Using Molecular Dynamics Simulations and Statistical Modeling to Explore Membrane Structure and Membrane-Protein Interactions

Shushan He

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2018

Reading Committee:

Lutz Maibaum, Chair

Sarah L. Keller

David Masiello

Program Authorized to Offer Degree:

Department of Chemistry

University of Washington

Abstract

Using Molecular Dynamics Simulations and Statistical Modeling to Explore Membrane Structure and Membrane-Protein Interactions

Shushan He

Chair of the Supervisory Committee:

Assistant Professor Lutz Maibaum

Department of Chemistry

The structure of cellular membranes gives rise to important biological phenomena, and understanding the (de)mixing behavior of multicomponent lipid bilayers is an important step toward unraveling the nature of spatial composition heterogeneities in cellular membranes and their role in biological function. In this dissertation we focus on the spatial organization and compositional heterogeneity of model membrane systems and their interaction with small, biologically relevant peptides. I employed both coarse-grained and atomistic molecular dynamics simulations to study the composition phase diagram of a quaternary mixture of phospholipids and cholesterol, as well as to sample the configurational space of peptide-membrane interactions. I investigate the mechanisms controlling membrane spatial heterogeneity and membrane protein interaction. This work yields important new insight into both the structural properties of lipid bilayer systems with spatial and compositional heterogeneity at vastly different length scales, which has prompted numerous publications in the field seeking for a plausible mechanism. This work will also provide perspectives on configurations of the $PAP_{248-286}$ peptide upon interacting with membranes, which, despite

its importance for human health, has not received as much attention from the research community as other amyloid-forming peptides. In this wide-open field such simulations will have significant scientific impact.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to thank Professor Lutz Maibaum for the opportunity to work on this research and for his input and guidance throughout my time in his research group. I would also like to thank the other members of the Maibaum group for their support. Within the Chemistry Department I have had the support of the members of the Amphiphiliphiles and Theory Suite groups. I would like to thank the members of both for the opportunity to receive input from different audiences and for expanding my knowledge within their fields. This work was facilitated through the use of advanced computational, storage, and networking infrastructure provided by the Hyak supercomputer system at the University of Washington. Parts of this work were performed using the Ilahie cluster, which was purchased using funds from a MRI grant from the National Science Foundation (CHE-1624430).

Chapter 1

INTRODUCTION

1.1 Membrane Structure and Spatial Heterogeneity

The plasma membrane is the barrier around the cell that separates its internal environment from the surroundings. The spatial organization of cellular membrane plays an important role in organizing important cellular components, as well as mediating the many processes that are vital for cellular organisms. The complex nature of the membrane system makes it a long-standing and challenging topic to study,[3,18,55,57,98–100,106] and the physical mechanism that gives rise to the many interesting phenomena, including spatial heterogeneity, as well as its role in biological function remain unclear.[18,79,97]

The enormous complexity of membrane structure of living cells requires that we start by understanding model lipid bilayer systems with controlled lipid compositions. Phospholipids, as amphipathic molecules and major component of biological membranes, are believed to contribute to composition heterogeneity through their complex phase behavior.[38,113] It is well-known that under some conditions model membrane systems can exhibit spatial heterogeneity at the micrometer length scale.[19,48,107,111–113,119] With a molecular dynamics (MD) simulation approach, as well as the continuum model we developed to help capture and understand the key features of spatially heterogeneous membrane systems, this work will focus on quantifying the onset and compositions of such phase separated lipid bilayer systems as a driving force of membrane heterogeneity on a four-component lipid mixture.

It has been noted that in some biologically relevant membrane systems, including the cellular plasma membranes, no such large domains have been observed with conventional microscopy.

However, results with experimental techniques on cell-derived Giant plasma membrane vesicles (GPMVs) have shown the existence of liquid-order-like and liquid-disorder-like domains at smaller length scales than those observed on model membrane systems,[9,110] and that such domains contribute to cell functions such as protein sorting.[19] Moreover, in model GUVs where membrane tensions are carefully controlled by tuning osmotic pressure, heterogeneity with smaller length scale and stripe-like features have been observed,[23] and a model based on competition between domain line tension and membrane curvature has been proposed.[2] As a collaborated effort with the Keller lab (UW Department of Chemistry) and the Levental lab (University of Texas Health Science Center at Houston Department of Integrative Biology and Pharmacology), this work will focus on interpreting and understanding the structural features of such model membranes systems with small domains, as well as elucidating the mechanisms controlling their length scales.

1.2 Structural Changes and Co-aggregation of PAP248-286 on Membranes

It is well known that peptides can undergo large structural changes when binding to surfaces. For example, when exposed to rigid, hydrophobic surfaces, peptides containing periodic leucine and lysine sequences adopt a helical conformation.[12,58,116] Cellular and model membranes also form surfaces for proteins and peptides to bind to. Their intrinsic heterogeneity, flexibility, and dynamics can lead to complicated binding mechanisms and new, emergent properties of the combined membrane/peptide system. For example, we have recently shown that the suppression of membrane undulations can result in an entropic, membrane-mediated attraction between proteins.[88]

PAP248-286, a 39-residue fragment of prostatic acid phosphatase, is an abundant seminal peptide that has been proposed to play a variety of roles, both beneficial and deleterious to human health. PAP is intrinsically disordered in solution and α-helical when bound to membranes and can also form β-sheet-rich amyloid fibrils. In its amyloid form, referred to as Semen-derived Enhancer of Viral Infection (SEVI), PAP248-286 potently enhances infection

by enveloped viral pathogens including HIV, herpes simplex virus and cytomegalovirus.[39,104] On the other hand, there is substantial evidence that PAP248-286 displays antimicrobial activity and participates in sperm quality control.[67] Note that while PAP248-286 effectively induces membrane leakage in vitro, its biologically relevant antimicrobial activity appears to derive from its ability to cause bacterial agglutination instead of direct cytotoxicity.[21,40] Our collaborator Dr. Abhinav Nath (UW Department of Medicinal Chemistry) and coworkers have also observed that PAP248-286 can inhibit biofilm formation by bacterial pathogens, an important virulence mechanism. A common feature of PAP's disparate activities is that they depend on this peptide's ability to bind biological membranes so as to alter cell-cell interactions.

PAP248-286 and similar positively-charged, amyloid-forming or antimicrobial peptides can interact with biological membranes in diverse and complex ways. Peptides can cause vesicle leakage or fusion, or induce changes in membrane curvature. Self-assembly is another important aspect of these interactions: some peptides can induce the aggregation or agglutination of lipid vesicles, while conversely membranes can induce amyloid fibril formation. Importantly, there is a growing body of evidence that self-assembly phenomena, such as the encapsulation of target cells or pathogens, can be more relevant to biological function than classical modes of action such as membrane lysis.[45,52,66] In this work we focus on the rich, experimentally observed structural changes that occur when $PAP_{248-286}$ binds to phospholipid bilayers.

1.3 Specific Aims

The work in this thesis is divided into four main components. The first focuses on a continuum model we employed to help us understand the structure and signature properties of two-dimensional phase separated systems with different geometric features. We then tested and highlighted some key structural properties using this model to illustrate how connections between different degree of membrane heterogeneity and signature properties can be built.

The second is a discussion of determining the onset and composition of phase separated model membrane systems using coarse-grained MD simulations, as well as developing a systematic way to detect such large scale heterogeneity in membrane simulations. The third focuses on interpreting and understanding the structural features of model membranes systems where large domains are not observed. The fourth part focuses on the interaction between the PAP248-286 peptide and phospholipid bilayers using atomistic simulations in an attempt to understand the conditions and mechanism for aggregate formation of the peptide.

1.3.1 Key Structural Features and Continuum Model of Heterogeneous Systems

In this part of the work we identified key structural features that can be calculated in order to illustrate the degree of spatial heterogeneity on a two-dimensional bilayer system. We then developed a "toy model" approach to understand connection between these structural features and the spatial heterogeneity on bilayer systems, in order to interpret our MD simulations on such systems. For these continuum systems we first assumed the existence of two distinct domains with uniform but different order parameters and stable, non-fluctuating domain boundaries. We then calculated various structural and compositional properties of these models and built the connection between the models parameters and measurable quantities in our simulations. Then, in a more realistic approach, we allowed the order parameters to fluctuate within the domains around some average value and re-evaluated the structural and compositional profiles of the model.

1.3.2 Determining the Onset of Phase Separation on Coarse-Grained Membrane Simulations

We used coarse-grained molecular dynamics simulations to study the composition phase diagram of a quaternary mixture of phospholipids and cholesterol. Employing the models described in the previous section, we compared and combined different statistical measures

of membrane structure to identify the onset of phase coexistence in composition space. Using the dependence of composition heterogeneities on the size of the system, along with information obtained from observation windows of different sizes, we applied statistical methods such as Gaussian clusters and developed a systematic approach to accurately distinguish phase coexistence from structured homogeneous phases.

1.3.3 Interpreting the Structural Signals of Membranes with Smaller Domains

In a collaborative effort with the Keller and Levental labs, we provided calculations and analysis on fluorescence image of GUVs in order to extract structural information from the model system, and to characterize the various length scale of the morphology that was previously identified as "modulated phase" or "nano-domains". We then evaluated characteristic length scales and fluorescence levels in these systems to test proposed mechanisms of modulated phases and microemulsions.

1.3.4 Structural Changes and Co-aggregation of PAP248-286 upon Binding to Membranes

As a collaboration with the Nath Lab, we discovered that both $PAP_{248-286}$ and the lipids show significant reorganization that depends both on the peptide:lipid ratio and on the composition of the bilayer. To explore the large ensemble of disordered conformations of the aggregate peptide/bilayer system, we used long time scale atomistic Molecular Dynamics simulations. We captured the molecular detail of the peptide membrane interaction in an attempt to investigate the binding mechanism, as well as the conditions and implications of aggregate formation. We first confirmed that the CHARMM36m[37] out-performs other popular force fields for atomistic $PAP_{248-286}$ simulation in that the peptide does not over-express α-helical structure in solution. We studied the binding between $PAP_{248-286}$ and the membrane, and its effect on the secondary structure of the peptide. For that purpose we simulated the peptide on top of lipid bilayers with various POPG compositions. We

found that both binding behavior and secondary structure of SEVI depend on the membrane composition of negatively charged POPG lipids.

Chapter 2

KEY STRUCTURAL FEATURES AND CONTINUUM MODEL OF SPATIALLY HETEROGENEOUS MEMBRANE SYSTEMS

In this work we identified and discussed key structural features one can employ to illustrate the degree of spatial heterogeneity on a two-dimensional bilayer system. We then took a model facilitated approach to understand connection between these structural features and the spatial heterogeneity on bilayer systems, in order to interpret our MD simulations on such systems. For these continuum systems we first assumed the existence of two distinct domains with uniform but different order parameters and stable, non-fluctuating domain boundaries. We then calculated various structural and compositional properties of these models and built the connection between the models parameters and measurable quantities in our simulations. Then, in a more realistic approach, we allowed the order parameters to fluctuate within the domains around some average value and re-evaluated the structural and compositional profiles of the model.

2.1 Density Radial Distribution Function

The density radial distribution function is a measure for density correlations between two points in space that are separated by a distance vector $\boldsymbol{r}$. On a two-dimensional system where we treat the density of the molecule of interest, α, as a continuous function in two-dimensional space, the density radial distribution function of α, or $g_\alpha(\boldsymbol{r})$, can be defined as:[27]

$$g_\alpha(\boldsymbol{r}) = \frac{\rho_\alpha^{(2)}(\boldsymbol{r})}{\rho_\alpha^{(1)}(\boldsymbol{r})^2} \tag{2.1}$$

where $\rho_\alpha^{(n)}$ is the partial equilibrium n-particle density, for particle α and:[27]

$$\rho_\alpha^{(n)} = \frac{N_\alpha!}{(N_\alpha - n)!} \frac{1}{\mathcal{Z}_{N_\alpha}} \int \mathrm{e}^{-\beta V_{N_\alpha}} \mathrm{d}\boldsymbol{r}^{(N_\alpha - n)} \tag{2.2}$$

where N_α is the total number of α moleulces in the system, Z_{N_α} is the configurational integral, and V_{N_α} is the potential part of the Hamiltonian of the system.

Similarly, for a system whose density can be described as a collection of discrete particles, we can define the partial molecular density in a slightly different way and adjust the expression accordingly. Cross correlation can also be calculated for particles of different types. More details of this are shown in Chapter 3 of this work.

2.2 Structure Factor

The structure factor provides information about the spatial organization of the membrane, and can be obtained from neutron or X-ray scattering experiments. Here we assume that each molecule can be considered as a point scatterer at position $\boldsymbol{r}_i$. While this is a gross approximation, it is justified because we are only interested in large-scale structure where detailed information from single molecules becomes negligible. We define the partial structure factor as[27]

$$S_{\alpha\beta}(\boldsymbol{k}) = \left\langle \frac{1}{N} \tilde{\rho}_\alpha(\boldsymbol{k}) \tilde{\rho}_\beta(-\boldsymbol{k}) \right\rangle \tag{2.3}$$

where

$$\tilde{\rho}_\alpha(\boldsymbol{k}) = \int \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{i\boldsymbol{k}\boldsymbol{r}} \rho_\alpha(\boldsymbol{r}) \tag{2.4}$$

is the Fourier transform of the one particle partial particle density function from eq (2.2). The structure factor and radial distribution function are related by [27]

$$S_{\alpha\beta}(\boldsymbol{k}) = x_\alpha(\delta_{\alpha\beta} + x_\beta)\rho \int g_{\alpha\beta}(r) \mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \mathrm{d}\boldsymbol{r} \tag{2.5}$$

where x_α and x_β are the mole fractions of particle species α and β, respectively, and $\delta_{\alpha\beta}$ is the Kronecker delta.

2.3 Local Density Distribution (LDD)

For a particle α, we define the local density distribution (LDD) as the probability distribution function of the instantaneous density of particle species α in a region with linear dimension w

$$P(\rho_\alpha) = \frac{1}{A}\int \mathrm{d}\boldsymbol{r}\,\langle\delta\left(\hat{\rho}_\alpha(\boldsymbol{r}) - \rho_\alpha\right)\rangle \tag{2.6}$$

where

$$\hat{\rho}_\alpha(\boldsymbol{r}) = \frac{1}{w^2}\int_{(w(\boldsymbol{r}),w(\boldsymbol{r}))} \mathrm{d}\boldsymbol{r}'\rho_\alpha\left(\boldsymbol{r}'\right) \tag{2.7}$$

is the local partial particle density averaged over a square observation window with side length w positioned at $\boldsymbol{r}$ on the membrane. A is the total area of the system. In practice, for a membrane system, (2.6) is calculated by sweeping the observation window laterally over each leaflet in every simulation frame.

Later in this chapter we apply the LDD calculation on the continuum model and show that the characteristics of the LDD profile depend both on the degree of segregation and on the observation window size, w. We provide detailed analysis of this dependence in the Results section in Chapter 3.

2.4 Continuum Membrane Models for Spatially Heterogeneous Systems

In order to interpret simulation data and obtain structural features from phase separated bilayer systems, we developed continuum phase models for heterogeneous systems. The geometry of the pattern can be stripy or circular as observed in the MD simulations, depending on the are fraction of the two phases. These models feature domains with uniform order parameters (*e.g.* phase density) within the domains and stable, non-fluctuating domain boundaries with different geometric shapes. It is noted that the model itself does not answer the questions about the nature of spatial heterogeneities on bilayer systems, nor will it alone determine whether a bilayer with a certain domain shape has undergone phase separation.

However this "toy model" approach will allow us to capture the key structural features of spatially heterogeneous systems. These features can later be used on heterogeneous membrane systems at various length scales as part of a scheme to systematically determine the nature of membrane heterogeneity based on physical principles.

2.4.1 Stripe Model

For a spatially heterogeneous membrane, a striped geometry occurs when the two coexisting phases have similar area fractions. We propose a simple two-phase model that can be mathematically defined as a two-dimensional density field for species α in phase A and B:

$$\rho_\alpha(\boldsymbol{r}) = \begin{cases} \rho_A & \text{if } 0 \le x < a \\ \rho_B & \text{if } a \le x \le L \end{cases} \tag{2.8}$$

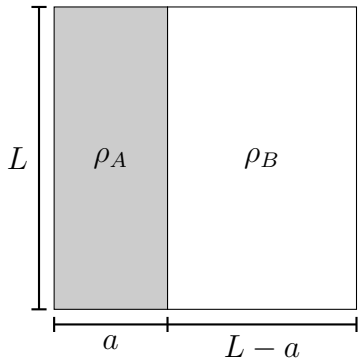

Figure 2.1: For a continuum model with a straight line boundary, ρ_A and ρ_B are the partial lipid densities of the two phases, respectively, on a $L \times L$ square 2-D membrane, with phase A being the smaller phase with dimensions $L \times a$, while phase B is the larger phase with dimensions $L \times (L - a)$. We assume that $a < \frac{L}{2}$.

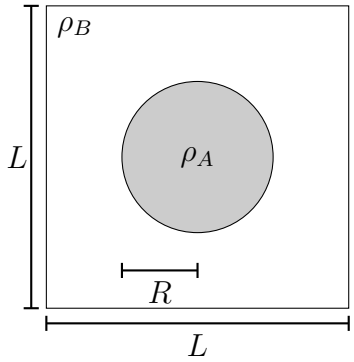

Figure 2.2: For a continuum model with a circular phase boundary the domain radius R relative to membrane size L is varied. When the domain is small and when domains are placed on the membrane randomly, systems with different number of circular domains show identical averaged numerically calculated structural properties (*e.g.* $g(r)$) over large number of iterations.

2.4.2 Disc Model

Circular domains/clusters were predicted and observed in simulations when one of the phases have a relatively small area fraction compared to the other phase. It can be easily shown that a circular domain is preferred for radius $R < \frac{L}{\pi}$, assuming a constant positive line tension between the phases. The two dimensional density profile can be characterized as:

$$\rho_\alpha(\boldsymbol{r}) = \begin{cases} \rho_A & \text{if } 0 \le r < R \text{ and } 0 \le \theta \le 2\pi \\ \rho_B & \text{if } R \le r \le \min\left\{\left|\frac{L}{2\cos\theta}\right|, \left|\frac{L}{2\sin\theta}\right|\right\} \text{ and } 0 \le \theta \le 2\pi \end{cases} \tag{2.9}$$

2.4.3 Discretized Model to Capture Density Fluctuation within Domains

We make variations of the aforementioned two continuum membrane models by implementing the membrane density not as a well-behaved, continuous function, but as a collection of delta

functions:

$$\rho_\alpha(\boldsymbol{r}) = \sum_{i=1}^{N_\alpha} \delta(\boldsymbol{r} - \boldsymbol{r}_i) \tag{2.10}$$

This discretized version of the membrane model will allow fluctuations within domains, yielding more comparable structural properties to those obtained from MD simulations. For a stripe model system, the discrete density field can be defined using a binomial distribution where the average density within each of the two stripe phases is a constant:

$$\langle \rho_\alpha(\boldsymbol{r}) \rangle = \begin{cases} \rho_A & \text{if } 0 \le x < a \text{ and } 0 \le y \le L \\ \rho_B & \text{if } a \le x \le L \text{ and } 0 \le y \le L \end{cases} \tag{2.11}$$

Similarly for a disc model we have:

$$\langle \rho_\alpha(\boldsymbol{r}) \rangle = \begin{cases} \rho_A & \text{if } 0 \le r < R \text{ and } 0 \le \theta \le 2\pi \\ \rho_B & \text{if } R \le r \le \min\left\{ \left|\frac{L}{2\cos\theta}\right|, \left|\frac{L}{2\sin\theta}\right| \right\} \text{ and } 0 \le \theta \le 2\pi \| \end{cases} \tag{2.12}$$

2.5 Understanding Observation Window Size Dependence in LDD Calculation Through Stripe Model Analytical LDD

2.5.1 Continuous Stripe Model Analytical LDD

We assume that the observation window is $w \times w$, and that $\rho_A > \rho_B$ (smaller phase is the "rich" phase in species α), then $\Delta\rho = \rho_A - \rho_B > 0$. For the calculation, we also choose an arbitrary value of $a = 0.4L$. Using the definition of the stripe membrane model in (2.8) and the LDD definition in (2.6), we have:

1. $0 < w \le a$:

$$P(\rho_\alpha) = \begin{cases} 0 & \text{if } \rho_\alpha < \rho_B \\ \frac{L-w-a}{L}\delta(\rho_\alpha - \rho_B) & \text{if } \rho_\alpha = \rho_B \\ \frac{2w}{L\,|\Delta\rho|} & \text{if } \rho_B < \rho_\alpha < \rho_A \\ \frac{a-w}{L}\delta(\rho_\alpha - \rho_A) & \text{if } \rho_\alpha = \rho_A \\ 0 & \text{if } \rho_\alpha > \rho_A \end{cases} \tag{2.13}$$

2. $a < w \leq L - a$:

$$P(\rho_\alpha) = \begin{cases} 0 & \text{if } \rho_\alpha < \rho_B \\ \frac{L-w-a}{L}\delta(\rho_\alpha - \rho_B) & \text{if } \rho_\alpha = \rho_B \\ \frac{2w}{L\,|\Delta\rho|} & \text{if } \rho_B < \rho_\alpha < \rho_B + \frac{a}{D}\Delta\rho \\ \frac{w-a}{L}\delta\left(\rho_\alpha - (\rho_B + \frac{a}{D}\Delta\rho)\right) & \text{if } \rho_\alpha = \rho_B + \frac{a}{w}\Delta\rho \\ 0 & \text{if } \rho_\alpha > \rho_B + \frac{a}{w}\Delta\rho \end{cases} \tag{2.14}$$

3. $L - a < w \leq L$:

$$P(\rho_\alpha) = \begin{cases} 0 & \text{if } \rho_\alpha < \rho_{\alpha,A} - \frac{L-a}{w}\Delta\rho \\ \frac{a-L+w}{L}\delta\left(\rho_\alpha - \left(\rho_{\alpha,A} - \frac{L-a}{w}\Delta\rho\right)\right) & \text{if } \rho_\alpha = \rho_{\alpha,A} - \frac{L-a}{w}\Delta\rho \\ \frac{2w}{L\,|\Delta\rho|} & \text{if } \rho_{\alpha,A} - \frac{L-a}{w}\Delta\rho < \rho_\alpha < \rho_{\alpha,B} + \frac{a}{D}\Delta\rho \\ \frac{w-a}{L}\delta\left(\rho_\alpha - (\rho_{\alpha,B} + \frac{a}{w}\Delta\rho)\right) & \text{if } \rho_\alpha = \rho_{\alpha,B} + \frac{a}{w}\Delta\rho \\ 0 & \text{if } \rho_\alpha > \rho_{\alpha,B} + \frac{a}{w}\Delta\rho \end{cases} \tag{2.15}$$

We notice that the bimodal density peak positions are invariant of the observation window size w when $w < a = 0.4L$ (the size of the rich/smaller phase). When w becomes greater than a, the peak corresponding to the rich phase density starts to shift towards the global density. Similarly, the peak corresponding to the poor (larger) phase begins shifting to the

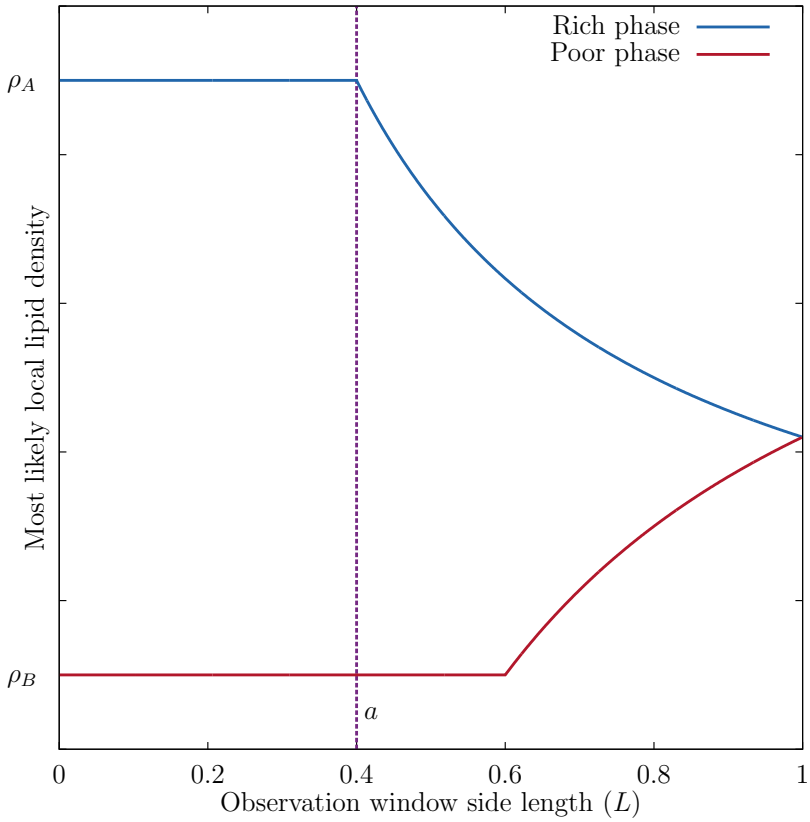

Figure 2.3: Most probable local lipid density derived from the analytical expression shown in Eqn. 2.13 - 2.15 for a system where $a = 0.4L$. The delta peak positions for the rich and poor phases are plotted as a function of the observation window size in terms of box size L.

center when the window size exceeds the length scale of the poor phase, and the two peaks meet at the average global density when w reaches the system size L. We extract the peak positions from the analytical solution (Eqn. 2.13 - 2.15) of LDD profile for the two phases of the continuum model, and the result is shown in Fig. 2.3.

