to LC occurs in these monolayers at around 308K. At 313K the monolayers fold, just as seen at 323K. When the temperature is reduced to 308K and 303K peptide aggregation still occurs; however changes in folding behavior are evident. As shown in Figure 3.9(a), at 308K folding occurs in only one monolayer (folding always occurs concurrently in both monolayers at 323K), resulting in shrinkage of the lateral dimensions of the simulation box, thereby forcing the other monolayer into the LC phase. At 303K, both monolayers condense into the LC phase (Figure 3.9(b)). At both 303K and 308K, the peptides in the LC phase monolayers are pushed further into the water subphase and no folding occurs within 500ns of simulation time. As the system size is increased, the monolayers are able to fold at lower temperatures. For the largest system size, DPPC monolayers containing SP-B_{1-25} fold at both 308K and 303K. At 303K coexistence between LC and LE phase is evident in the largest system; however as the undulations grow the monolayers disorder before folding. To test whether folds and LC-LE coexistence could be observed simultaneously, temperature was further lowered to 298K. At 298K LC-LE coexistence was evident; however the monolayer did not fold within the simulation time, and a few of the SP-B_{1-25} molecules were forced completely out of the monolayer and into the subphase. The collapse of LC phase DPPC monolayers has been previously simulated at temperature of 300K, using the MARTINI model [31]. This collapse occurred by formation a small buckling deformation and required a larger negative surface tension and system size than utilized in our simulations.



(a) (b)

Figure 3.9: Decreasing the temperature affects the occurrence of folding. DPPC monolayers containing 256 DPPC and 4 SP-B_{1-25} molecules each are shown at 308K (a) and 303K (b), after 500ns under small negative surface tension.

When the surface tension is increased, the monolayers gain stability against collapse. Larger systems (1024 or 2304 lipids) of DPPC mixed with POPG, SP-B_{1-25}, or SP-C with a small negative, zero, and small positive surface tensions always yield order parameters that are below the transitional order parameter identified above. However folding is not always observed for these simulations with zero or small positive surface tensions within the simulation time, because of increased stability against collapse. Folding would likely occur if these simulations were run for a long time.

As shown in Table 3.1, at the largest system size (2304 lipids per monolayer) both 1:1 DPPC:POPG and DPPC monolayers containing 36 SP-B_{1-25} molecules fold under a small negative surface tension. If the surface tension is set to zero, folding occurs readily in the SP-B_{1-25} -containing monolayer (within 100ns), with a group of nearby peptides (not a well defined aggregate) acting as a nucleation site for fold formation. However, 1:1 DPPC:POPG monolayers exhibit small undulations but