2.5.2 Discrete Stripe Model Numerical LDD

The box is set up using a grid point representation of the discrete density field, where the particle density distribution satisfied the distribution around the average density of the two

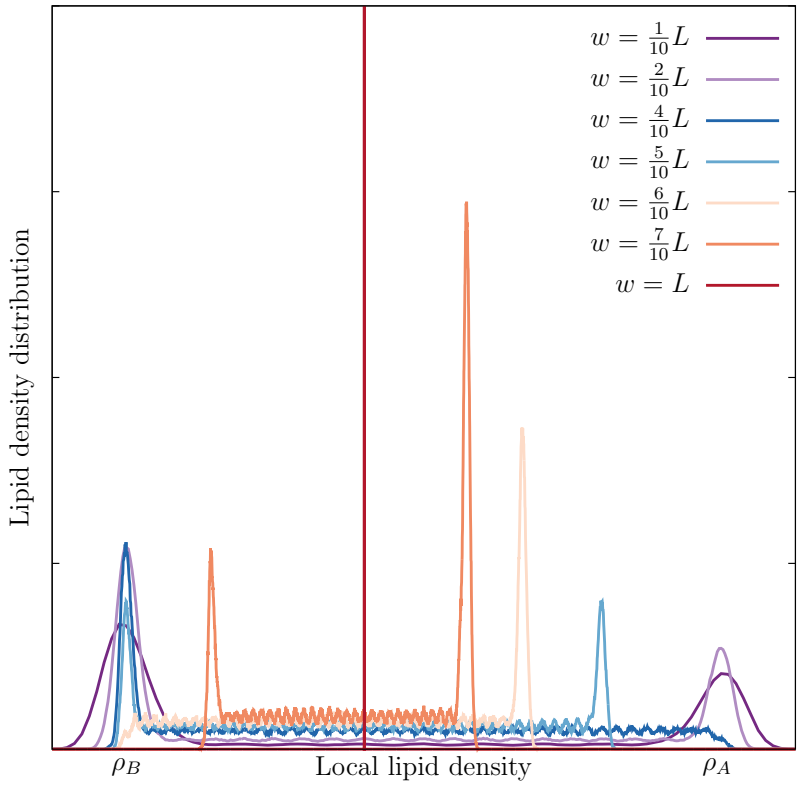

Figure 2.4: Local lipid density distribution calculated for a stripe model system described in Eqn. S6-S7. The discrete field of density distribution satisfies a binomial distribution around the phase average density. The width of the stripe a is set to $\frac{4}{10}L$. The average density inside the smaller stripe is ρ_A, while the average density outside is set to ρ_B. Note that as observation window size w increases and becomes greater than a the peak representing the average density of the smaller stripe (ρ_A) starts shifting towards the center. We observe similar effects on the ρ_B peak as w grows beyond the length scale of the bigger stripe.

individual phases respectively, given in Eqn. (2.11). The result of this calculation is shown in Figure 2.4. Here the width of the smaller stripe (where $\langle \rho_\alpha \rangle = \rho_A$) is $a = 0.4L$, and we observe that for $w > a$ the ρ_A peak starts shifting towards the system average density. Similarly for $w > L - a$ the ρ_B peak also starts shifting, as predicted by the continuum model. We thus confirm that, for the model system with two different phases where each has a binomially distributed density of particles, the peak positions and their movement as the observation window size changes are consistent with the result predicted by the analytical/continuum model shown in Figure 2.3. This model of finite resolution causes the flat region between the two peaks to fluctuate, as discrete "particles" pop in and out of the sampling window.

2.5.3 Discrete Disc Model Numerical LDD

Due to the discrepancy in the geometry between a square observation window and circular domains, analytical solution to the LDD profile of the disc model is not trivial. However similar trend can be observed when we calculate LDD for a disc model with discrete particle density. To confirm that the analytical equations we derived earlier applies to the circular model we calculated the local density histogram using the discrete density field defined in Eqn. (2.12). The box is set up using a grid point representation of the discrete density field, where the particle density distribution satisfied a normal distribution around the average density of the two individual phases respectively. The result of this calculation is shown in Figure 2.5. Here the radius of the circular domain (where $\langle \rho_\alpha \rangle = \rho_A$) is $R = \frac{1}{4}L$, and we observe that for $w > \sqrt{2}R$ (or $w > 0.35L$) the ρ_A peak starts shifting towards the system average density. Note that for the circular modal this threshold w value is no longer simply the width of the domain, but rather $\sqrt{2}R$ due to the discrepancy in geometry between a square observation window and circular domains. Similarly, for $w > L - \sqrt{2}R$ (or $w > 0.65L$) the ρ_B peak also starts shifting, as seen in the stripe model. We thus confirm that, for the disc model system with two different phases where each of the phases has a binomially distributed density of particles, the LDD peaks' positions and movements are once again

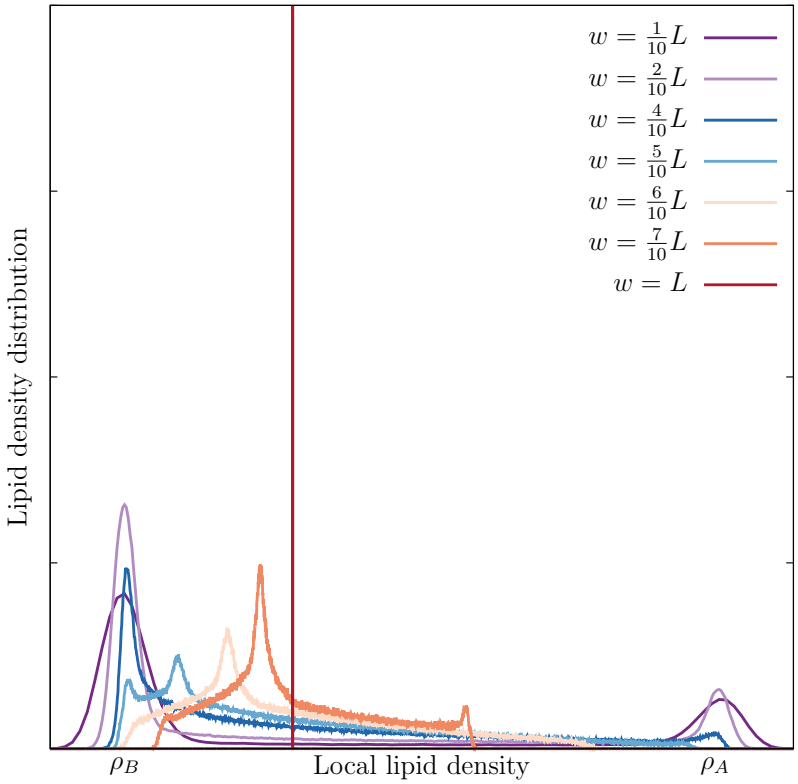

Figure 2.5: Local lipid density distribution calculated for a model system described in Eqn. S8-S9. The discrete field of density distribution satisfies a Gaussian distribution around the phase average density. The radius of the circular domain R is set to $\frac{1}{4}L$. The average density inside the circular domain is ρ_A, while the average density outside is set to ρ_B. Due to the special circular geometry of the domain, we note that as observation window size w increases and becomes greater than $\sqrt{2}R \approx 0.35L$, the peak representing the average density of the smaller stripe (ρ_A) starts shifting towards the center. We observe similar effects on the ρ_B peak as w grows beyond the length scale of the phase outside the circular domain ($w > L - \sqrt{2}R \approx 0.65L$).

consistent with the analytical/continuum model shown in Fig. S2.3 with a minor correction on the threshold w value. It is noticeable that, however, the center portion of the LDD becomes more complicated and a third peak starts to emerge at large w values due to the circular geometry of the domain boundary. This presents little challenge since for our analysis we will focus only on the position and shifting of the two peaks representing the pure phases.

2.6 Radial Distribution Function and Lipid Heterogeneity - Model Study

2.6.1 Stripe Model Analytical RDF

We then derive the radial distribution function (RDF) of the model membrane from the density field:

$$g_\alpha(r) = \frac{\rho_\alpha^{(2)}(r)}{\rho_\alpha^2} \tag{S10}$$

$$= \begin{cases} \frac{2}{\pi\rho^2} kr + m & \text{if } 0 \le r < a \\ \left[\frac{2}{\pi\rho^2}(n-m)\arccos\left(\frac{a}{r}\right) + k\left(r - \sqrt{r^2-a^2}\right)\right] + m & \text{if } a \le r \le \frac{L}{2} \end{cases} \tag{S11}$$

where ρ is the average density:

$$\rho_\alpha = \langle \rho_\alpha(\boldsymbol{r}) \rangle \tag{S12}$$

$$= \frac{a}{L}\rho_A + \frac{L-a}{L}\rho_B \tag{S13}$$

and other parameters:

$$k = -\frac{(\rho_A - \rho_B)^2}{L} \tag{S14}$$

$$m = \frac{a}{L}\rho_A^2 + \frac{L-a}{L}\rho_B^2 \tag{S15}$$

$$n = \frac{L-a}{L}\rho_B^2 + \frac{2a}{L}\rho_A\rho_B \tag{S16}$$

The linear portion of the analytical expression of $g(r)$ predicts a monotonically decreasing RDF for a two-phase coexisting system with linear boundaries. We showed that the slope

of the linear part is directly related to $(\rho_A - \rho_B)$, or the difference between the lipid partial densities of the coexisting phases, which gives rise to a new perspective to interpret the RDF of lipid bilayer systems apart from the molecular level interactions.

2.6.2 Disc Model Analytical RDF

Due to the rather complicated geometric features of this model, deriving the analytical form of the RDF is not always practical. We thus implemented a discretized representation of the density function for the model in a two-dimensional grid box, and used it to calculate numerical radial distributions and structure factors for the model through a set of randomly generated and sampled configurations.

$$g_\alpha(r) = \begin{cases} \frac{(\rho_A - \rho_B)^2}{2\rho^2 L^2}\left[4R^2 \arccos\left(\frac{r}{2R}\right) - r\sqrt{4R^2 - r^2}\right] & \\ +\frac{\rho_B}{\rho^2 L^2}\left[\rho_B L^2 + 2\pi R^2 (\rho_A - \rho_B)\right] & \text{if } r \leq 2R \\ \\ \frac{\rho_B}{\rho^2 L^2}\left[\rho_B L^2 + 2\pi R^2 (\rho_A - \rho_B)\right] & \text{if } r > 2R \end{cases} \tag{S17}$$

Although a simple linear relationship does not exist between the lipid density contrast parameter and the functional form of the radial distribution, the density contrast still shows up as a pre-factor of the functional form and thus can be extracted from numerically fitting the radial distribution with as few as three unknown parameters, $\rho_A - \rho_B$, ρ_B, and R.

Chapter 3

IDENTIFYING THE ONSET OF PHASE SEPARATION IN QUATERNARY LIPID BILAYER SYSTEMS FROM COARSE-GRAINED SIMULATIONS

3.1 Introduction

The role of spatial composition heterogeneities in biological membranes is a long-standing problem in membrane biophysics.[3,18,55,57,98–100,106] The physical mechanism that gives rise to these heterogeneities as well as their role in biological function remain poorly understood.[18,79,97] Phospholipids, as major constituents of biological membranes, are believed to contribute to composition heterogeneity through their complex phase behavior.[38,113] It is well-known that under some conditions model membrane systems can exhibit spatial heterogeneity at the micrometer length scale.[19,48,107,111–113,119] This micron-scale heterogeneity is characterized by the formation of two distinct lipid phases both in perturbed cell membranes and in model membranes:[35,92] liquid-ordered (L_o) regions are enriched in cholesterol and lipids with saturated tails and high melting temperatures, while lipids with unsaturated tails and lower melting points are typically found in liquid-disordered (L_d) regions. These regions correspond to two coexisting thermodynamic phases, as shown for example by fluorescence microscopy experiments on giant unilamellar vesicles (GUVs):[112] starting from a homogeneous state at high temperature, the membrane spontaneously separates into liquid-ordered and liquid-disordered domains when cooled below a characteristic transition temperature, and these domains then diffuse and coalesce until complete phase separation is achieved. A multitude of experiments employing different techniques and membrane compositions have established that many model membranes can undergo phase separation into

coexisting L_o and L_d regions.[8,32,62,113,119]

In cellular plasma membranes no such large domains have been observed with conventional microscopy. However, experiments on cell-derived Giant plasma membrane vesicles (GPMVs) have shown the existence of liquid-order-like and liquid-disorder-like domains[9,110] and that such domains contribute to cell functions such as protein sorting.[19] With experimental techniques such as Förster resonance energy transfer (FRET), neutron or X-ray scattering, and super-resolution fluorescence, researchers have detected characteristic signals that are consistent with the existence of heterogeneity on nanometer length scales on unperturbed membranes of live cells.[22,92,93,102]

To reconcile the discrepancy in length scale of lipid spatial heterogeneity between different membrane systems and the connection between membrane heterogeneity and lipid composition, numerous studies have been performed on model membrane systems, especially unilamellar vesicles, with well-controlled lipid compositions. These experiments have provided insight into the mechanism that controls the length scale of heterogeneous domain formation using fluorescence microscopy, FRET, NMR, or other techniques.[4,25,30,32,48,49,54,56,77,112] It has been observed that model membrane systems composed of ternary mixtures of dipalmitoylphosphatidylcholine (DPPC), 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), and cholesterol show micron-scale heterogeneity at optical resolution, characterized as L_d–L_o phase coexistence below the miscibility temperature.[112] However, a mixture of DPPC, cholesterol and 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) did not exhibit optically observable lipid heterogeneity, while FRET and electron spin resonance (ESR) experiments suggest that heterogenities exist on nanometer length scales.[32]

To further probe this difference in length scale of membrane heterogeneity, four-component lipid mixtures were studied.[1,23,49] Here the onset of lipid heterogeneity can be investigated by manipulating the compositions of two unsaturated lipids (POPC and DOPC). Microscopy experiments on GUVs reveal the existence of an additional phase, characterized by stripe-like features on the vesicle surface, in between the phase-separated and homogeneous (but

nanoscopically ordered) regions.[48] Several models have been proposed to rationalize the emergence of such stripe-like morphologies, including the competition between line tension and curvature effects,[23] critical fluctuations,[36,110] and the existence of a nearby microemulsion phase.[96] Interestingly, these features were not observed in neutron scattering experiments on much smaller vesicles.[31]

Due to the elusive nature of submicron scale domains and difficulties in experimentally resolving and interpreting scattering data, computer simulations of multicomponent bilayers provide a promising approach to study their spatial organization. Along with previous simulations of cholesterol-containing membranes,[71,73,76,85] an all-atom membrane simulation on the microsecond time scale suggests that cholesterol plays an important role in L_o-L_d phase coexistence in a ternary lipid mixture.[101] In addition, several groups have used the coarse-grained MARTINI model[59,61] to simulate such membranes, and have confirmed that this model can qualitatively reproduce experimentally observed phase behavior.[1,6,81,86] Many of these studies use the polyunsaturated lipid ,2-dilinoleoyl-sn-glycero-3-phosphocholine (DUPC or DLiPC), instead of DOPC because the MARTINI model fails to reproduce the experimentally observed phase separation of DPPC/DOPC/cholesterol membranes. It is worth noting that a more recent modification of the MARTINI DOPC force field parameters[17] resolves this issue and produces more accurate phase behavior of bilayer systems that is consistent with experimental observations.

In this work, we present results from coarse-grained molecular dynamics (MD) simulations. We focus on a quaternary mixture of phosphatidylcholine (PC) lipids and cholesterol with different levels of unsaturation. Specifically, we use fully saturated DPPC, one-chain singly unsaturated POPC, doubly unsaturated DUPC, and cholesterol (CHOL) in a planar model lipid bilayer. We employ a statistical mechanics approach to investigate the phase coexistence behavior and the structural properties of nanoscopic and intermediate regimes of this mixture. We find a transition between the homogeneous and phase-separated regimes as we increase the DUPC content of the system. For the latter, we show how one can obtain the

compositions of the coexisting liquid-ordered and liquid-disordered phases both from partial density correlation functions and from local density distribution functions. To characterize the nature and length scale of spatial heterogeneity for membrane systems and to identify the onset of lipid phase segregation, we apply these approaches to a series of quaternary lipid mixtures at various unsaturated lipid contents, as well as to systems of different sizes. We observe that the difference in composition between lipid domains is proportional to the global composition of the doubly unsaturated lipid (DUPC). Furthermore, we show that local density and composition distribution functions strongly depend on the size of the observation window chosen for the analysis, which allows us to quantify the length scale of heterogeneity. This length scale then plays a crucial role in identifying the nature of the heterogeneity. The hallmark feature of a system undergoing phase coexistence is that the length scale of heterogeneity is proportional to the size of the system, while a homogeneous system does not show such a dependence even in the presence of long-ranged correlations. Finally, we combine the results from these different analysis methods to identify the onset of phase separation in this quaternary lipid mixture.

3.2 Methods

3.2.1 Molecular Dynamics Simulation

Force Field

We use the coarse-grained MARTINI 2.0 force field, which speeds up bilayer simulations by as much as 3−4 orders of magnitude compared to atomistic models.[59,61] An explicit solvent representation was included in the force field with a four-to-one mapping of water molecules into a MARTINI water bead.

Bilayer Composition

Our choice of quaternary mixture is inspired by the experimental and computational studies of refs[48] and,[1] respectively. We study symmetric lipid bilayers that consist of one-third DPPC molecules, one-third cholesterol molecules, and the remaining third is made of varying amounts of POPC and DUPC molecules. To vary the lipid composition and observe its effect on membrane phase behavior, we move along the POPC/DUPC binary axis with the compositions of the other two species fixed. We define the composition variable

$$\chi = \frac{[\text{DUPC}]}{[\text{DUPC}] + [\text{POPC}]} \tag{3.1}$$

where [DUPC] and [POPC] denote the global partial lipid densities. We vary χ by changing the relative composition of POPC and DUPC while keeping [DPPC] and [CHOL] constant.

By changing χ from 0 to 1, we follow a path in composition space that starts at a ternary DPPC/POPC/cholesterol membrane and that ends at a ternary DPPC/DUPC/cholesterol membrane. The former is homogeneous, while the latter is known to separate into coexisting ordered and disordered phases. At intermediate values of χ, the membrane is a quaternary mixture of DPPC, POPC, DUPC, and cholesterol.

Simulation Parameters and Setup

Simulations were performed using the GROMACS software package (version 4.6.5).[10,78] The time step of the simulation was set to be 20 fs, which is typical for MARTINI simulations.[59] van der Waals and electrostatic interactions were truncated at 1.2 nm, with a smooth decay of the former starting at 0.9 nm. Temperature was controlled by a stochastic velocity rescaling thermostat.[14,109] Pressure was controlled by a semi-isotropic Berendsen barostat[10] with a reference pressure of 1 bar, which effectively maintains zero surface tension for the membrane.

The construction of the simulated systems proceeded in multiple steps. First, we used the INSANE program[115] to build small patches of lipid bilayers consisting of 24 DPPC, 24 POPC, and 24 cholesterol molecules, together with 618 MARTINI solvent particles. This system was equilibrated at 350 K, 5-fold replicated in each membrane direction, and then again equilibrated at 350 K. We then replaced a number of randomly selected POPC molecules with DUPC to obtain starting structures for our systems at different χ values. The resulting structures contained 1800 lipids and 14540 solvent particles and had a side length of approximately 40 nm. They served as starting points for production runs at 298 K.

For membranes with high DUPC content, visual inspection shows that large domains comparable to the length scale of the system are formed, which is a clear indication of phase coexistence. For systems with intermediate DUPC content, identifying the nature of spatial heterogeneity in lipid distribution is less obvious. To quantify the phase behavior of these systems, we developed a toolkit of several statistical analyses to probe the ensemble properties of the equilibrated bilayer systems. These quantities, including lipid partial density correlation functions (or radial distribution functions), lipid partial structure factors, and lipid local density/composition distributions, along with the continuum model and clustering analysis described in the Supporting Information, allow us to identify the onset of phase separation.

3.2.2 Radial Distribution Function

As a slight variation from the general definition of radial distribution function provided in Chapter 2, we defined the partial molecular density field of lipid species α for each leaflet as

$$\rho_\alpha(\boldsymbol{r}) = \sum_{i=1}^{N_\alpha} \delta(\boldsymbol{r} - \boldsymbol{r}_i) \tag{3.2}$$

where $\boldsymbol{r}_i$ is the (x, y) projection of the center of mass position between the two glycerol-ester beads of the i-th molecule of lipid species α and N_α is the total number of lipids α in the

leaflet. We used the MDAnalysis software package to identify the two leaflets of the bilayer.[24] For cholesterol molecules, due to their tendency to reside in and sometimes flip-flop between the two leaflets, the assignment is based on the molecular orientation defined by the vector pointing from the center of geometry of the molecule to the hydrophilic hydroxyl group.

We calculated the partial radial distribution function (RDF), also known as the partial pair correlation function, between lipid species α and β according to[16,27]

$$g_{\alpha\beta}(\boldsymbol{r}) = \frac{N}{\rho N_\alpha N_\beta} \left\langle \sum_{i=1}^{N_\alpha} \sum_{j=1}^{N_\beta}{}' \delta(\boldsymbol{r} - \boldsymbol{r}_{ij}) \right\rangle \tag{3.3}$$

where $\boldsymbol{r}_{ij}$ is the (x, y) projection of the intermolecular distance between the center of mass positions of the ith and jth lipid molecules and the prime symbol indicates that the $i = j$ term is excluded if $\alpha = \beta$. N_α and N_β are the total number of molecules for lipid type α and β, respectively, N is the total number of molecules in the leaflet, and ρ represents the average lipid density over the entire leaflet. Angular brackets denote the equilibrium average. The function (3.3) was then radially averaged to obtain $g_{\alpha\beta}(r)$.

In our density function definition, we followed the convention that the lipid position $\boldsymbol{r}_i$ is the center of mass between its two glycerol-ester beads,[86] projected onto the (x, y) plane. Other definitions of molecular position will yield slightly different pair correlation functions; however, these small variations are insignificant for the purpose of this study.

3.2.3 Structure Factor

As previously shown in Chapter 2, we defined the partial structure factor as[27]

$$S_{\alpha\beta}(\boldsymbol{k}) = \left\langle \frac{1}{N} \tilde{\rho}_\alpha(\boldsymbol{k}) \tilde{\rho}_\beta(-\boldsymbol{k}) \right\rangle \tag{3.4}$$

where

$$\tilde{\rho}_\alpha(\boldsymbol{k}) = \int \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{i\boldsymbol{k}\boldsymbol{r}} \rho_\alpha(\boldsymbol{r}) \tag{3.5}$$

is the Fourier transform of the partial lipid density function from eq (3.2). The structure factor and radial distribution function are related by [27]

$$S_{\alpha\beta}(\boldsymbol{k}) = x_\alpha(\delta_{\alpha\beta} + x_\beta)\rho \int g_{\alpha\beta}(r)\mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\mathrm{d}\boldsymbol{r} \tag{3.6}$$

where x_α and x_β are the mole fractions of lipid species α and β, respectively, and $\delta_{\alpha\beta}$ is the Kronecker delta.

3.2.4 Local Density Distribution (LDD)

We defined the local density distribution (LDD) exactly as described in Chapter 2, Eqn. (2.6), and it was calculated by sweeping the observation window laterally over each leaflet in every simulation frame. Also in Chapter 2 we showed that the characteristics of the LDD profile depend both on the degree of lipid segregation and on the observation window size, w. We provide detailed analysis of this dependence in the Results section of this chapter.

3.2.5 Local Composition Distribution (LCD)

In addition to the LDD calculation, we defined the local composition distribution (LCD) as a multidimensional probability distribution of each lipid species within an observation window of variable size. Although the LCD does not provide a local lipid density profile that is directly comparable to results from model RDF analysis, a significant advantage of LCD over LDD is that its multidimensional nature allows us to quantify the compositional correlations between lipid species. This allows us to better identify the enriched and depleted species in each domain, as well as to enforce the lever rule when we obtain the ternary lipid phase diagram for our systems.

Similar to the LDD, we also calculate the LCD function with the same observation window algorithm to analyze lipid composition on the bilayer system. Just like in the calculation of LDD, we employ observation windows at various sizes to sample the local composition of

each individual lipid species within the observation window, and the window then sweeps both leaflets separately for the entire bilayer to acquire statistical information on the local composition throughout the membrane. We compute the multidimensional probability distribution

$$P(\mathbf{\Gamma}) = \frac{1}{A}\int \mathrm{d}\boldsymbol{r} \left\langle \prod_{\alpha} \delta\left(\hat{\Gamma}_{\alpha}(\boldsymbol{r}) - \Gamma_{\alpha}\right) \right\rangle \tag{3.7}$$

where

$$\hat{\Gamma}_{\alpha}(\boldsymbol{r}) = \frac{N_{\alpha,w}(\boldsymbol{r})}{N_{tot,w}(\boldsymbol{r})} \tag{3.8}$$

is the local partial lipid composition of lipid species α within an observation window of size w centered at position $\boldsymbol{r}$ and $N_{\alpha,w}(\boldsymbol{r})$ and $N_{tot,w}(\boldsymbol{r})$ are the number of lipid α molecules and total number of all lipid molecules within the observation window centered at $\boldsymbol{r}$, respectively. $\mathbf{\Gamma}$ is the three- or four-dimensional vector of local compositions for each lipid species, and A is the total area of the membrane.

3.2.6 Gaussian Mixture Model

To analyze the multidimensional LCD profile for our four-component lipid mixture and quantify the onset of phase separation, we use a Gaussian mixture model as implemented in the Python Scikit-learn machine learning package developed by Pedregosa et al.[75] to identify the nature of the distribution of lipid composition. Inspired by the continuum bilayer model (described in the Supporting Information), we perform three population Gaussian cluster analysis to the multidimensional local lipid composition distribution in order to identify the lipid-rich and lipid-poor phases and their corresponding compositions. The Bayesian information criterion (BIC) for Gaussian mixture models was used to ascertain whether a three-population fit or a one-population fit is more appropriate,[46] corresponding to phase-separated and homogeneous membranes. The BIC is defined as

$$\mathrm{BIC} = n \cdot \ln\left(\hat{\sigma_e^2}\right) + k \cdot \ln(n) \tag{3.9}$$

where n is the number of data points in the random variable, k is the number of free parameters used in the fit, and $\hat{\sigma_e}^2$ is the error variance, defined as

$$\hat{\sigma_e^2} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x_i})^2 \tag{3.10}$$

where x is the random variable or, in this case, the local lipid composition. Details of the Gaussian mixture and BIC analysis are shown at the end of this chapter, in the Supplemental section.

3.3 Results

3.3.1 Domain Formation and Local Density Distribution Depend on Composition

Final configurations of four (out of 11) MD trajectories at a simulation time of 20 μs are shown in Figure 3.1. At low χ value, the system exhibits no apparent long-range order. As χ increases, we start to observe domains enriched in DUPC. Further increasing the DUPC composition causes formation of long-lasting, large-scale DUPC-rich and DUPC-poor regions. Due to the geometry of the rectangular simulation box together with the use of periodic boundary conditions, the shape of the coexisting domains can be either circular or striped, depending on the area fraction of the domains. Typically, these domains are in registry across leaflets after sufficient equilibration time, as reported in previous studies of phase separation in mixed lipid bilayers.[1,82,86]

To quantify the bilayer heterogeneity, we calculate the local lipid density distribution (LDD) function of the four lipid species. We calculate the number density of each lipid type within an observation window that is smaller than the simulation box size. Statistics on the local lipid density were collected at different locations of the observation window throughout the entire membrane and over the last 5μs of the simulation trajectories. Intuitively, one would expect the distribution to be unimodal if the membrane is homogeneous, i.e., the bilayer system has the same composition everywhere. On the other hand, if the bilayer consists of

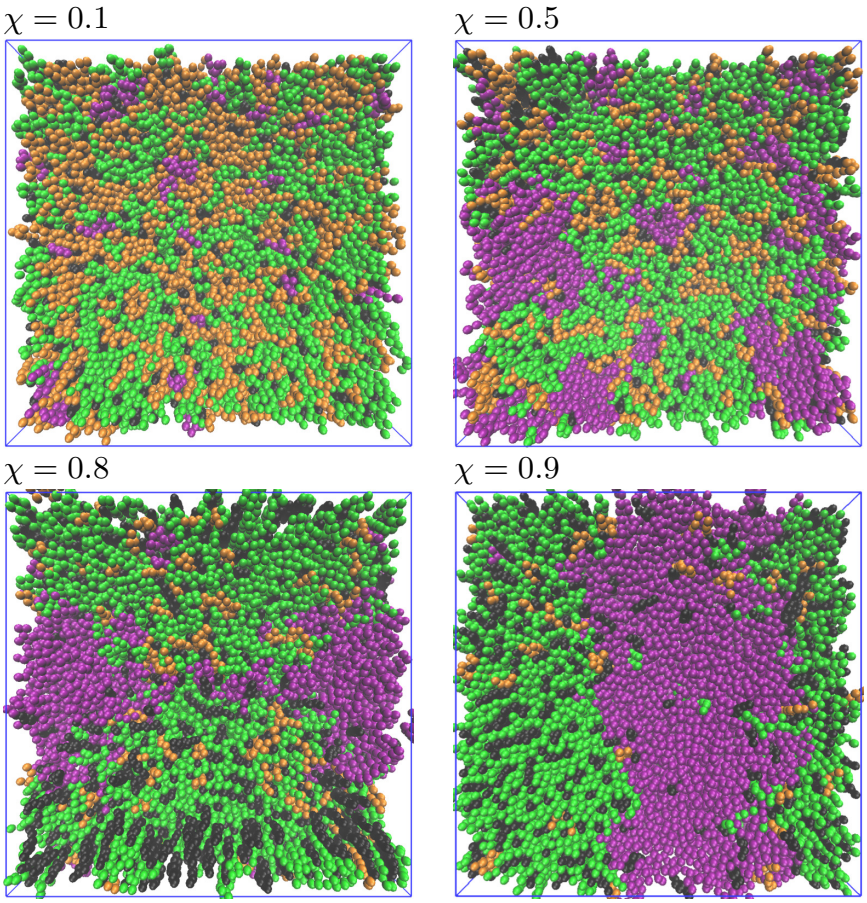

Figure 3.1: Simulation snapshots of quaternary lipid bilayers taken 20 μs after the initial temperature quench with increasing χ values of 0.1 (top left), 0.5 (bottom left), 0.8 (top right), and 0.9 (bottom right), with DPPC (green), DUPC (purple), POPC (orange), and cholesterol (black). At low χ, the system exhibits no apparent long-range order. As χ increases, we start to observe small clustering of domains enriched in DUPC. Further increasing the DUPC content induces phase separation into DUPC-rich and DUPC-poor regions.

distinct regions where a lipid species is either enriched or depleted, we expect a bimodal distribution of partial lipid densities indicating two distinct populations in the lipid density distribution. To make this intuitive picture quantitative, we devised a simple continuum model (shown in the Supporting Information) to capture the spatially heterogeneous nature of the membrane and to analyze it by deriving an analytical form of the LDD for such heterogeneous systems. The results are then used to analyze the lipid density distribution from bilayer MD simulations.

Results of the LDD calculation for all 4 molecular species and 11 studied membrane compositions are shown in Figure 3.2. We find unimodal local density distributions of DPPC molecules (top panel) in quaternary systems at low χ values $(0 - 0.4)$. This shows that no large-scale composition heterogeneity exists at and beyond the length scale of the 4 nm$\times$4 nm window we used. At higher χ (> 0.6), a bimodal pattern starts to emerge, indicating two distinct populations of local lipid densities, indicating the coexistence of DPPC-rich and DPPC-poor regions. We also observe an increase in the difference between the two peaks of the bimodal DPPC local density distribution for higher χ values, which suggests that the compositions in the two regions become increasingly different as more POPC is replaced by DUPC. It is noticeable that, as the distribution shifts from the unimodal regime to the bimodal regime, the emerging low density peak does not retain the Gaussian-type distribution of the unimodal peak. This is caused by the small number of DPPC molecules in the L_d domain and the small observation window size. In this case, the density distribution can no longer be properly described by a Gaussian distribution but should rather be modeled by a Poisson distribution, both of which are illustrated in a discrete membrane model in the Supporting Information. No significant difference is observed between the two leaflets, and the reported data were averaged across both leaflets. The local density distributions of DUPC molecules show a similar progression from unimodal to bimodal behavior as the composition variable χ increases, which strongly suggests a similar heterogeneous partitioning of DUPC molecules as in the case of DPPC. In contrast to DPPC, the global density of DUPC

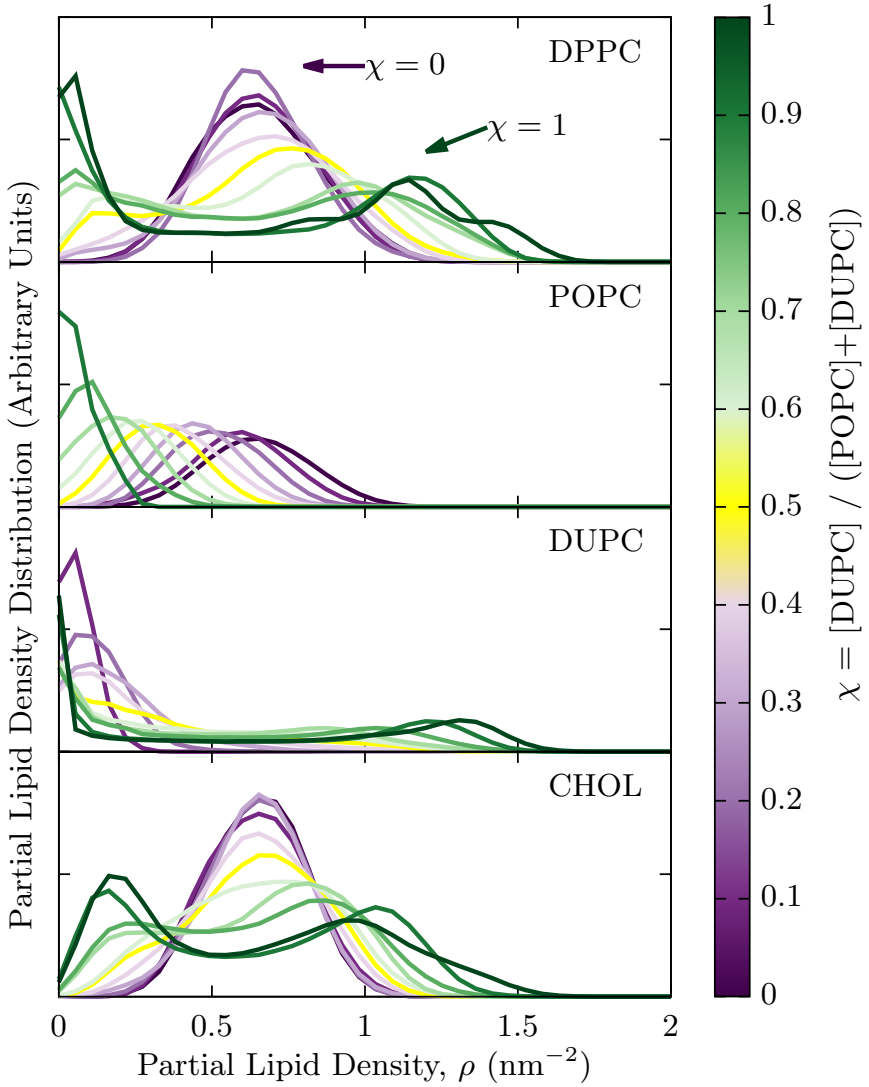

Figure 3.2: Local density population distributions calculated using the sampling window algorithm (with window size 4 nm × 4 nm) of all simulated bilayer systems for all lipid types. The colored spectrum and the arrows indicate the value of χ of the system, from high (green) to low (purple), with the halfway point singled out (yellow). For compositions with a low content of DUPC ($\chi < 0.5$), unimodal distributions of partial lipid density are observed for all lipid types, while the peak for DUPC is less Gaussian-like due to insufficient sampling of DUPC molecules at very low DUPC composition. For compositions with high χ values, bimodal distributions of lipid density are found for DPPC, DUPC, and cholesterol. For POPC, the density distribution remains unimodal while being shifted to the low density end due to a decrease in global POPC density, and the peak also gets distorted away from its Gaussian shape due to decreased sample size.

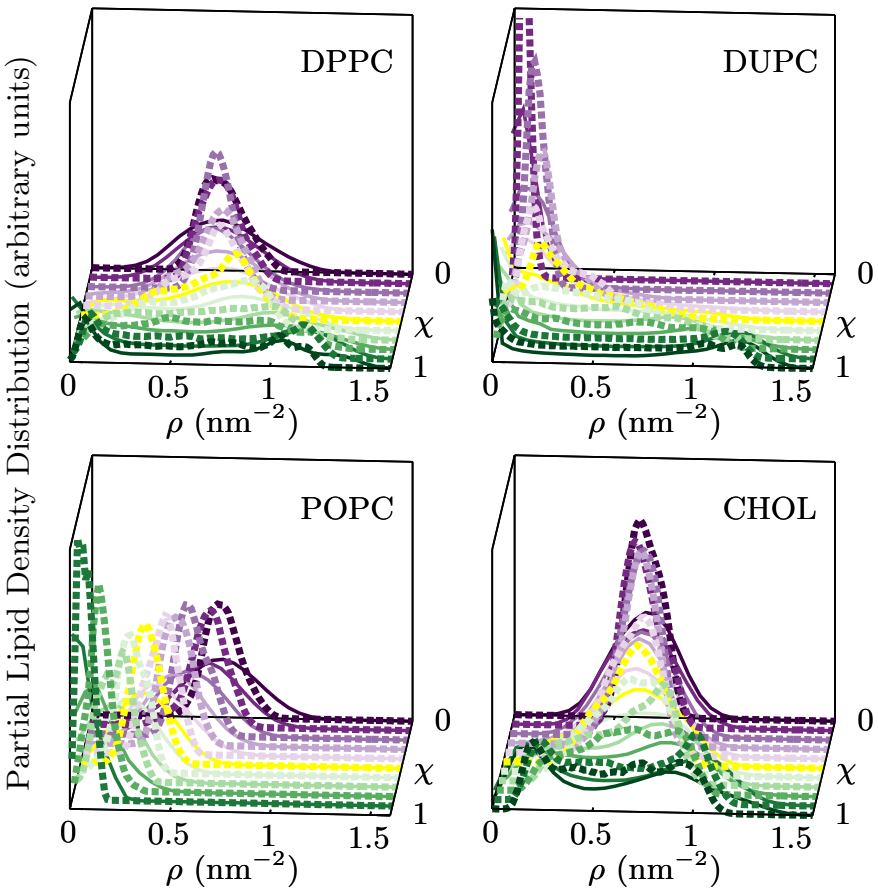

Figure 3.3: LDD calculations for all lipid species at various compositions with different window sizes, one of which was shown in Figure 3.2 with 4 nm × 4 nm observation window (solid lines) and the other one with a bigger observation window of 8 nm×8 nm (dashed lines). The color scheme is based on the composition variable χ and follows the same spectrum in Figure 3.2. At large χ values, both LDD profiles show bimodal distributions of lipid density for DPPC, DUPC, and cholesterol. It is noted that, for a bigger observation window size, the bimodal-to-unimodel transition in the density distribution happens at larger χ values compared to LDD with a smaller observation window. For DPPC, DUPC, and cholesterol LDDs at small χ values, as well as for POPC LDDs at all compositions, we observe a unimodal density distribution, indicating a homogeneous distribution at the observed length scales. The calculations with a bigger observation window show a narrower distribution in local density than those with a smaller window at all compositions.

changes with χ; therefore, the center of the peak shifts as χ changes, and the histograms look qualitatively different between DPPC and DUPC as a result.

For POPC molecules, no bimodal distribution is observed over the entire χ range. Due to the change in global POPC composition as we traverse the χ axis, the POPC composition peak shifts, similar to the case of DUPC. However, the local density distribution of POPC remains unimodal, which suggests that the POPC lipid does not preferentially partition into any of the coexisting phases in a significant way. When the global POPC density is low at small χ values, we again see deviations from Gaussian behavior as discussed previously.

The LDD profiles of cholesterol are very similar to those of DPPC. At small χ values ($\chi < 0.5$), the distributions show unimodal, Gaussian-like features. As χ increases, the distribution becomes bimodal. The average cholesterol density, or the mean of the distribution, is independent of χ because the overall cholesterol composition is kept constant. The separation between the cholesterol-rich domain density and the cholesterol-depleted domain density is not as large as that of the two phospholipids, indicating that the preferential partitioning of cholesterol between the two regions is not as strong as that of DPPC and DUPC. Another feature of the cholesterol LDD that is consistent with this observation is that the low density peak cholesterol is much better resolved and retains a Gaussian-type line shape around its maximum. Both observations indicate that the cholesterol content in the two coexisting regions is more similar than the content of the two phospholipids.

A closer investigation of the LDD profile for both the bilayer simulations and the simple continuum model (shown in the Supporting Information) reveals that lipid density distribution profiles depend strongly on the scale of potential heterogeneities and the observation window size by which this density profile is obtained. The same LDD calculation was performed with a different observation window size of 8 nm $\times$ 8 nm. Results are shown alongside those obtained with a smaller window in Figure 3.3. By using a larger sampling window size, we observe that, if a lipid composition shows unimodal density distributions with a smaller observation window, the larger window still indicates spatial homogeneity. It is noticeable

that the distribution gets narrower as the window gets larger, which is a consequence of the central limit theorem. More interestingly, for some systems where the LDD profile for the smaller window indicates two coexisting phases, that of the larger window shows no strong evidence of a bimodal distribution. Further analysis suggests that the two peaks in a bimodal density distribution start to merge when the observation window size exceeds a χ-dependent threshold value. Comparison with the simple continuum model described in the Supporting Information shows that this threshold size is an indicator for the length scale of the heterogeneity.

In the following two sections, we will first focus on quantifying the difference between coexisting phases and domains in order to understand the nature of the relationship between membrane composition and membrane morphology. Then, we will elucidate the dependence of LDD on observation window size and its implication on identifying the nature, as well as quantifying the degree, of lipid spatial heterogeneity in bilayer systems.

3.3.2 Using Density Correlation Functions to Identify the Onset of Phase Separation

Pair correlation or radial distribution functions (RDFs) provide a different route to studying membrane heterogeneity. Figure 3.4 shows results obtained from our coarse-grained simulations.

At short distances, the RDF provides information on molecular packing. It is noticeable that in this regime the RDF varies systematically with membrane composition. For example, in the case of DPPC, the height of the nearest-neighbor peak increases with χ. This behavior is expected for a system that undergoes a transition from a homogeneous to a phase-separated regime, where in the latter the DPPC molecules are concentrated in one domain and depleted in the other. The RDF of cholesterol shows long-range order at high χ values, indicating nearly crystalline order. This is a known artifact of the MARTINI coarse-grained model that has been rectified in recent versions of this force field.[38]

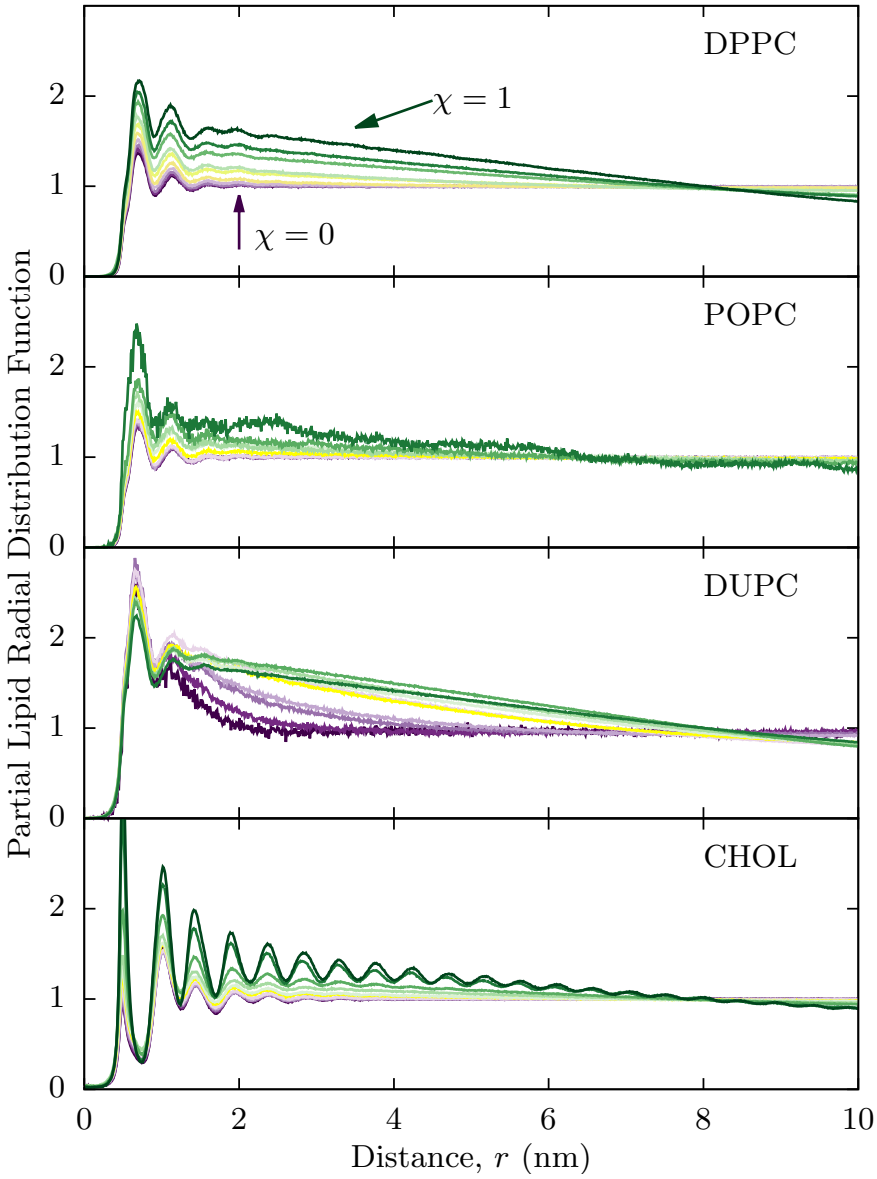

Figure 3.4: Partial radial distribution functions of lipids in bilayer systems for different molecule types at various compositions. The color scheme is the same as that in Figure 3.2, ranging from low χ (purple) to high χ (green) value, with $\chi = 0.5$ (yellow) being the halfway point. At low χ values, the RDF approaches unity, indicating composition homogeneity over large distances. At high χ values, the RDFs for DPPC, DUPC, and cholesterol show a characteristic long-range linear decay, a hallmark of phase separation, while the RDF of POPC converges to unity at long range for all χ values.

At large distances, the RDFs of DPPC, DUPC, and cholesterol exhibit a long-ranged, nearly linear decay to values less than unity at compositions rich in DUPC. This indicates that the system cannot be homogeneous at those large χ values, because homogeneity invariably translates into flat pair correlation functions at large distances, indicating the finite range of density correlations in bulk fluids. The fact that such uniformity is not observed even at length scales comparable to the size of the system is a strong indicator for thermodynamic phase separation.

To obtain further insight from these RDFs, we analyze in the Supporting Information a simple continuum model of two coexisting phases, both with spherical and stripe domains. We derive analytical expressions for the pair correlation function, which exhibits a nearly linear, slowly decaying behavior at large length scales, similar to those shown in Figure 3.4. The analysis of the continuum model shows that the slope of the linear part of the RDF is related to the difference $\Delta\rho$ in density of a lipid species between the two coexisting regions. With this analysis at hand, we can extract this important quantity for the quaternary membrane studied in our simulations.

At each simulated composition, the partial RDFs were calculated, and the long-ranged linear portions were fit to analytical functions derived from the models. Results of those fits are shown in Figure S6. Using the relationship obtained from the continuum model eq S11, we calculate the density contrast from the fitting parameters. Figure 3.5 shows the results of this approach for DPPC and compares them to the difference in peak positions of the LDD profiles discussed in the previous section. Both approaches find significant density contrast between the two phases at large values of χ, and their estimates of $\Delta\rho$ agree quantitatively. For $\chi \leq 0.4$, however, the two approaches give conflicting results: while the LDD for these compositions is unimodal, which points toward a single homogeneous phase, the linear fit to the pair correlation function yields a negative slope, which gives rise to small but nonzero estimates of the density contrast. In this regime, the RDF approach becomes unreliable, because the slope of the linear fit and therefore the estimate for $\Delta\rho$ become sensitive to the

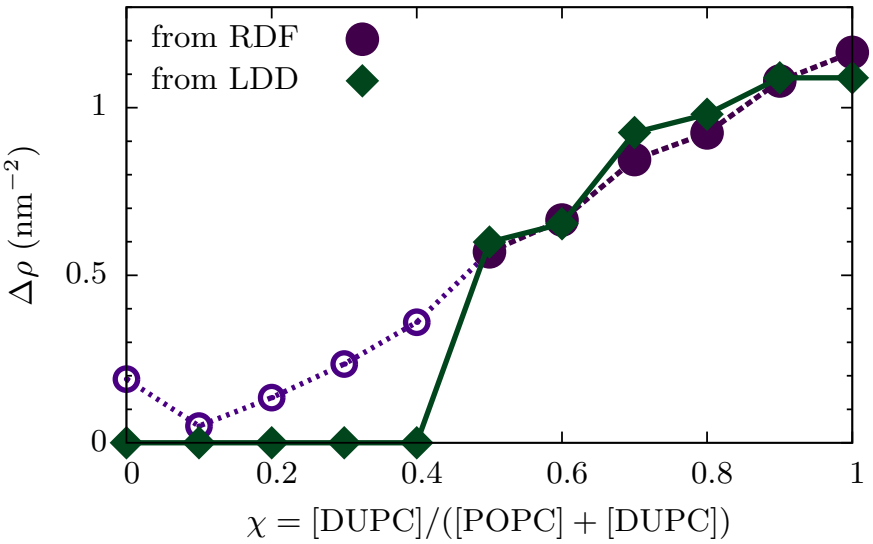

Figure 3.5: Density difference of DPPC between coexisting phases, calculated using the slope of the radial distribution function (Figure 3.4) and from the local density distribution function (Figure 3.2). At large values of χ, both the RDF and the LDD results show a significant density contrast between the DPPC-enriched and DPPC-depleted regions. The magnitude of this contrast depends on composition. At low values of χ, the RDF approach becomes less reliable (open symbols) due to difficulties in identifying and fitting the linear region of the density correlation function.

range over which the RDF is fitted. Furthermore, this method assumes that phase-separated domains, if they exist, form two rectangular stripes in the system. While this is the case for separated systems with high interface tension, this assumption breaks down if the interface tension becomes small and comparable to thermal fluctuations. The LDD method does not require such assumptions; we believe it is therefore the better approach to establish the presence of coexisting domains and the density contrast between them.

3.3.3 Analysis of Structure Factors

A complementary way to detect and characterize lipid heterogeneity is the analysis of the partial structure factors $S_{\alpha\beta}(\boldsymbol{k})$, defined in eq (3.4). While containing the same information as the pair correlation function, the structure factor is of interest for multiple reasons.

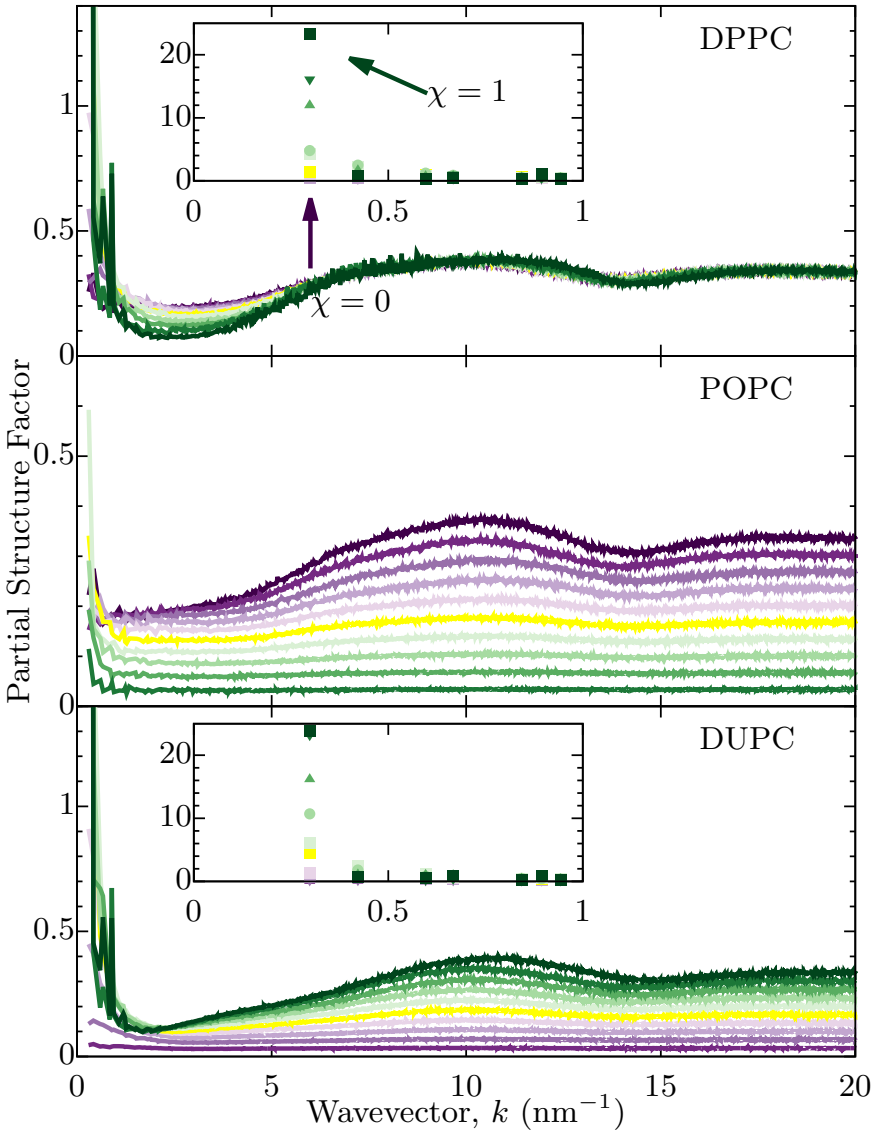

Figure 3.6: Radially averaged 2-d partial structure factors for all studied compositions and lipid species. The color scheme is the same as that in Figure 3.2. For lipid species with high concentrations (DPPC, DUPC at high χ, and POPC at low χ), we observe a broad peak at ~ 10 nm^{-1}, which originates from nearest neighbor packing. At very large k values, the structure factor converges to the global mole fraction of the given lipid, without accounting for cholesterol molecules. Large values of the structure factor at low k indicate significant density fluctuations on the scale of the system. Insets show that the magnitude of such density fluctuations is significantly larger for DPPC and DUPC than for POPC at high χ, while they are comparable at low values of χ.

First, it is directly related to data obtained in neutron or X-ray scattering experiments.[5,31] Second, the structure factor can be used to distinguish between unstructured fluids, structured fluids, and modulated phases. Recent work has suggested that seemingly homogeneous bilayers might in fact be microemulsions that are ordered on length scales not discernible in optical experiments.[90,95,96] The hallmark feature of this microemulsion phase is a peak in the structure factor at nonzero wavevector on length scales larger than those related to molecular packing.

Figure 3.6 shows partial structure factors for the three phospholipids in all studied systems. They show significant changes in bilayer structure as the composition is varied. Their normalization is such that they converge at large wavevectors to a lipid species' mole fraction in the membrane. As the parameter χ changes from 0 to 1, this mole fraction changes for POPC and DUPC but not for DPPC, which is why the partial structure factors of the latter overlap while those of the first two lipids are separated by a constant offset.

We are primarily interested in the low-wavevector (large-length-scale) regime. Both the DPPC and DUPC partial structure factors exhibit 5−10-fold higher low-wavevector intensities at large vs low χ values, indicating long-range order of density fluctuations for phase-separated systems and large-scale segregation of DPPC and DUPC lipid molecules. The POPC structure factor shows no significant variation in the low-wavevector regime with increasing χ. This suggests, as one would expect from the LDD and RDF results, that POPC does not participate strongly in the segregation of lipids, while DPPC and DUPC show strong partitioning into heterogeneous domains.

The nature of this apparent heterogeneity cannot be identified from the analysis of partial structure factors alone. A sharp increase of $S(k)$ as $k \to 0$ is consistent with two coexisting thermodynamic phases, separated over the length scale of the entire system. Alternatively, a peak in $S(k)$ at small but nonzero wavevector is consistent with a single, structured phase such as a microemulsion.[90,95,96] Our simulations cannot distinguish between these two scenarios, which would require additional data at lower wavevectors to test whether the observed

behavior of $S(k)$ is part of a monotonic increase at low wavevectors or the high-k flank of a peak in the structure factor. This data can only be provided by simulations spanning much larger system sizes, which are currently unfeasible.

Combining results from LDD, RDF, and structure factor calculations, along with a model-based analysis of lipid density distributions, we have established that both the length scale and the lipid composition of heterogeneous lipid domains depend on the membrane composition. Although the phase behavior at the near-ternary end points of χ has been well studied by experiments and simulations (including this work), the nature of this heterogeneity remains unclear. As indicated by the high intensity at low wavevectors in the partial structure factors, the size of heterogeneous domains is apparently often bound by the simulation box size, which does not allow us to distinguish between multiple segregated phases and a single structured fluid. While performing simulations on significantly larger systems is out of reach, we can obtain related information by varying the size of the observation window inherent to the LDD analysis method while keeping the system size fixed.

3.3.4 Dependence of Local Density Distributions on Observation Window Size

As previously discussed and illustrated in Figure 3.3, the LDD depends on the size w of the observation window chosen for the analysis. This dependence can be used to gain further insight into the nature of lipid heterogeneity in mixed bilayer systems. For example, Figure 3.7 shows the position of local maxima in the LDDs of the four lipid types as a function of w for two different membrane compositions.

For bilayers rich in DUPC ($\chi = 0.9$, top panel), we find that when the observation window is sufficiently small ($w < 8$ nm) two distinct populations of local lipid densities are observed for DPPC, DUPC, and cholesterol, indicating a separation of the bilayer into regions enriched and depleted of those species. POPC, on the other hand, does not participate in this separation. As predicted by the analytical form of the LDD in a simple continuum model (shown

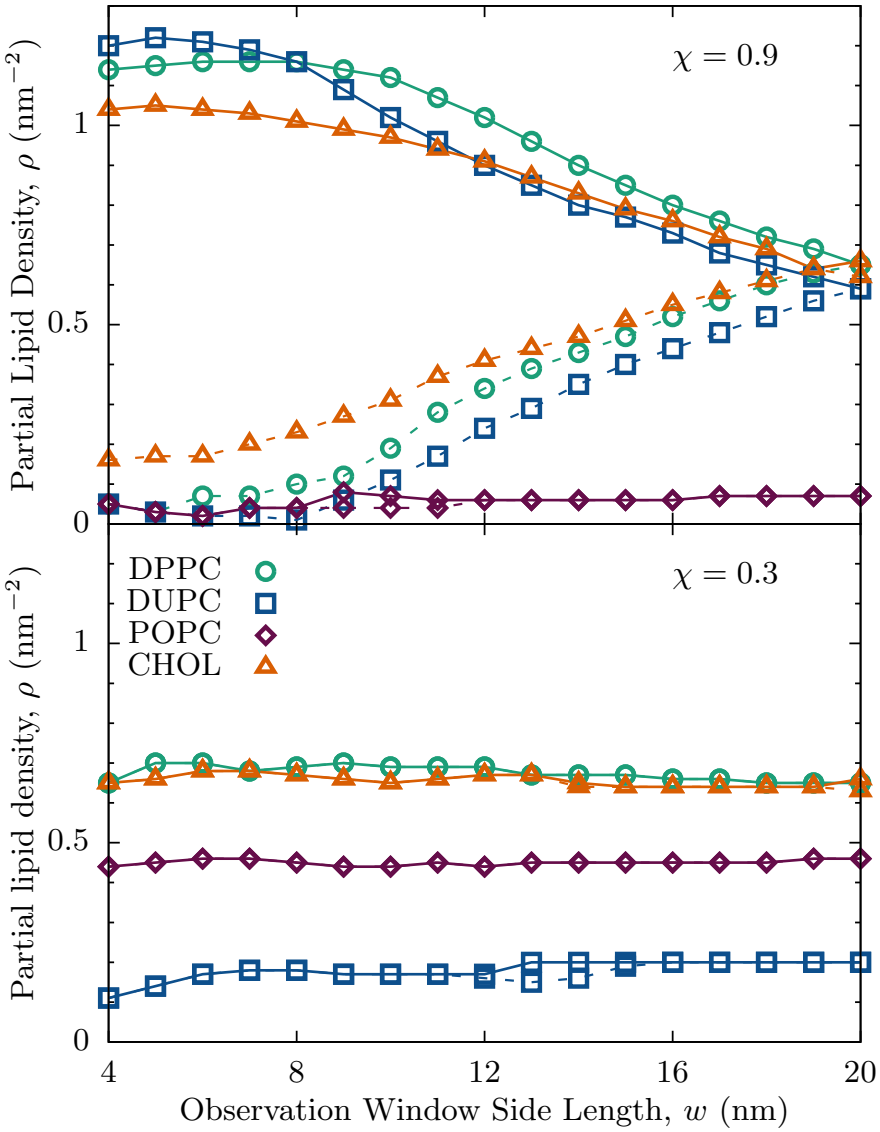

Figure 3.7: Peak positions in the local density distribution (LDD) function for various lipid species as a function of observation window size w. In a segregated system ($\chi = 0.9$, top panel), the LDDs for DPPC, DUPC, and cholesterol have two local maxima corresponding to the densities in distinct enriched or depleted domains, whereas the LDD of POPC only has a single peak independent of w. In a homogeneous system ($\chi = 0.3$, bottom panel), the LDDs of all species have only a single peak independent of observation window size.

in the Supporting Information), when the observation window size goes beyond the length scale of the heterogeneity, the spacing between the two peaks decreases, as the calculation now fails to sample the domain bulk lipid density due to the large observation window size. For the strongly phase-segregated system, the rich-phase density only becomes identical to the poor-phase density when the observation window is as big as the entire system. This suggests that the length scale of lipid density heterogeneity scales with the simulation box size, which is indicative of thermodynamic phase separation.

For bilayers with low DUPC content ($\chi = 0.3$, bottom panel), we find that the LDDs of each lipid species exhibit only a single maximum whose position is nearly independent of the observation window size w. These results suggest that the bilayer is laterally homogeneous over the entire range of length scales that we considered. It is noticeable that when the observation window gets very small (close to 4 nm) some lipids, especially DUPC, show a slightly lower than average local density. This is most likely due to the statistics of low numbers in the limit of low overall DUPC density and small observation windows. We attribute the brief appearance of a DUPC-poor peak around $w = 14$ nm to the uncertainty inherent in identifying local maxima in the noisy LDD that contains statistical uncertainty.

To further study the relationship between the degree of lipid heterogeneity and observation window size, we compute the lipid density contrast $\Delta\rho$ between the enriched and depleted regions for each lipid type as a function of global composition χ. These calculations are similar to the LDD calculation shown in Figure 3.5 for DPPC but are now done for all lipid types at various compositions and observation window sizes. Results are shown in Figure 3.8. First, we see that for reasonably small window sizes ($w < 11$ nm, about half of the simulation box size) the system shows no sign of heterogeneity for compositions with small χ, as the lipid density contrast between the two regimes remains zero. Second, for large χ, the system shows strong segregation, as shown by large differences in local lipid densities. These values correspond to the end points on the tie-line projected onto the axis of each lipid species. It is noticeable that for smaller observation window sizes ($w = 4$, 6, and 8 nm) the density

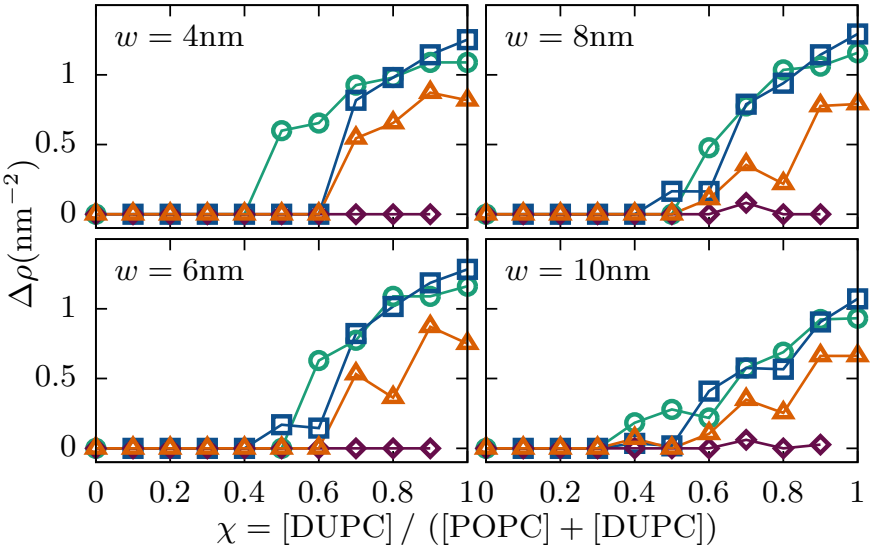

Figure 3.8: Density contrast between enriched and depleted local lipid densities as a function of observation window size. Symbols and colors are the same as those in Figure 3.7.

contrast for the same lipid type converges to the same value as χ approaches 1 consistently across different window sizes. This indicates that these observation windows are small enough to capture the bulk phase densities. For $w = 10$ nm, however, the maximum contrast at $\chi = 1$ starts to decrease compared to the value from smaller windows, suggesting that the observation window has grown beyond the length scale of heterogeneity of the system.

Our analysis of the local density distribution functions shows that in order to obtain accurate estimates of bulk lipid densities in coexisting domains one has to use an observation window size that is significantly larger than the size of an individual lipid but that cannot exceed the length scale of the heterogeneity. In the case of thermodynamic phase coexistence, the latter is determined by the size of the system. Characteristic changes in the LDD as the size of the observation window is varied can be used to demonstrate phase coexistence, and to distinguish it from other mechanisms that may lead to spatial heterogeneity.

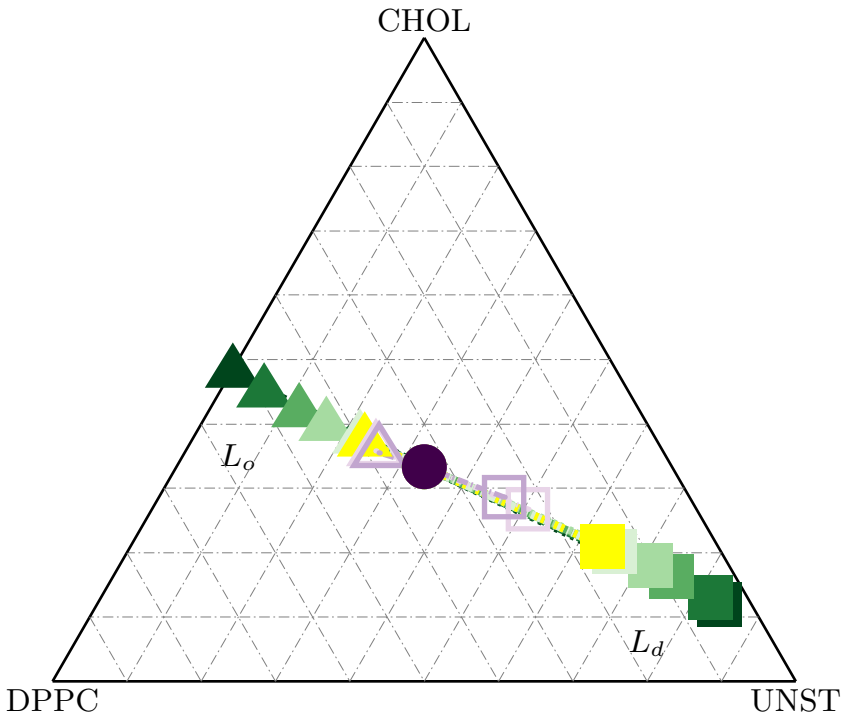

Figure 3.9: Multicomponent lipid phase diagram of the quaternary lipid mixture as obtained from the Gaussian mixture model analysis with observation window size $w = 5$ nm. UNST stands for the sum of unsaturated POPC and DUPC lipids. The colors are the same as those in Figure 3.2. At large χ values, the Bayesian information criterion (BIC) indicates strong evidence for two distinct populations of lipid species (solid symbols), and the system exhibits coexistence of a liquid-ordered (L_o) phase rich in DPPC and cholesterol (triangles) and a liquid-disordered (L_d) phase rich in unsaturated lipids (squares). At intermediate χ values, the evidence for phase separation is very weak (hollow symbols). At low χ, the LCD shows a unimodal distribution of lipid compositions and the BIC supports the identification of a single, uniform phase (solid circle).

3.3.5 Obtaining Composition Phase Diagrams Using Local Composition Distributions

The mixing behavior of multicomponent systems is best illustrated in composition phase diagrams. Compositions (or mole fractions) are generally preferred over molecular densities due to the built-in constraint that the mole fractions of all species must add to unity, thereby reducing the dimensionality from four to three. We therefore adapt our algorithm to calculate the LDD to compute local composition distribution (LCD) functions, defined by (3.7). From our simulation data, we obtain a large number of samples of local lipid compositions. As described in the Methods section and in the Supporting Information, these samples are then analyzed by fitting to a Gaussian mixture model with either one or three population centers and the Bayesian information criterion (BIC) is used to determine which model best describes the data. This approach allows us to distinguish a homogeneous system from a heterogeneous system in a statistically meaningful way.

To simplify the analysis and graphical representation of the mixing behavior, we chose to combine the two unsaturated lipids POPC and DUPC into a single component, denoted UNST, thereby further reducing the dimensionality from three to two. This simplification is justified by the LDD calculations shown in Figure 3.7, because POPC does not show significant separation across all studied observation length scales. Therefore, no information is lost by combining POPC with any other lipid species, or by leaving it out of the LCD calculation entirely. We chose to combine POPC with DUPC because the total number of these two lipids is conserved across all compositions considered in this study.

Figure 3.9 shows the ternary DPPC / UNST / cholesterol phase diagram obtained from the LCD analysis. At $\chi \geq 0.5$, the BIC indicates signifiant evidence for a multipopulation model of local compositions, which is consistent with a phase-separated system. As illustrated in the Supporting Information, the three populations correspond to the compositions of two coexisting bulk phases and an additional broad band that connects them, which stems from samples in which the observation window contains the interface between two domains.

One phase is rich in DPPC and cholesterol, while the other is rich in unsaturated lipids. These phases form the end-points of tie-lines that span the coexistence region in the phase diagram. We identify them with the liquid-ordered (L_o) and liquid-disordered (L_d) phases, respectively.

For $0.3 \leq \chi \leq 0.4$, there is still some evidence for multiple populations in composition space but that evidence is very weak. For $\chi \leq 0.2$, the BIC indicates that a three-population model overfits the data and that a one-population model more accurately represents the data. To these systems, we therefore assign only a single phase, with a composition equal to the global mole fractions of lipids.

To further study the role of POPC, we repeat the Gaussian mixture model analysis on the four-component compositions of all lipid species. Figure 3.10 shows a different rendering of the quaternary phase diagram obtained by projection onto the DPPC−POPC−DUPC plane, thereby ignoring the contribution of cholesterol. As shown previously, cholesterol preferentially partitions with DPPC into the L_o phase, and by leaving cholesterol out of the analysis, we therefore decrease the composition contrast between the two phases. Both phase diagrams are obtained from local composition data calculated with an observation window size of $w = 5$ nm. This window size is chosen because it is small enough to avoid the domain interface for most observations and because it is big enough to allow sufficient sampling of lipid molecules. This method is valid as long as the domains are at least as large as the chosen observation length scale. We find that the BIC scores again indicate a regime of strong evidence for phase separation ($0.5 \leq \chi \leq 1$), a regime of weak evidence for two lipid composition populations ($\chi = 0.4, 0.3$), and a regime of complete homogeneity in lipid compositions ($\chi < 0.3$). We also observe that within the strong segregation regime the POPC molecules partition weakly into the liquid-ordered phase, as indicated by the tilted tie-lines in Figure 3.10.

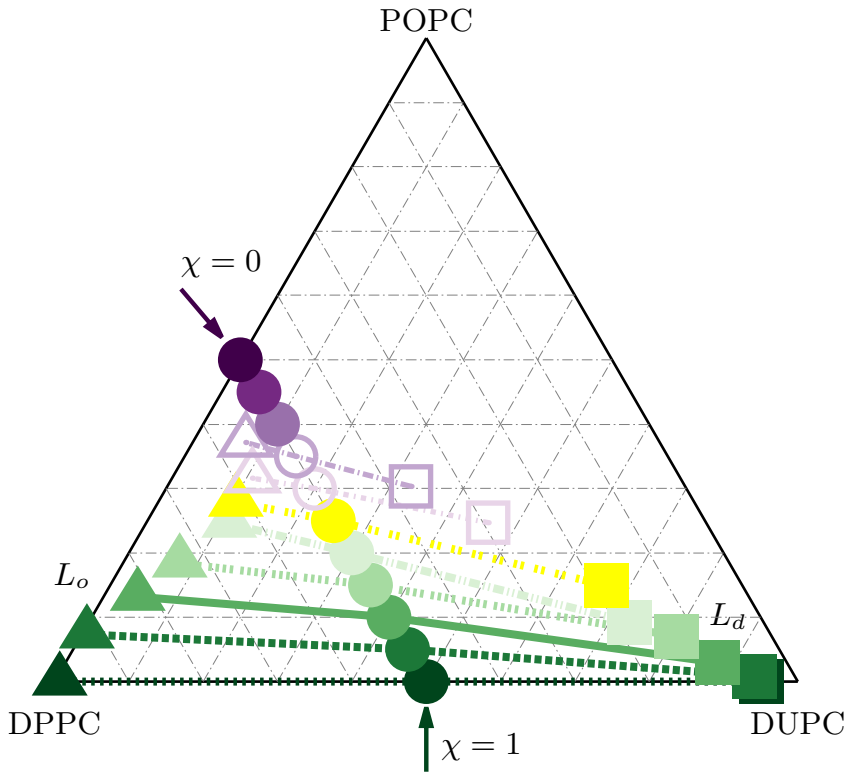

Figure 3.10: Multicomponent phase diagram obtained using a four-dimensional Gaussian mixture model, and neglecting the cholesterol component for ease of graphical illustration. The chosen observation window size is $w = 5$ nm. Colors and symbols are the same as those in Figure 3.9. In this representation, the $\chi = 0$ system lies on the DPPC/POPC axis, while $\chi = 1$ corresponds to a point on the DPPC/DUPC axis. The latter shows strong separation into a liquid-ordered and a liquid-disordered phase. The contrast between these phases decreases with decreasing χ, and eventually vanishes as the system becomes homogeneous. The tie-lines spanning the coexistence region are slightly tilted, indicating a weak preference of POPC for the ordered phase.

3.3.6 Identifying Phase Coexistence via System Size Dependence of Composition Distributions

We previously showed that the peak positions in one-dimensional LDDs at large χ have the dependence on the size w of the observation window that one would expect for a strongly segregated system. To pinpoint the mechanism that causes this segregation, we now study the dependence of local distribution functions on the size L of the system. The hallmark characteristic of phase separation is that the associated length scale of heterogeneity is proportional to the system size. In addition, it has recently been shown that phase separation can be artificially suppressed in computer simulations if the size of the system is too small due to the different scaling of mixing entropy and interface energy.[74]

To verify that phase separation is indeed the underlying mechanism for the observed heterogeneity in our system, we performed an additional set of simulations on much larger membrane systems, obtained by replicating the initial conditions of the previous simulations twofold in each membrane direction. Once equilibrated, we repeated the Gaussian mixture analysis described in the previous section on this new data set.

Figure 3.11 compares the compositions obtained from this analysis for both system sizes and a wide range of observation window sizes for a strongly segregated system ($\chi = 1$). The two graphs are essentially the same, even though the ranges of the observation window size are different. At small w, we obtain stable lipid compositions in the two coexisting lipid domains. For this ternary mixture, we observe 52% DPPC, 48% cholesterol, and essentially no DUPC in the liquid-ordered domain as well as 85% DUPC, 11% cholesterol, and 4% DPPC in the liquid-disordered domain.

As illustrated in the simple continuum model, the apparent peaks in the LDD or LCD functions start to shift toward the bulk density/composition when the sampling window size surpasses the length scale of the heterogeneity. For the smaller system, we observe that this shift occurs at a length scale of 6 nm, while, for the large system, it starts when w

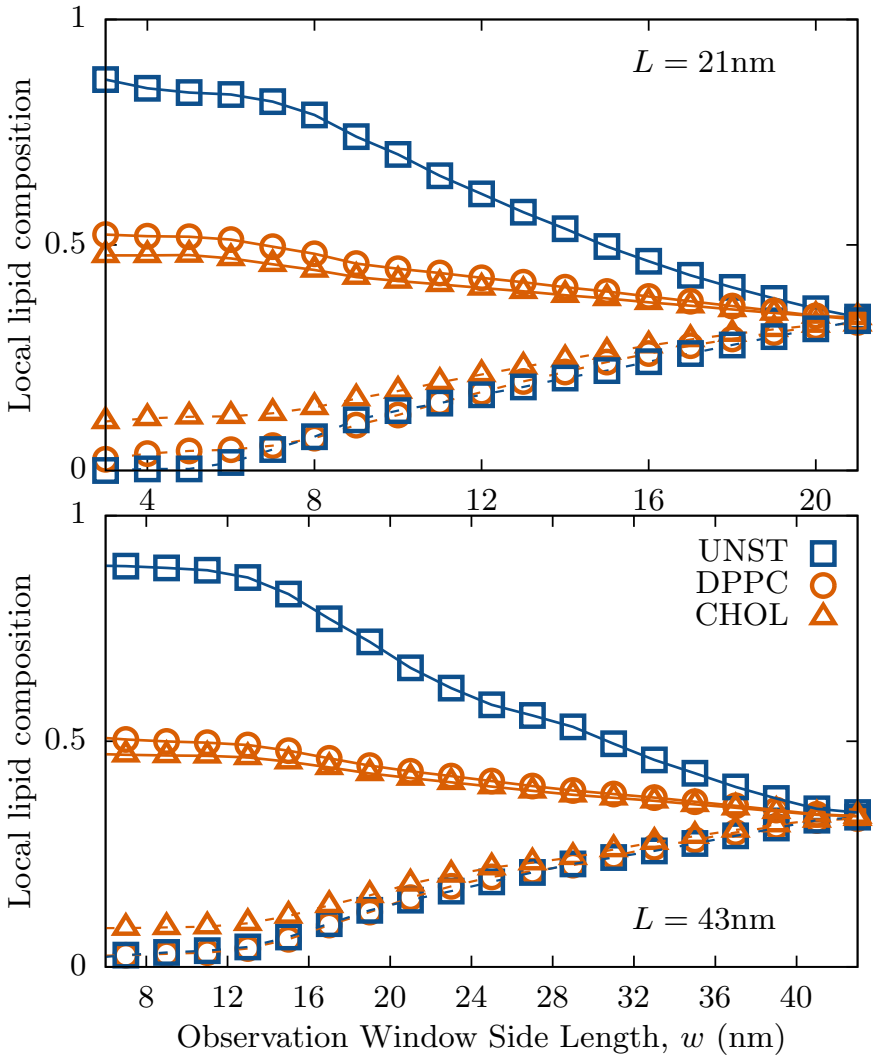

Figure 3.11: Local bilayer compositions as a function of observation window size obtained from the Gaussian mixture model for a small ($L = 21$ nm, top) and a large ($L = 43$ nm, bottom) membrane system at $\chi = 1$. The two graphs are essentially the same, which shows that the length scale of membrane heterogeneity is proportional to the system size, as is expected for thermodynamic phase separation.

exceeds 12 nm. This finding indicates that the length scale of heterogeneity of the strongly segregated system grows proportionally with the system size, which is further proof that phase separation is the mechanism that leads to this heterogeneity.

Finally, we note that the excellent agreement between the two graphs shown in Figure 3.11, which were obtained from independent simulations, indicates that the Gaussian mixture analysis yields robust results for the compositions of coexisting phases.

3.4 Discussion

The coarse-grained simulations presented in this Chapter capture similar trends along the χ axis as previous studies on multicomponent lipid bilayer systems in that (1) the scale of lateral heterogeneity, characterized by domain formation, increases with χ, as observed both in experiments[31,49] and in simulations,[1,6,17,26] (2) the composition difference between coexisting phases also increases with χ, in agreement with previous experimental and computational studies,[1,49] (3) signatures of phase separation at high χ such as the overall form of radial distribution functions are in qualitative agreement with previous simulations,[86] and (4) the phase diagram obtained from our simulations is in qualitative agreement with that obtained previously.[1]

However, this work also presents new features and perspectives of the problem. In the phase-separated regime, our model-facilitated analysis of the partial radial distribution functions enables us to quantitatively capture the composition information from coexisting lipid phases by relating it to the linear decay of the RDF over long distances. Both the size of the domains and the difference in their compositions can be extracted by comparison with the appropriate continuum model, and the results show a strong dependence of the length scale of bilayer heterogeneity on the lipid composition as described by χ. We have shown how the transition from unimodal to bimodal behavior of local distribution functions can be used to identify the onset of phase separation. The system size dependence of these functions is known to

contain information about the phase diagram, in particular the precise location of the critical point.[11,13] These functions also depend on the size of the chosen observation window, which we have utilized to verify the onset of phase separation. In addition, Bayesian analysis of Gaussian mixture models can further enhance the estimates of the phase boundary and of the compositions of the coexisting phases.

For compositions poor in DUPC, we have found a well-mixed, homogeneous phase characterized by short-ranged composition correlations and unimodal distribution functions. This result is at variance with neutron scattering experiments on a similar quaternary lipid mixture, which have revealed domains as small as 10 nm in size,[31] and that should therefore be detectable in our simulations. We also find no evidence for a structured fluid phase, such as a microemulsion, in this regime. However, the systems that we simulated are too small to rule out the presence of such a phase. The typical length scale of a bilayer microemulsion is expected to be on the order of 100 nm, which is significantly larger than the systems we can simulate over sufficiently long time scales. The signature of such a phase is a peak in the structure factor at nonzero wavevector. While we do find a significant increase in $S(k)$ at small k, our observations do not reach small enough wavevectors to determine whether this increase continues monotonically or whether it reverses, which would yield the characteristic peak.

Further studies, both experimental and computational, will be required to fully elucidate the nature of this material, which at least on the small length scales studied in our simulations seems homogeneous. The results presented in this work contribute to this quest by using novel approaches to accurately determine the boundaries of the phase coexistence region in this quaternary lipid bilayer.

3.5 Supplemental

3.5.1 Fitting Analytical Model to Simulation RDFs

Simulation DPPC partial RDFs are fitted with analytical RDFs derived from appropriate continuum models (equations shown in Chapter 2) for phase separated systems with $\chi = 0.5$ to $\chi = 1.0$. The linear portion of the simulation RDFs are identified based on the quality of initial linear regression fit, and the best linear region of the functions is used to extract the phase composition contrast as well as other parameters to get the compositions. We obtain good agreement with expected results for the fitted values. Structure factors are also calculated using the parameters obtained from the fit, and qualitative dependence of low wave-vector intensities on domain size and phase compositions is observed. Due to the simplicity of the model this agreement cannot be easily shown quantitatively for low wave-vector region of the analytical structure factor, so the results are not included in this analysis.

3.5.2 Gaussian Mixture Model Analysis and Bayesian Information Criterion

We perform Gaussian mixture model analysis on the multi-dimensional local composition profile of the system. For this example we fit the 3-component lipid local composition distribution with three 3-dimensional Gaussians, each representing the density distribution of the L_o phase (yellow), the L_d phase (red), and the flat region in between (brown). The mean of each Gaussian fit is shown as a black star on the histogram.

To test whether there is statistical evidence for three distinct evidence over a single population, we perform Bayesian Information Criterion (BIC) calculation show in Fig. 3.14. Since the change in BIC is calculated by:

$$\Delta\mathrm{BIC} = \mathrm{BIC}_{1-\mathrm{population}} - \mathrm{BIC}_{3-\mathrm{population}} \tag{S18}$$

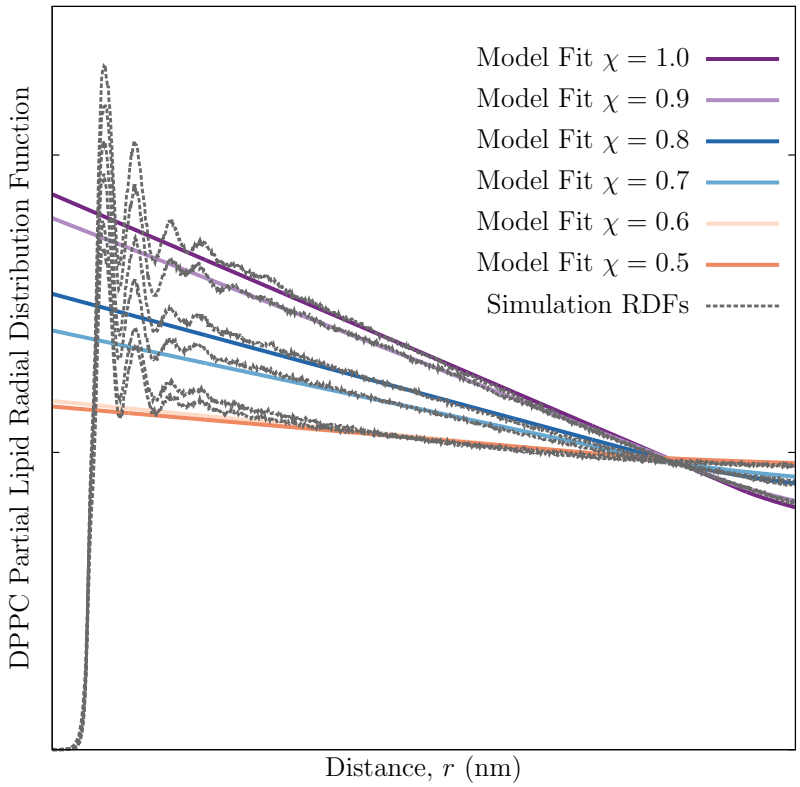

Figure 3.12: Simulation DPPC partial RDFs (gray dashed lines) fitted with RDFs derived from appropriate analytical models for phase separated systems with $\chi = 0.5$ to $\chi = 1.0$. The linear portions of the RDFs are fitted with the functions derived from the continuum models using a regression algorithm. Good linear fits are obtained for systems where phase separation is expected.

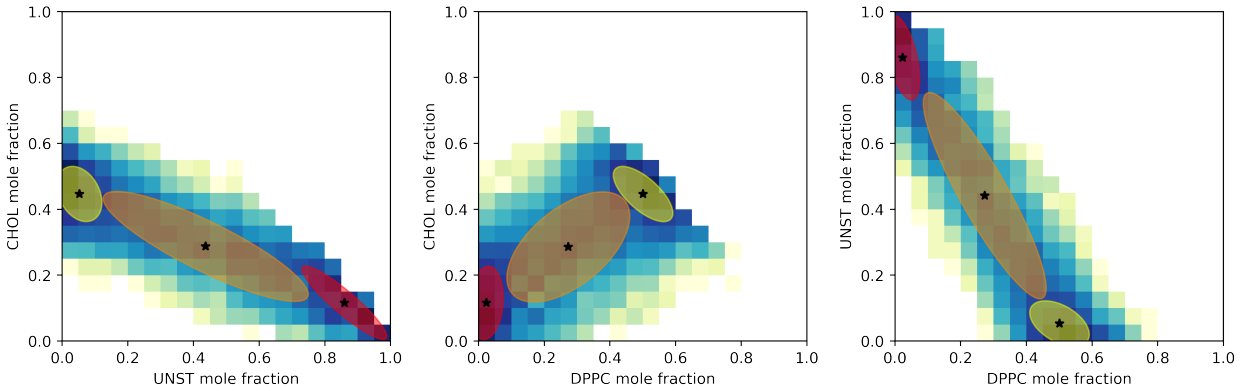

Figure 3.13: Gaussian mixture model analysis on multi-dimensional local lipid composition distribution for $\chi = 0.9$. The three panels are three dimensional composition histograms of DPPC, cholesterol and unsaturated lipids (POPC and DUPC) projected onto the three possible two dimensional axis. On top of the lipid density distribution histogram, we show three ellipses whose centers represent the means of the three-population Gaussian fit (marked with black stars). The long and short axis of the ellipses represent the value of the eigenvectors of the covariance matrices of the three-population Gaussian fit

and that lower BIC score indicates a better fit. A large value of ΔBIC indicates that the three-population model is justified. According to previous study on Bayes factors, the strength of evidence for the three-population model (against the one-population model) can be categorized by the ΔBIC values:[46]

ΔBIC_{1-3}	Evidence for three-population model
0 to 2	No/very weak evidence
2 to 6	Positive
6 to 10	Strong
> 10	Very strong

For $\chi \leq 0.2$, ΔBIC is close to zero, which suggests there is no evidence for three populations, and that a homogeneous distribution can best describe the density distribution of the system. For $\chi \geq 0.5$, especially when $\chi \geq 0.8$, the large value of ΔBIC indicates that a three-population Gaussian fit captures the features of the distribution much better than the

one-population model, and that the composition distribution is in fact heterogeneous. For intermediate χ values (0.3 and 0.4) there is weak evidence that the system is heterogeneous rather than homogeneous.

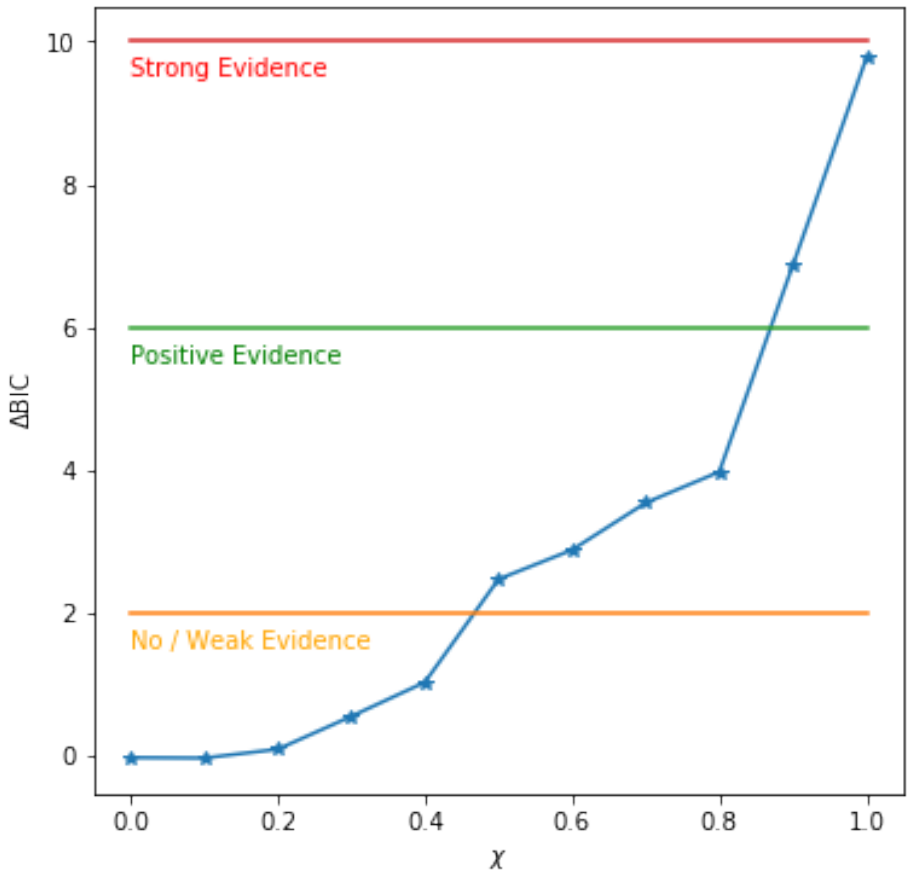

Figure 3.14: Change in BIC from a three-population Gaussian fit to a one-population Gaussian fit. For $\chi \leq 0.2$, ΔBIC is close to zero, which suggests there is merely any evidence that a three-population Gaussian fit is justified and a homogeneous distribution can best describe the density distribution of the system. For $\chi \geq 0.5$, especially when $\chi \geq 0.8$, the large value of ΔBIC indicates that a three-population Gaussian fit captures the features of the distribution much better than the one-population model, and that the composition profile is truly heterogeneous.

Chapter 4

INTERPRETING THE STRUCTURAL SIGNALS AND CHARACTERISTIC LENGTH SCALES OF SMALL DOMAINS IN CELL-DERIVED MEMBRANES AND SYNTHETIC MODEL MEMBRANES

NOTE: This chapter is based on published work by C. E. Cornell, A. D. Skinkle, S. He, I. Levental, K. R. Levental, and S. L. Keller.[15] Figures reproduced with permission.

4.1 Introduction

Although GUVs under most conditions exhibit micron-scale, coexisting liquid ordered (L_o) and liquid disordered (L_d) domains, which sizes are characterized by the size of the system, are observed, under some other conditions GUVs exhibit domains with smaller length scales. Our collaborators in the Keller lab and Levental lab,[15] as well as other researchers,[48] have observed these morphologies on both GUVs and GPMVs. In this work we analyze and interpret the fluorescence images of these vesicles. Our goal is to elucidate the mechanisms by which characteristic length scales varu in modulated phases and microemulsions.

As discussed in previous literature,[90,91,95,105] one major distinction between a modulated phase and a microemulsion is that, although both yield a peak at nonzero wavevector in a structure factor calculation and both show oscillating features in their density radial distribution function ($G(r)$), the decay in the magnitude of such an oscillation is significantly different. As shown in the schematics in Figure 4.1, the $G(r)$ of the modulated phase exhibits a slower, power law decay whereas the $G(r)$ of the microemulsion yields a faster, exponential decay. The difference arises due to a difference in the length scale and directionality of

the density fluctuation.The designation of modulated phase versus microemulsion does not determine the shape of the domains ("dots" or "discs" versus "stripes"). This observation is similar to the one in Chapter 3 in which there was no direct connection between the onset of phase separation in bilayer systems and the domain shape. In this chapter we focus on the methods we used to evaluate the characteristic length scales and fluorescence levels of images from both GUVs and GPMVs prepared by our collaborators, in order to test the proposed mechanisms of modulated phases and microemulsions.

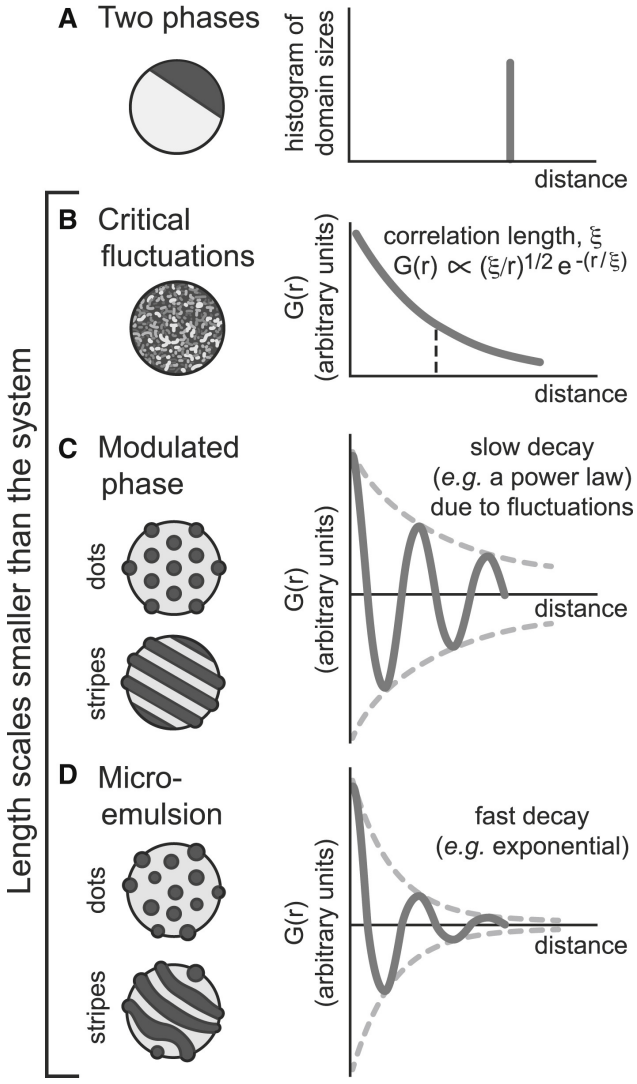

Figure 4.1: Schematic of different types of membrane behavior reproduced from published work by Cornell *et. al.*[15] (A) GUVs exhibiting two coexisting liquid phases that scale with the system size. (B) GUVs exhibiting critical fluctuations show a broad range of correlation length scales as indicated by the $G(r)$. (C) GUVs in modulated phases yield an oscillating $G(r)$ bounded by a slow, power law decay. (D) GUVs in microemulsion show an oscillating $G(r)$ bounded by a fast, exponential decay.

4.2 Methods

4.2.1 Experimental

GUVs and GPMVs were prepared at various temperatures and osmotic pressure by our experimental collaborators using electroformation and canine kidney cell culture, respectively. The GUVs were prepared using a synthetic lipid mixture of DiPhyPC/DPPC/cholesterol. The excess area and asymmetry in GUV membrane that are needed in order to induce the formation of dots and stripes was achieved by loading excess DPPC into GUV outer leaflets using DPPC-loaded cyclodextrin. GPMVs were not loaded with extra lipids. The surface tension for both types of the vesicles was carefully controlled by tuning the osmotic pressure of the exterior glucose solution for GUVs, and by controlling the solution NaCl concentration for GPMVs. The detailed procedures by which the vesicles were prepared are not the focus of this thesis and are thus not included. For more information, please refer to our publication.[15] Sample fluorescence images for GUVs and GPMVs were also obtained by our collaborators, and are shown in Figure 4.2 and Figure 4.3.

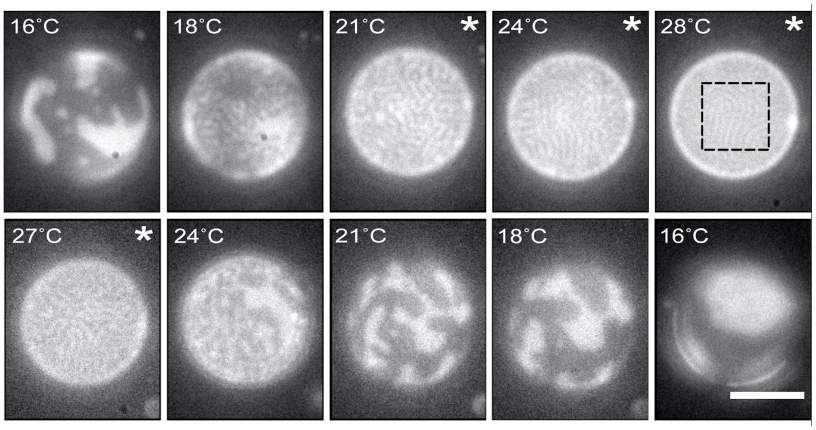

Figure 4.2: Fluorescence image for a single GUV going from 16 °C to 28 °C (top row), and then from 27 °C to 16 °C (bottom row). Scale bar, 20 μm.

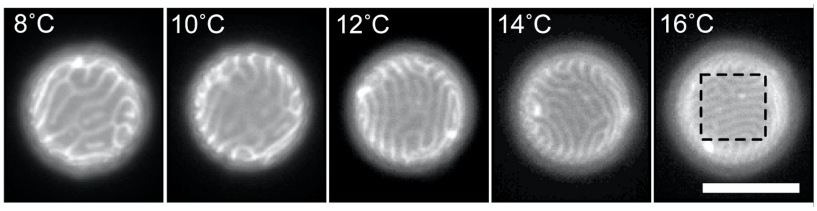

Figure 4.3: Micrographs of a single GPMV from 8 to 16 °C. Scalre bar, 10 μm.

4.2.2 Image Analysis

To analyze and interpret the fluorescence images obtained, we calculate strucutral properties such as radial distribution functions and structure factors of the images. Following the general definition given in Chapter 2, we modified these calculations for the two-dimensional images.

We defined the two dimensional radial distribution function (RDF) for the image as:

$$G\left(\boldsymbol{r}\right) = \frac{\langle \delta\rho(\boldsymbol{r'}+\boldsymbol{r})\delta\rho(\boldsymbol{r'})\rangle}{\langle \delta\rho(\boldsymbol{r})\rangle^2} \tag{4.1}$$

where $\delta\rho(\boldsymbol{r})$ is the contrast between the two dimensional image grey value vector and the image average grey value:

$$\delta\rho(\boldsymbol{r}) = \rho(\boldsymbol{r}) - \bar{\rho} \tag{4.2}$$

and

$$\bar{\rho} = \frac{\int \mathrm{d}\boldsymbol{r}\rho(\boldsymbol{r})}{\int \mathrm{d}\boldsymbol{r}}. \tag{4.3}$$

To implement the above definition for our pixelated image, we put the grey value vector into matrix representation:

$$\rho(\boldsymbol{r}) = \rho_{i,j} \tag{4.4}$$

where $\boldsymbol{r}$ is the position vector represented by a matrix element as a pixel at i-th row and

j-th column. The average density can be calculated as:

$$\bar{\rho} = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\rho_{i,j}}{mn} \tag{4.5}$$

and the contrast:

$$\delta\rho(\boldsymbol{r}) = \delta\rho_{ij} = \rho_{i,j} - \bar{\rho} \tag{4.6}$$

Thus, the 2-D $G(\boldsymbol{r})$ becomes:

$$G(\boldsymbol{r}) = G(r_{i,j}) = \frac{\left(\sum_{k,l=1}^{m,n} \delta\rho_{k,l} \cdot \delta\rho_{k+i,l+j} \right)}{\left(\sum_{k,l=1}^{m,n} \delta\rho_{k,l} \right)^2} \tag{4.7}$$

for all $i \in (1, m)$ and $j \in (1, n)$. The one dimensional RDF, $G(r)$, is the two dimensional function that reports modulation in the direction perpendicular to the stripes and is averaged over the direction parallel to the stripes in the image, assuming translational invariance:

$$G(r) = G(r_i) = \sum_{j=1}^{n} \frac{G(r_{i,j})}{n} \tag{4.8}$$

Due to the finite size of the image, the calculation yields fewer data points as the separation between two correlated pixels increases and approaches the image size, thus the quality of $G(r)$ degrades as r increases.

We also calculated the one dimensional structure factor as a Fourier transform of the 1-D $G(r)$ of the image using the generic definition:[27]

$$S(k) = 1 + \bar{\rho} \int G(r) \mathrm{e}^{-i\boldsymbol{k} \cdot \boldsymbol{r}} \mathrm{d}\boldsymbol{r} \tag{4.9}$$

The actual Fouier transform is carried out using the fast Fourier transform function provided in the NumPy v 1.14 Python software package.[41,72]

4.3 Results

4.3.1 Radial Distribution Function and Structure Factor Calculation Confirms that the Characteristic Length Scale is Inversely Related to Temperature

As a review of the aspects of with phase-separated membranes we discussed in Chapter 3, we note that for vesicles in which L_o/L_d phases coexist, the compositions of the lipids in the two phases become more similar as temperature increases, causing the line tension between the Lo and Ld phases to decrease, and lowering the energetic cost to form an interface between the two phases.

For GUV and GPMV membranes with significant excess area, we quantify the characteristic wavelength of stripes by computing $G(r)$ and $S(k)$ of the images, where the characteristic wavevector (k) appears as a peak in the $S(k)$. For example, the image taken from Figure 4.4 A at 16^{circ}C) yields the $S(k)$ in Figure 4.4 C, which has a peak corresponding to a characteristic wavelength of $0.89 \pm 0.08\,\mu$m. The autocorrelation, $G(r)$, was calculated by aligning the striped domains in the y-direction for optimal contrast, then thresholding the gray value from the fluorescence image to enhance signal to noise ratio and to better capture the characteristic wavelength. In Figure 4.5 B, only the local maxima of the peaks in $G(r)$ (shown as squares) were used to fit the exponential curve and the power law. Both curves fit with statistical significance based on the R^2 statistics. The structure factor, $S(k)$, in Figure 4.5 C was calculated using fast one-dimensional Fourier transform of the radially averaged $G(r)$ shown in Figure 4.5 B. The same calculations were done for other temperatures and the results are shown in Figure 4.8.

Similar calculations were done for GUVs, with a sample calculation shown in Figure 4.5 and the rest in Figure 4.7.

By examining all the characteristic wavelengths calculated from the $S(k)$ peak wavenumber, in Figure 4.6 A and B, we showed that the characteristic wavelengths decrease with tempera-

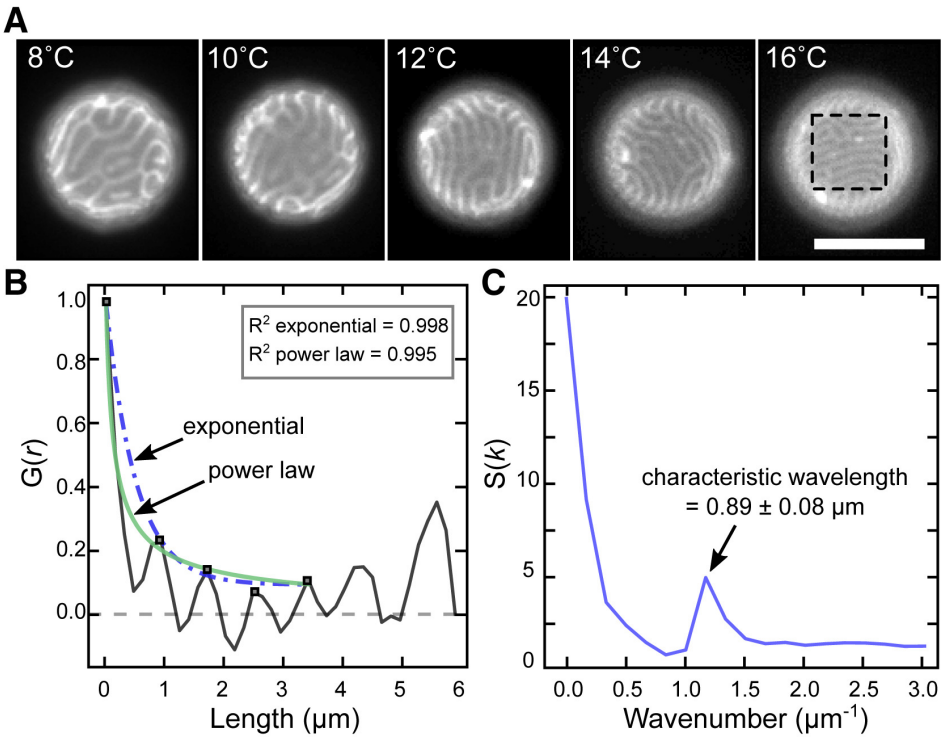

Figure 4.4: Sample calculations for $G(r)$ and $S(k)$ for a GPMV at 16 °C. (A) The same micrographs as shown in Figure 4.3, with the area within the dashed box at 16 °C used for the sample calculation. Scale bars, 10 μm. (B) Autocorrelation, $G(r)$, of the area in the dashed box at 16 °C, where five maxima (squares) are fitted to an exponential curve and a power law, yielding similar R^2 coefficients. (C) Structure factor, $S(k)$, calculated from the same $G(r)$ in (B). The peak in $S(k)$ corresponds to the Fourier mode of the characteristic wavelength, which is 0.89 $\pm$ 0.08 μm, where the uncertainty is half the length of a pixel.

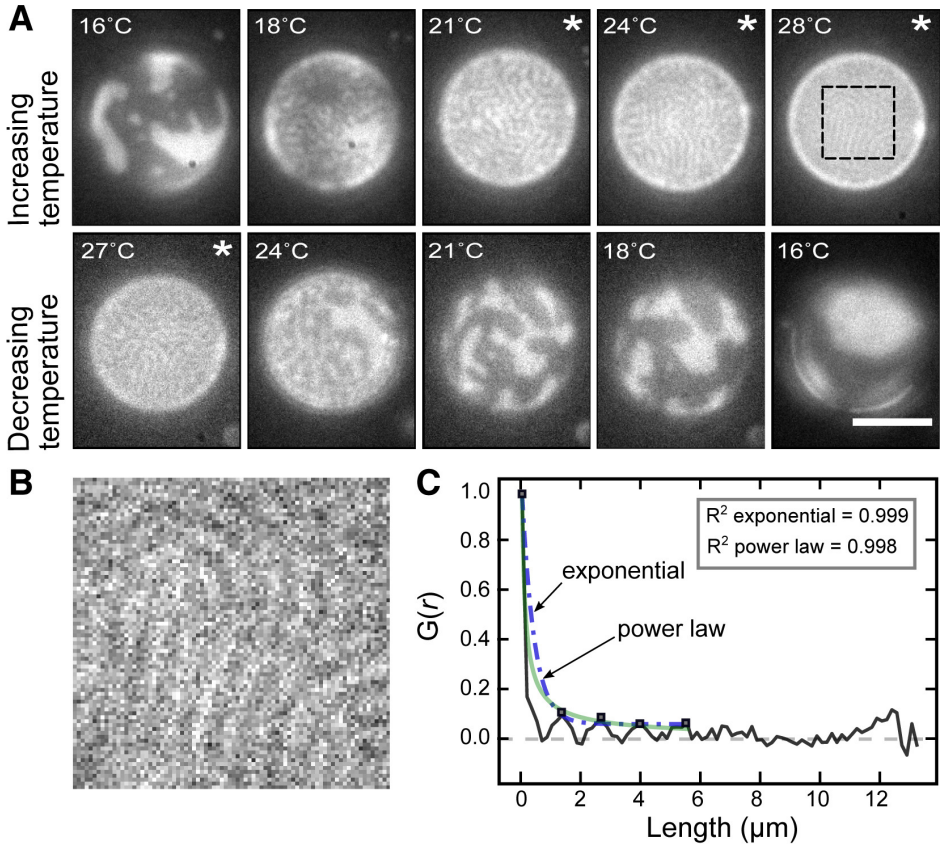

Figure 4.5: Sample calculations for $G(r)$ and $S(k)$ for a GUV at 28 °C. (A) The same micrographs as shown in Figure 4.2, with the area within the dashed box at 28 °C used for the sample calculation. Scale bars, 20 μm. (B) An enlarged image of the dashed box at 28 °C. (C) Autocorrelation, $G(r)$, of the area in the dashed box at 28 °C, where five maxima (squares) are fitted to an exponential curve and a power law, yielding similar R^2 coefficients.

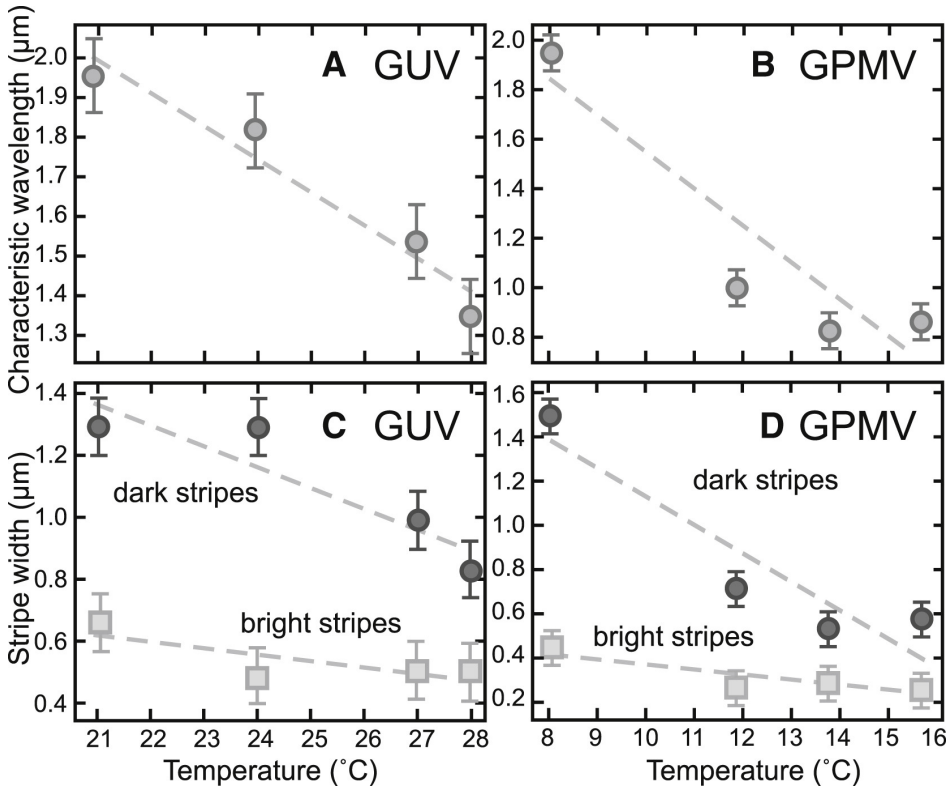

Figure 4.6: Characteristic wavelengths decrease with increasing temperature. (A) and (C) correspond to the single GUV in Figure 4.2 (using only images from temperatures marked with asterisks). (B) and (D) correspond to the single GPMV in Figure 4.3 (excluding the out of focus image at 10 °C).

ture for both GUVs and GPMVs. Since each wavelength corresponds to the summed widths of one dark stripe and one white stripe, the width of a dark stripe corresponds to wavelength minus the width of the bright stripe, where the width of the bright stripe corresponds to the half-width of the first peak in the $G(r)$. Thus the widths of both the dark and bright stripes decrease with temperature for both GUVs and GPMVs (Figure 4.6, C and D).

Figure 4.4 B and 4.5 C also illustrate that when domains are wrapped on a sphere with diameter $\sim 10\,\mu$m, the distinction between a modulated phase and microemulsion is difficult to assess from a single image. Visually, the small-scale stripes in Figsure 4.2 and 4.3 show similarities to simulations of a modulated phase by Sholomovitz and others.[95] However,

within uncertainty, the sample $G(r)$ calculations shown in 4.5 C and 4.5 B are equally well bounded by a function that decays with a power law, as one would expect from a modulated phase, and with an exponential, as a microemulsion would.

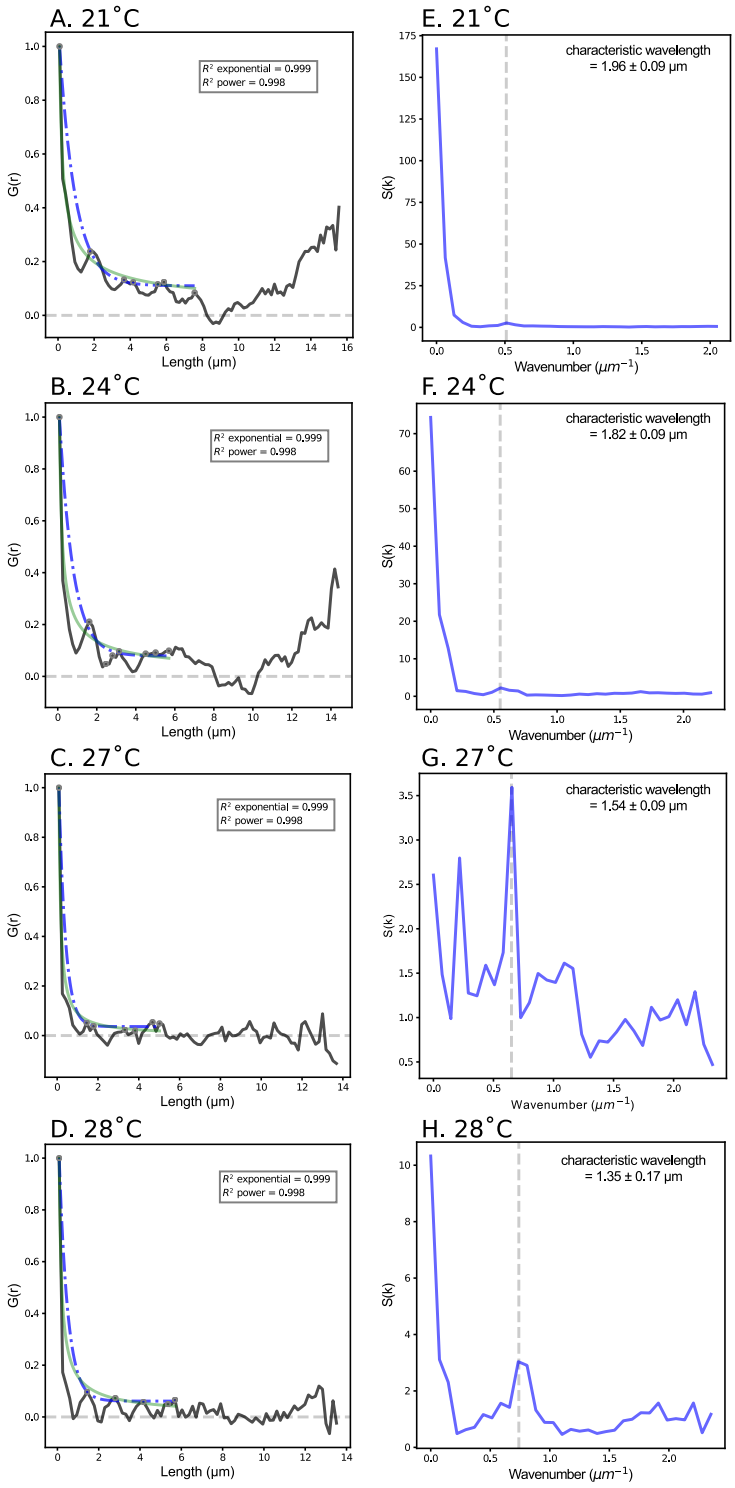

Figure 4.7: G(r) and S(k) for images GUVs at four of the temperatures shown in Figure 4.2 and 4.5, the green curves are fits to a power law and the dot-dashed curves are fits to an exponential.

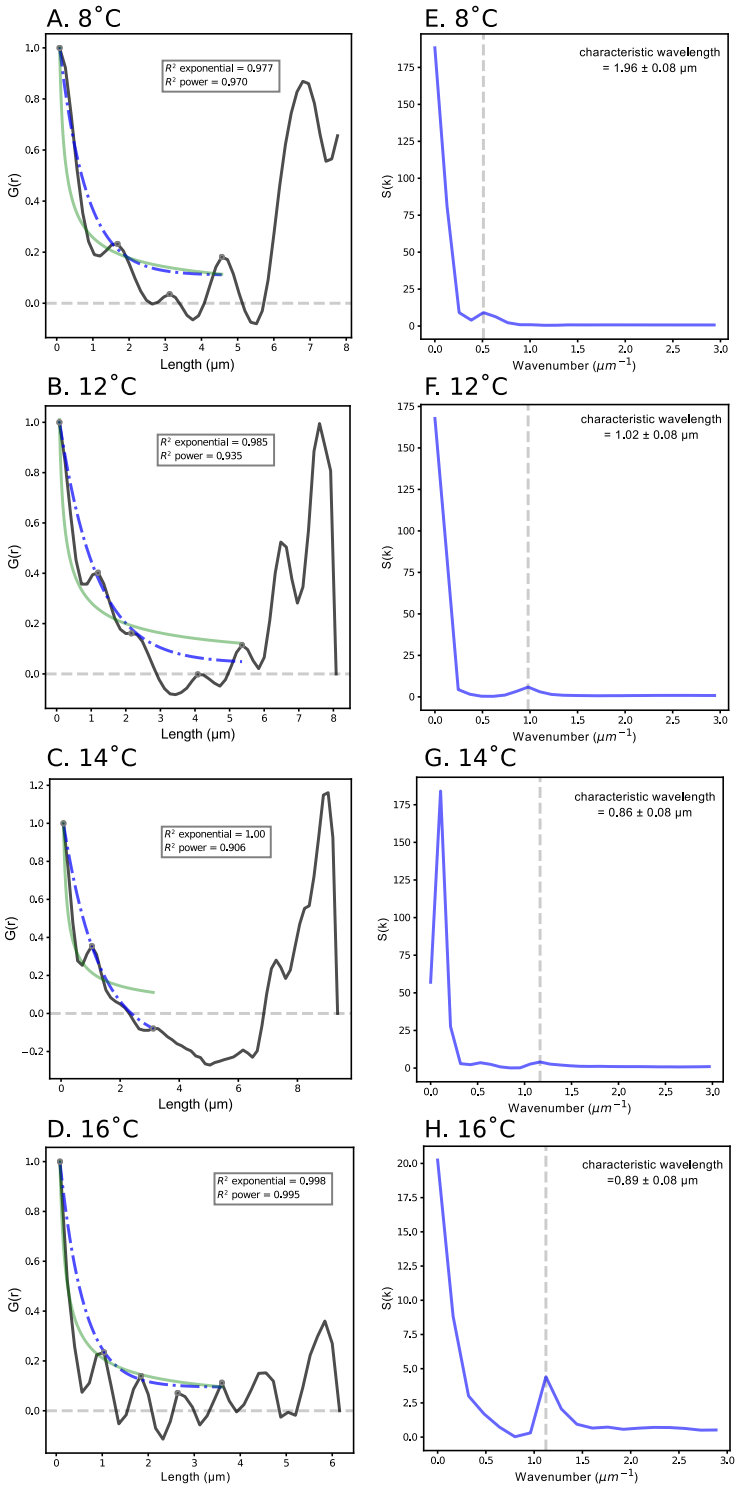

Figure 4.8: G(r) and S(k) for images GPMVs at four of the temperatures shown in Figure 4.3 and 4.4, the green curves are fits to a power law and the dot-dashed curves are fits to an exponential.

4.4 Discussion

In this chapter, we evaluated how the characteristic lengths of small-scale features in membranes of GUVs and GPMVs vary with temperature, using structural properties such as density autocorrelation functions and structure factors to analyze fluorescence images. We found that the length scales increase as temperature decreases and as tension increases (the tension study is not included in this thesis). Our results are partly consistent with both the Schick theory of microemulsions,[89–91,95] and with some predictions of the theory of Harden et al. for modulated phases.[28] For both morphologies, we observed exceptions in the structural, including inter-leaflet, properties of our vesicles. Discussion regarding these discrepancies is not directly relevant to the analysis presented in this work and was thus not included. Although it is also beyond the scope of this work, we propose that researchers could plausibly differentiate between modulated phases and microemulsions by collecting a series of images over time of a stationary vesicle exhibiting small domains and then could compute same structural properties presented in this work. If the domains were due to a microemulsion, then the averaged Fourier coefficients in $S(k)$ would approach zero even for those wave vectors at which the structure factor exhibits a peak. If the average of those Fourier coefficients were nonzero, then the domains would be due to a modulated phase, as presented extensively in Chapter 3 of this thesis, as well as in previous literature.[95] Independent of whether the vesicles that we investigate exhibit a modulated phase or a microemulsion, using the analysis presented above, we tested how the length scales in the system vary with physical parameters such as temperature. This helped us and our experimental collaborator to evaluate theoretical predictions associated with mechanisms of the formation of small domains.

4.5 Supplementary

Note: This code was published in the supplementary information of the paper by C. E. Cornell, A. D. Skinkle, S. He, I. Levental, K. R. Levental, and S. L.

Keller.[15]

4.5.1 Python Code for G(r) and S(k) Calculations

""" This script will take an 8-bit image and compute a pair correlation function, G(r), of the pixels in the image. First, a pair correlation is calculated along the x-direction. This generates a 2D pair correlation with correlations between particles in X and Y. The code then plots one slice of the G(r) in the x-direction against the radius.

The pair correlation function will be normalized around 0 and this code can also calculate an envelope function of the G(r). The code fits either an exponential or a power law function to the G(r) or envelope and gives error bounds. This version truncates the envelope function after the first several peaks of the G(r)

```
import matplotlib.pyplot as plt
import numpy as np
from scipy.misc import imread
from scipy.optimize import curve_fit

def calcrdf(phi):
        ''' This function calculates the radial distribution function
        of an image composed of an array of pixels. It returns a 2D
        G(r) integrated over the average contrast intensity '''

        ## Set up the size of the array
        Lx = phi.shape[0]
        Ly = phi.shape[1]

        ## Define field as contrast from average
```

```
        ## pixel intensity (can be negative)
        dphi = phi - np.average(phi)

        ## Calculate the average contrast intensity
        diphiavg2 = np.average(dphi**2)

        ## Define the array for a 2D G(r)
        rho2 = np.zeros((Lx+1, Ly+1))

        ## Calculate the 2D G(r)
        for dx in np.arange(Lx):
                for dy in np.arange(Ly):
                        phi2 = dphi
                        rho2[dx,dy] = np.average(phi2[0:Lx-dx,0:Ly-dy]
                         * phi2[dx:Lx,dy:Ly])

        return rho2/diphiavg2

def exponential(x, a, b, c):
        return a * np.exp(-b * x) + c

def power(x, a, b, c):
        return a * (x**(-b)) + c

def power2(x, a, b):
        return a * (x**(-b))

if __name__ == "__main__":
```

```
## Load in the image
im = imread("GUV_28 deg.tif")
phi = im

## Calculate the 2D G(r) using the RDF function
dgr_2d = calcrdf(phi)

## Slice the 2D G(r) in the x-direction, where
## dy=0 to obtain a 1D G(r)
dgr_1d = dgr_2d[0]

## Calculate the numerical derivative of G(r)
## to find the maxima to fit to an exponential
## or power law curve
gr_deriv = np.diff(dgr_1d)

## Roll the derivative over one to see where
## it changes sign
gr_derivposition = np.array(gr_deriv > 0, dtype=int)
gr_derivposition2 = np.roll(gr_derivposition, 1)
np.array(gr_derivposition - gr_derivposition2 == -1, dtype=int)
maxX = np.array(gr_derivposition-gr_derivposition2 == -1,
dtype=int)

## Define the array in units of pixels
## (will need the max number of pixels)
r = np.arange(80)

## Define the array in units of microns
```

```
## (will need the resolution of the image)
r = r*0.17

## Redefine r to shift by half a pixel length.
## This is due to an error in fitting a power law
## when there are 0's present in r. The
## uncertainty of the self correlation is on the
## order of the size of the pixel (0.17   m )
r = r+0.17/2

## Redefine the bounds of the 1D G(r)
g_r = dgr_1d[:80]

## For plotting, set the maxX at 0 equal to 1
## and find the x values where maxX is 1
## and the y values where x is 1 multiplied
## by the 1D G(r)
maxX[0] = 1
envelopeX = r[maxX==1]
envelopeY = maxX[maxX==1]*g_r[maxX==1]

## Truncate the envelope at ~6   m
## and delete erroneous point at 1 pixel noise
envelopeX = envelopeX[:7]
envelopeY = envelopeY[:7]
print(envelopeX)
envelopeX = np.delete(envelopeX,4)
envelopeY = np.delete(envelopeY,4)
```

```
## Fit the data to an exponential curve or power law
poptE, pcovE = curve_fit(exponential, envelopeX, envelopeY)#,
p0=[1,1],maxfev=10000)
fitE = exponential(envelopeX, *poptE)
poptP, pcovP = curve_fit(power2, envelopeX, envelopeY,
p0 =[1,1])
fitP = power2(envelopeX, *poptP)
#print(popt)
#print(pcov)

# Calculate the error based on a taylor expansion
# and the most weighted parameters (a and c)
perrE = np.sqrt(np.diag(pcovE))
upperE = fitE + (perrE[0]+perrE[2])
lowerE = fitE - (perrE[0]+perrE[2])

perrP = np.sqrt(np.diag(pcovP))
upperP = fitP + (perrP[0])
lowerP = fitP - (perrP[0])

# Calculate the residual and the variance to obtain
# the R-squared value for both fits

residualE = np.sum((envelopeY - fitE)**2)
varianceE = np.sum((envelopeY- envelopeY.mean())**2)
r_squaredE = 1 - (residualE/varianceE)
print(r_squaredE)

residualP = np.sum((envelopeY - fitP)**2)
```

```
varianceP = np.sum((envelopeY - envelopeY.mean())**2)
r_squaredP = 1 - (residualP/varianceP)
print(r_squaredP)

### Plot the G(r)

# Make a plot with space for subplots
fig, axes = plt.subplots(1,1,sharex=True, sharey=True,
figsize=(6,4))

# Plot the G(r) and the fit with error bounds
x = np.linspace(envelopeX.min(),envelopeX.max(),200)
axes.axhline(0, color = "black", linestyle ='--', alpha=0.2 )
axes.plot(r, g_r, color = "black", alpha = 0.7, label = 'data')
axes.plot(x, exponential(x,*poptE), color = "blue",
        alpha = 0.7, linestyle = '-.', label = 'exponential fit')
axes.plot(x, power2(x,*poptP), color = "green", alpha = 0.4,
        label = 'power law fit')
axes.plot(envelopeX,envelopeY, marker = "o", linestyle = "None",
markersize = 5, markeredgecolor = "grey",
        markerfacecolor = "None")
axes.set_ylabel("G(r)", fontname="Arial", fontsize=15)
axes.set_xlabel("Radius (  m )", fontname="Arial",
        fontsize=15)
axes.text(8, 0.85, '$R^2$ exponential = 0.999 '
        +'\n $R^2$ power = 0.998', style='normal',
        bbox={'facecolor':'None', 'alpha':0.5, 'pad':3})

# Save the figure as a PDF
```

```
plt.savefig("GUV_28deg_combined_truncate.pdf")
plt.show()

## Extra code to compute the S(k)

## Calculate the S(k) by taking a Fourier
## Transform of the magnitude squared of the
## 1D G(r)
rho_k = np.fft.fft(dgr_1d)
s_k = np.real(rho_k)**2 + np.imag(rho_k)**2


## Truncate the s_k to cut off the
## mirrored part of the plot
s_k = s_k[0:19]
```

Chapter 5

MEMBRANE BINDING OF SEVI AMYLOID PRECURSOR PEPTIDE PAP248-286: STRUCTURAL CHANGES AND COAGGREGATION

5.1 Introduction

It is well known that peptides can undergo large structural changes when binding to surfaces. For example, when exposed to rigid, hydrophobic surfaces, peptides containing periodic leucine and lysine sequences adopt a helical conformation.[12,58,116] Cellular and model membranes also present surfaces to which proteins and peptides bind. The intrinsic heterogeneity, flexibility, and dynamics of the peptide-membrane systems can lead to complicated binding mechanisms and new, emergent properties of the combined membrane/peptide system. Researchers have recently shown that the suppression of membrane undulations can result in an entropic, membrane-mediated attraction between proteins.[88]

$PAP_{248-286}$, a 39-residue fragment of prostatic acid phosphatase, is an abundant seminal peptide that has been proposed to play a variety of roles, both beneficial and deleterious to human health. PAP is intrinsically disordered in solution; it shows helical characteristics when bound to membranes and can also form β-sheet-rich amyloid fibrils. In its amyloid form, $PAP_{248-286}$ are usually referred to as Semen-derived Enhancer of Viral Infection (SEVI), and as a result of the presence of these amyloid fibrils formed from the self-assemby of $PAP_{248-286}$, which is abundant protein secreted into semen from the prostate.,[39] $PAP_{248-286}$ potently enhances infection by enveloped viral pathogens including HIV, herpes simplex virus and cytomegalovirus.[39,104] On the other hand, there is substantial evidence that PAP248-286 displays antimicrobial activity and participates in sperm quality control.[67]

Note that while $PAP_{248-286}$ effectively induces membrane leakage in vitro, its biologically relevant antimicrobial activity appears to derive from its ability to cause bacterial agglutination instead of direct cytotoxicity.[21,40] Our collaborators in the Nath Lab (UW Department of Medicinal Chemistry) have also observed that $PAP_{248-286}$ can inhibit biofilm formation by bacterial pathogens, an important virulence mechanism. A common feature of PAP's disparate activities is that they depend on this peptide's ability to bind biological membranes so as to alter cell-cell interactions as observed by our collaborators. Another published study also confirms this feature of $PAP_{248-286}$ and further investigates its interaction with nano-particles. **??**

$PAP_{248-286}$ and similar positively-charged, amyloid-forming or antimicrobial peptides can interact with biological membranes in diverse and complex ways. Peptides can cause vesicle leakage or fusion, or induce changes in membrane curvature.[84,104] Self-assembly is another important aspect of these interactions: some peptides can induce the aggregation or agglutination of lipid vesicles, while conversely membranes can induce amyloid fibril formation.[39,104] Importantly, there is a growing body of evidence that self-assembly phenomena, such as the encapsulation of target cells or pathogens, can be more relevant to biological function than classical modes of action such as membrane lysis.[45,52,66]

It has been recently observed by Nath and coworkers that $PAP_{248-286}$ forms peptide-lipid co-aggregates in a reversible, charge-dependent manner at particular peptide:lipid ratios (microcsopy images shown in Figure 5.1). While superficially similar co-aggregates have been observed for other peptides such as tau and α-synuclein,[34,87] the mechanism by which they form remains unknown. $PAP_{248-286}$ lipid co-aggregates, which we term "messicles", display some amyloid-like characteristics (thioflavin T fluorescence and β-sheet structure), but form much more quickly than bona fide $PAP_{248-286}$ amyloid and are incapable of seeding fibril formation. Messicles are heterogeneous aggregates that span a range of length scales from nm to μm (shown in the two center panels in Figure 5.1), and form in minutes at particular ratios of $PAP_{248-286}$ and negatively-charged liposomes at a POPC:POPG ratio

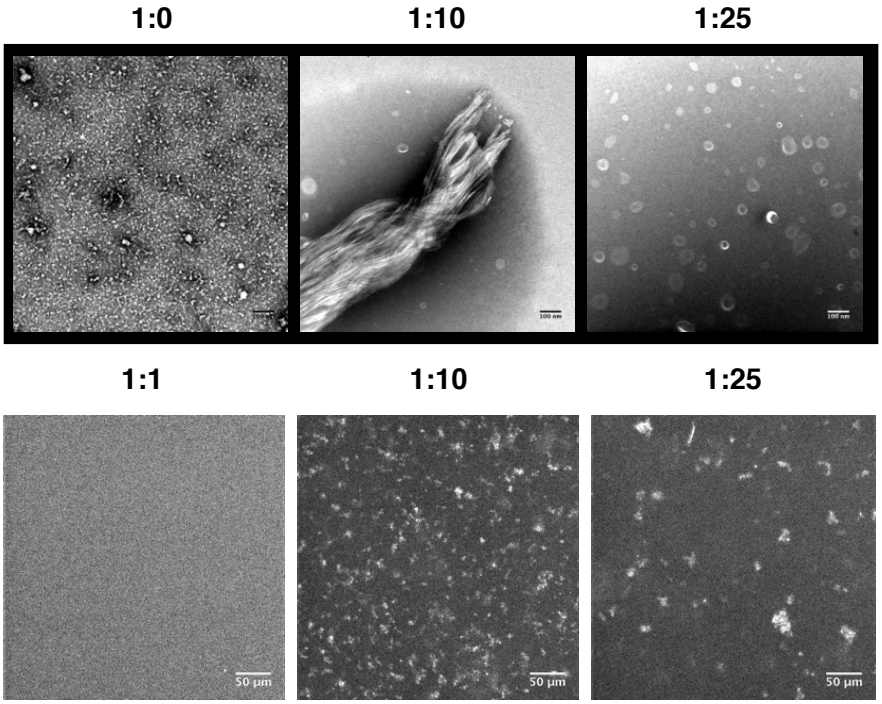

Figure 5.1: Experimental microscopy images of $PAP_{248-286}$ at different peptide to lipid ratios, ranging from peptide only (top left panel), peptide:lipid one to one (bottom left panel), peptide:lipid one to ten (two center panels), and peptide:lipid one to twenty five (two right panels). Messicles form at peptide:lipid ratios around 1:10, and span length scales ranging from nanometers (electron microscopy, top panels) to microns (light microscopy, bottom panels). In all systems, lipid mixture is composed of 30% POPC and 70% POPG. Scale bars are 100 nm for the top three electron microscopy images, and 50 μm for the bottom three light microscopy images, all taken by Vane and coworkers in the Nath Lab.

of 3:7. Addition of more peptide or lipid causes the dissociation of pre-formed messicles on similar timescales (shown in Figure 5.1). We hypothesize that messicles form when the density of surface-bound peptide passes a threshold that favors the formation of PAP248-286-mediated liposome-liposome contacts. In this work, we use our model to explore whether and how $PAP_{248-286}$ conformation is affected by crowding on a membrane. Specifically, we run atomistic simulations of SEVI at various concentrations on top of charged lipid bilayers and observe any resulting conformational changes to the peptide as well as to the membrane.

5.2 Methods

5.2.1 Molecular Dynamics Simulation

Force Field

In this work, we used atomistic Molecular Dynamics simulations to capture the molecular detail of the SEVI-membrane interaction. Our goal is to investigate the binding mechanism, as well as the conditions and implications of messicles formation. In preliminary studies we first confirmed that the CHARMM36m[37] out-performs other force fields for atomistic SEVI simulation in that the peptide does not over-express α-helical structure in solution. We ran long (microseconds) equilibrations of SEVI in solution using multiple force fields, starting from membrane bound SEVI configuration obtained from the Protein Data Base (PDB). From experimental results, we expected the helical membrane bound structure to unfold and become disordered when $PAP_{248-286}$ was equilibrated in solution, and we confirmed that the CHARMM36m force field reproduces that behavior.

Simulation Parameters, Setup, and Membrane Compositions

First we constructed the initial configuration of $PAP_{248-286}$ from the Protein Data Base using the NMR structure of $PAP_{248-286}$ in a membrane environment.[69] We picked this particular

configuration due to a lack of reliable $PAP_{248-286}$ configurations in a solution environment comparable to that adopted by our collaborators. To ensure accurate sampling of the structural change of $PAP_{248-286}$ as it approaches the membrane, and to confirm that consistent secondary structures were obtained compared to experimental results, we equilibrated the configuration in long ($> 5\,\mu$s) free $PAP_{248-286}$ simulation in KCl solution at the same ionic strength (150 mM) used by our experimental collaborators.

For simulations of $PAP_{248-286}$ on membranes, our choice of lipid mixture was guided by our experimental collaborators in the Nath Lab. We built symmetric lipid bilayers that consist of neutral POPC and negatively charged POPG molecules, at 1) 100% POPC, 2) 50%POPC/50% POPG, and 3) 30% POPC/70% POPG. We set up bilayer simulations at these lipid compositions using the Charmm-GUI membrane builder,[53,117] where the distribution of each type of the lipids was randomized. We then combined those lipids with the previously equilibrated solution $PAP_{248-286}$ structure by inserting and solvating the $PAP_{248-286}$ $\sim$ 5 nm on top of the bilayer. This step was facilitated by the Charmm-GUI web program. Then the system was equilibrated using similar conditions as in the solution. From our preliminary results, we found that the $PAP_{248-286}$ membrane systems tend to require long simulation times to equilibrate and explore configuration space ($> 10\,\mu$s). It was also noted that, upon approaching and binding to some of the membranes, $PAP_{248-286}$ tend to get stuck in configurations determined by the initial binding process. To obtain optimal sampling of the secondary structure, we developed a scheme to run 20 independent pre-equilibrated production simulations for each charged membrane, where $PAP_{248-286}$ was initially inserted sufficiently far away (3 - 5 nm) from the membrane and was allowed to approach the membrane independently throughout the course of the simulation. The results were then combined and analyzed to provide statistical insight into the configurational space of the peptide upon binding.

All simulations involved in this work were conducted using 2 fs time step, which is typical for atomistic simulations using CHARMM force fields.[37] Van der Waals and electrostatic

interactions were truncated at 1.2 nm, with a smooth decay of the former starting at 0.9 nm. Temperature was controlled by a stochastic velocity rescaling thermostat.[14,109] Pressure was controlled by a semi-isotropic Berendsen barostat[10] with a reference pressure of 1 bar, which effectively maintains zero surface tension for the membrane. All simulations were performed using the GROMACS software package (version 4.6.5).[10,78]

5.2.2 Peptide-Lipid Contact

To quantify the level of interaction between the peptide and the lipid molecules on the membrane, we defined and calculated the peptide-lipid contact using the definition that a peptide-lipid contact is counted if the closest distance between any part of a lipid and any part of the peptide is within 1 nm. One lipid molecule can at maximum contribute 1 peptide-lipid contact. This contact is then checked between the peptide and all the lipid molecules in the system at any given instance of the simulation. Finally the average peptide-lipid contact is calculated by average the intantaneous peptide-lipid contacts over the entire course of the equilibrated simulation trajectory.

5.2.3 Peptide Secondary Structure

For identification and characterization of the peptide secondary structure, we adopt the DSSP database and definition.[42,43] The DSSP database and program is based on a pattern recognition process of hydrogen bonded and geometrical features extracted from X-ray coordinates. The output is based on a compilation of the primary structure, including sulfur-sulfur bonds, secondary structure, and solvent exposure. The algorithm recognizes repeats of the elementary hydrogen bonding patterns "turn" and "bridge", based upon which the cooperative secondary structure is determined. Repeating turns are defined as "helices", repeating bridges are recognized as "ladders", and connected ladders are "sheets". Geometric structure is defined in terms from differential geometry of torsion and curvature. Local

chain chirality is the torsional handedness of four consecutive alpha carbon positions and is positive for right handed helices and negative for ideal twisted β-sheets. Curved pieces are defined as "bends". Solvent exposure is given as the number of water molecules in possible contact with a residue. To simplify our calculation while simultaneously capturing all the structural features that are important to our analysis, we identify and characterize four main categories of secondary structures: 1) random coil, 2) β-sheet, 3) bend and turn, and 4) helix, where, under rigorous DSSP definition, we combine both β-sheet and β-bulge in to the second category, both bend and turn into the third category, and 3-helix, α-helix, and $5/\pi$-helix into the fourth category.

5.3 Results

5.3.1 $PAP_{248-286}$ in Solution and on Membranes

In order to validate the initial configuration of the peptide against experimental results, we first simulated $PAP_{248-286}$ in 15 mM KCl solution (shown in the top left panel in Figure 5.2). We then simulated $PAP_{248-286}$ on top of lipid bilayers with various POPG compositions (after the peptide was equilibrated in solution with no apparent excessive helical structure compared to experimental results). We find that binding behavior and secondary structure of the peptide depend on the mole fraction of negatively charged POPG lipids in the membrane. We observe the formation of extended structure as the peptide binds to the negatively charged bilayer, whereas no apparent strong binding or secondary structure change is observed for $PAP_{248-286}$ on a neutral membrane (bottom left panel in Figure 5.2).

From our membrane simulations with POPG presence (two right panels in Figure 5.2), we observe rapid binding of the peptide to the bilayer and subsequent conformational changes, including the formation of extended beta sheet structure in the peptid. These results indicate that the simulations capture the experimentally observed behavior. In all of the peptide on membrane systems, neutral and charged, the membranes remain roughly flat and no

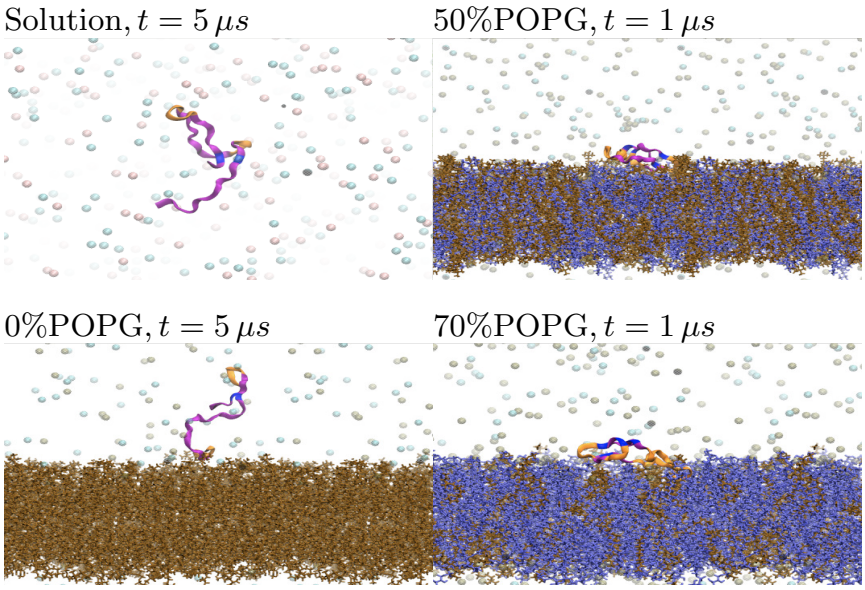

Figure 5.2: Simulation snapshots of $PAP_{248-286}$ systems taken at 1μs or 5μs after the initial temperature quench. $PAP_{248-286}$ is shown in solution (top left), on a pure neutral POPC bilayer with 100% POPC (bottom left), on a bilayer with 50% POPC and 50% POPG (top right), and on a bilayer with 30% POPC and 70% POPG (bottom right). POPC (gold) and POPG (lavender) are color coded, as is the peptide by its secondary structure at the time when the snapshot was taken, with random coil (purple), beta sheet (orange), bend and (blue), and helix (red). In an abnsence of POPG, the peptide exhibits extended structure, and random coil (purple) appears to be the prominent configuration. When the bilayers contains POPG, the peptide binds more tightly and extended beta sheet structure is observed.

significant deformation is observed upon interacting with the peptide. In the next subsection, I will examine if the peptide causes any lateral lipid re-organization.

5.3.2 Peptide lipid contacts does not indicate non-ideal mixing or recruitment of POPG molecules by $PAP_{248-286}$

To determine if the peptide binding causes any lipid rearrangement, especially POPG rearrangement, around the peptide due to electrostatic interaction, we calculated the peptide-lipid contacts using a cutoff value r_c of 1 nm. For each of the three peptide-membrane systems, we tabulated (Table 5.1) the average number of contacts between the peptide and lipids of any type, the average number of contacts between the peptide and POPG lipids, and the percent of all contacts that were with POPG. As expected based on the simulation snapshot, the POPC (0% POPG) membrane shows few contacts with the peptide (4.39 on average) with large fluctuations (5.54). This suggests that the peptide is barely touching the membranes that are neutral (without any POPG). For the two peptide-membrane systems containing POPG, we observe that the average number of peptide-lipid contacts are much higher (25.46 *pm* 2.47 and 26.02 $\pm$ 3.81, respectively), indicating a tighter binding between the peptide and the membrane surface. These two numbersare within the simulation variance of each other, which suggests that the 50% POPG membrane provides attractive interactions that are strong enough for the peptide to saturate its entire possible contact surface with the lipid molecules, without significantly deforming the membrane or the peptide. Because the variances for peptide-lipid contacts for POPG-containing membranes are smaller than those for the neutral membrane, the peptide structure does not fluctuate as much on top of these negatively charged membranes, which again suggests tighter binding.

By examining the percentage of all contacts that were with POPG lipids, we find that there is a small enrichment of POPG molecules (56.5% $\pm$ 8.6%) around the peptide compared to the global POPG composition of 50%, which suggests that the peptide recruits a small portion of POPG molecules via electrostatic interactions. The enrichment is also observed for the

	0% POPG	50% POPG	70% POPG
Average peptide-lipid contact	4.39 ± 5.54	25.46 ± 2.47	26.02 ± 3.81
Average peptide-POPG contact	0	14.38 ± 1.69	19.42 ± 3.00
POPG contact%	0%	56.5% ± 8.6%	74.6% ± 15.9%

Table 5.1: Peptide-lipid contacts and peptide-POPG lipid contacts calculated using a cutoff value r_c of 1 nm. The percent of all contacts that were between the peptide and POPG (POPG contact% in the thrid row) are calculated by dividing the second row by the first row. In the first column we notice that there is no significant contact between $PAP_{248-286}$ and the lipids in a neutral POPC membrane. On the two right columns the average contacts are 25.46 for 50% POPG and 26.02 for 70% POPG, respectively.

peptide on a 70% POPG membrane. However the small magnitude of this enrichment of POPG compared to the uncertainties suggests that the degree of spatial lipid reorganization by the peptide via electrostatic interaction is not significant. No evidence of non-ideal mixing or aggregation of PG lipids is found.

5.3.3 Peptide Secondary Structure

We calculated the secondary structure for $PAP_{248-286}$ in all systems and environments using the DSSP database discussed in section 5.2.3. The initial DSSP characterizations for every residue on the peptide are shown in the supplementary figures 5.5 - 5.9. Visually it is clear that $PAP_{248-286}$ does not adopt an extended structure in solution or on a neutral POPC membrane. In contrast, on membranes of 50% POPG and 70% POPG, $PAP_{248-286}$ adopts a long-lasting helical and beta sheet structure (α-helix at residues 10 - 15, and β-sheet near residues 2 and 25), as shown in supplementary Figures 5.7 and 5.8. However, as we discussed in section 5.2.1 of this work, for $PAP_{248-286}$ on charged POPG membranes, the peptide gets stuck in its configuration space for microseconds time scale upon binding to the bilayer. This poetentially leads to persistent and drastically different protein structural conformations. Compare secondary structure for two independent systesm with $PAP_{248-286}$

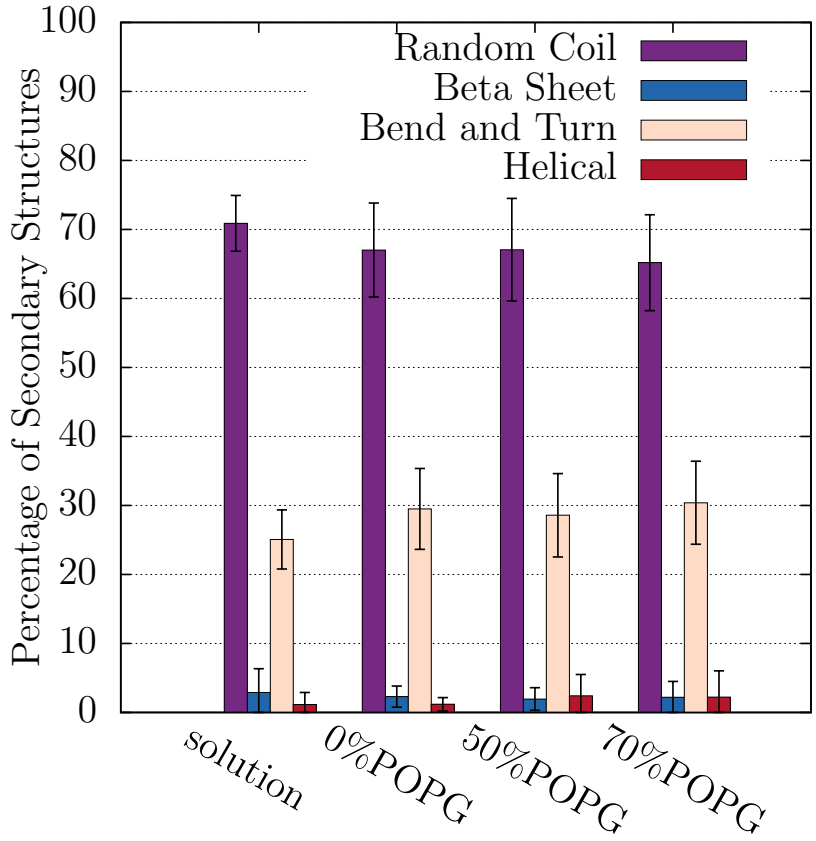

Figure 5.3: Percentages of each of the four categories of the $PAP_{248-286}$ peptide's secondary structures, including random coil (purple), beta sheet (blue), bend and turn (pink), and helix (red) averaged over time for different membrane systems.

on 70% POPG membrane shown in Figure 5.8 and 5.9. In 5.8 we observe persistent helical structure between residues 10 and 15 for μs time scale, whereas in In 5.9 no such conformation can be observed for the first 2 μs.

In order to sample peptide configuration efficiently and to find the complete configurational space of the peptide, we analyzed peptide secondary structure for the 20 independent simulations for each membrane composition. We then identified the peptide secondary structure from these trajectories and calculated the percent of residues exhibiting characteristics of one of the four major categories of secondary structures (Figure 5.3). We observe no significant

difference in the percent secondary structure for the peptide in the four environments. However, this result is likely caused by $PAP_{248-286}$ being stuck at various different conformations in large portion of the simulations upon binding with the charged membranes. This can be indicated by the relatively large variances in the percent residues in helical conformation on peptides in charged membrane environments, shown in Figure 5.3. When residue conformations are calculated form $PAP_{248-286}$ that are stuck in these local environments, it is probably that important structural information on the residues get averaged out over the entire peptide, and over all the trajectories (20 trajectories for each of the two systems with charged membranes).

In order to investigate structural features of the peptide beyond the percent of each structure for entire molecule, we then calculated the percet of time every peptide residue spends in one of the four secondary structures individually over the equilibrated portion of all trajectories. Complete results for this analysis are shown in supplementary section 4.5, Figure 5.10 - 5.13. We observe no significant variation in the percent time each residue spends in random coil or bent and turn conformations from simulations of $PAP_{248-286}$ in four different environments. The error bars for these two conformations are large compared to the percent time value, and for the rest of this chapter we will not focus on the random coil or bend and turn conformations. Instead we investigate the percent time each residue spend in the helical conformation in solution, pure POPC, 50% POPG, and 70% POPG systems. In Figure 5.4, we found that residues 11 to 16 show significant difference in helical characteristics between charged membrane environment (both 50% POPG and 70% POPG) and neutral environment (both solution and membrane containing 100% POPC). To further investigate the determining factor of peptide structural change upon binding, we calculated p-values using the residue-specific secondary structure percentages. From these values, we identify key residues that define peptide secondary structure. For example, the p-values for helical structure shown in Table 5.2, we observe extended helical structure from residue 11-16 with $\pm$ 1 residue variation (shown in Table 5.2) The rest of the p-value test results are included

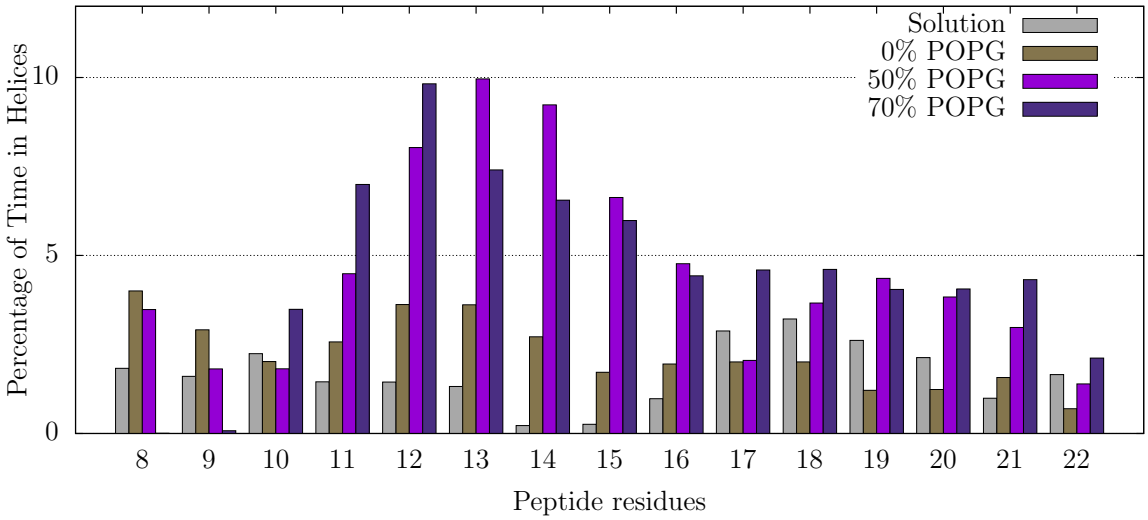

Figure 5.4: Average percent time that each of the 15 residues from residue 8 to residue 22 of $PAP_{248-286}$ spends in helical conformation for the peptide in solution (silver), on a membrane with 100% POPC (gold), on a membrane containing 50% POPGv(light purple), and on a membrane containing 70% POPG.

P-values	11	12	13	14	15	16	17	18	19
Solution POPC	0.39	0.30	0.27	0.27	0.24	0.83	0.26	0.94	0.85
Solution 50%POPG	0.03	0.02	0.03	0.01	0.03	0.20	0.52	0.58	0.93
Solution 70%POPG	0.22	0.02	0.015	0.014	7e-3	4e-12	3e-11	0.95	0.16
POPC 50%POPG	0.07	0.04	0.025	0.026	0.76	0.70	0.84	0.69	0.93
POPC 70%POPG	0.53	0.14	0.02	0.03	0.03	8e-12	0.03	0.55	3e-06
50%POPG 70%POPG	0.51	0.36	0.57	0.37	0.32	0.96	0.96	0.44	0.22

Table 5.2: P-values comparing different environments in the peptide's helical structure for residues 11 - 19. A small p-value (< 0.05 indicates that helical characteristics of the given residue is significantly different in the two environments involved.) The highlighted entries suggest that persistent helical structure is identified for peptide on bilayers with charged POPG lipids when compared to the solution and neutral membrane structure.

in the supplementary section of this chapter.

5.4 Discussion

In the atomistic simulations presented in this Chapter, $PAP_{248-286}$ on negatively charged POPC:POPG bilayers behaves in a manner similar to that reported by our experimental collaborators in the Nath Lab. Specifically: (1) the solution structure of $PAP_{248-286}$ is mostly a disordered, random coil (2) $PAP_{248-286}$ on neutral membraneof pure POPC shows no evidence of peptide-membrane binding or significant structural change compared to the peptide in solution (3) strong membrane binding is observed for $PAP_{248-286}$ on charged bilayers containing POPG, although no significant clustering or recruitment of POPG molecules around the peptide is observed, and (4) via residue-specific secondary structure characterization, we find statistically significant evidence that persistent helical structure forms in systems with negatively charged POPG in the bilayer, supporting our initial hypothesis that $PAP_{248-286}$ binds to membranes and forms co-aggregates by extended structure, an amyloid-like characteristic.

Our work presents new features and perspectives of the problem. To overcome the computational resource barrier and the slow sampling of the peptide configurational space, we performed multiple independent MD simulations to model the peptide-membrane binding process. This allows us to optimize sampling and utilize the simulation data to yield statisitical insight on the behavior of each residue during the binding process and in different binding environments. This is a big leap towards understanding the binding mechanism of $PAP_{248-286}$ melecules on charged bilayers when combined with the knowledge of the sequence of the peptide. A future challenge is to simulate multiple $PAP_{248-286}$ molecules on a much larger membane surface in order to probe the mechanism by which $PAP_{248-286}$ peptides aggregate. To fully elucidate all details of the mechanism by which $PAP_{248-286}$ binds to membrane and forms aggregates, super-computer facilitated computational studies will be required to simulate and sample the structural properties of co-aggregate formation in-

volving multiple $PAP_{248-286}$ molecules. The results presented in this work contribute to this long-term goal by using statistical approaches to accurately determine the residue-specific structural characteristics of the binding process with a single $PAP_{248-286}$ peptide.

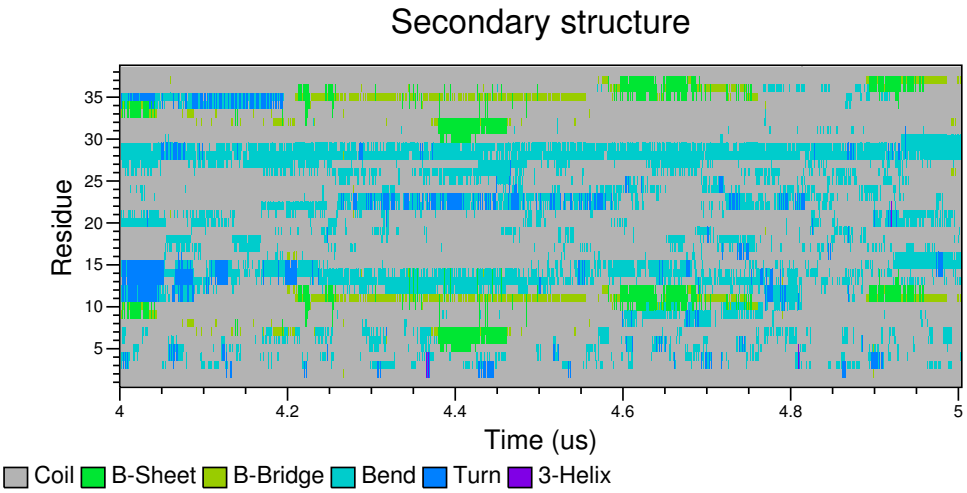

Figure 5.5: Secondary structure calculation for the last 1 μs of a simulation of $PAP_{248-286}$ in solution. Vertical axis spans the first to the thirty-ninth residues of $PAP_{248-286}$, and the horizontal axis is simulation time in μs. Different colors dsignate different secondary structure adopted by each residue: 1) "coil" is randome coil 2) "B-sheet" and "B-bridge" are for *beta*-sheet in our definition 3) "Bend" and "Turn" are combined in our bend and turn category, and 4) "3-Helix" is the helical structure.

5.5 Supplementary

5.5.1

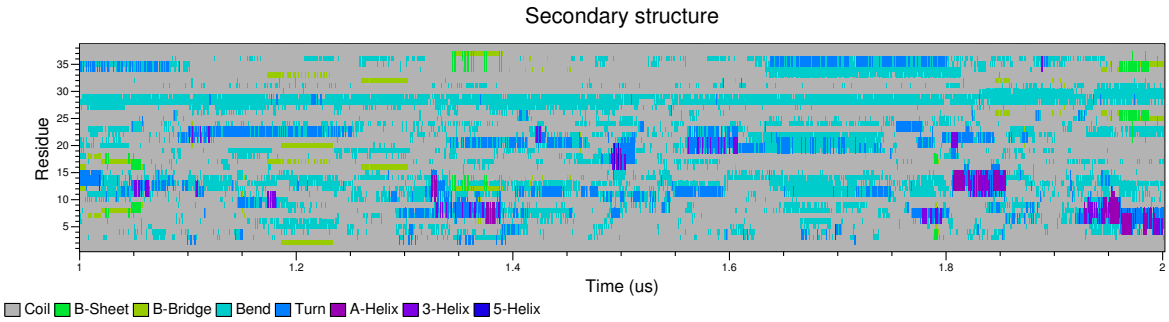

Figure 5.6: Secondary structure calculation for the last 1 μs of a simulation of $PAP_{248-286}$ on POPC membrane simulation. Vertical axis spans the first to the thirty-ninth residues of $PAP_{248-286}$, and the horizontal axis is simulation time in μs. Different colors dsignate different secondary structure adopted by each residue: 1) "coil" is randome coil 2) "B-sheet" and "B-bridge" are for *beta*-sheet in our definition 3) "Bend" and "Turn" are combined in our bend and turn category, and 4) "A-Helix", "3-Helix", and "5-Helix" together make up the helical structure category.

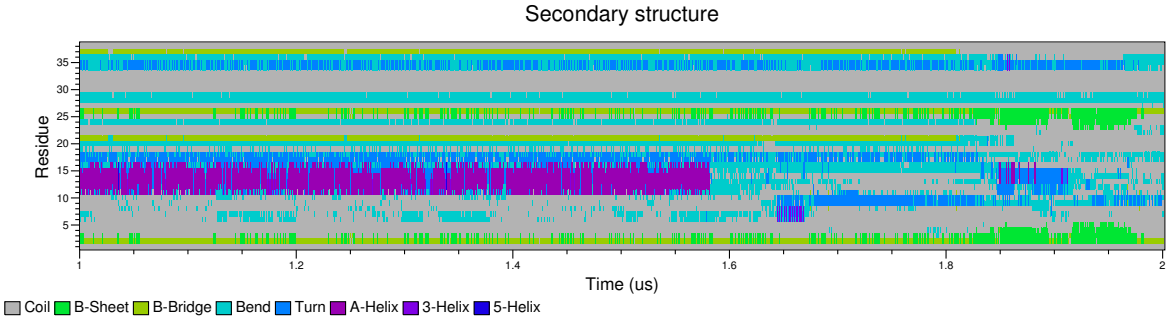

Figure 5.7: Secondary structure calculation for the last 1 μs of a simulation of $PAP_{248-286}$ on 50% POPG membrane simulation. Vertical axis spans the first to the thirty-ninth residues of $PAP_{248-286}$, and the horizontal axis is simulation time in μs. Different colors dsignate different secondary structure adopted by each residue: 1) "coil" is randome coil 2) "B-sheet" and "B-bridge" are for *beta*-sheet in our definition 3) "Bend" and "Turn" are combined in our bend and turn category, and 4) "A-Helix", "3-Helix", and "5-Helix" together make up the helical structure category.

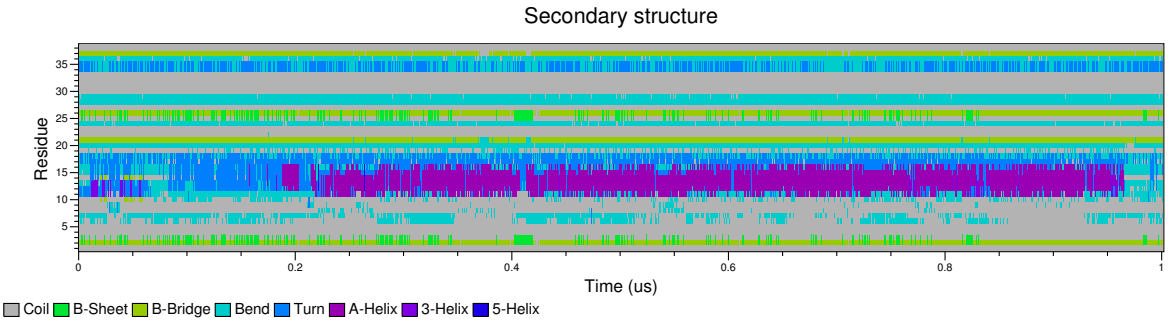

Figure 5.8: Secondary structure calculation for a 1 μs of a simulation of $PAP_{248-286}$ on 70% POPG membrane simulation, in which the peptide shows persitent helical structure. Vertical axis spans the first to the thirty-ninth residues of $PAP_{248-286}$, and the horizontal axis is simulation time in μs. Different colors dsignate different secondary structure adopted by each residue: 1) "coil" is randome coil 2) "B-sheet" and "B-bridge" are for *beta*-sheet in our definition 3) "Bend" and "Turn" are combined in our bend and turn category, and 4) "A-Helix", "3-Helix", and "5-Helix" together make up the helical structure category.

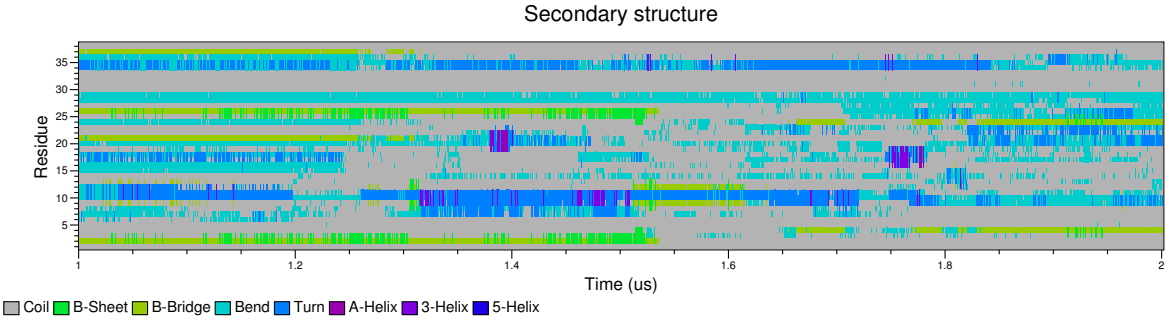

Figure 5.9: Secondary structure calculation for the last 1 μs of a simulation of $PAP_{248-286}$ on 70% POPG membrane simulation, in which the peptide shows random coil characteristics. Vertical axis spans the first to the thirty-ninth residues of $PAP_{248-286}$, and the horizontal axis is simulation time in μs. Different colors dsignate different secondary structure adopted by each residue: 1) "coil" is randome coil 2) "B-sheet" and "B-bridge" are for *beta*-sheet in our definition 3) "Bend" and "Turn" are combined in our bend and turn category, and 4) "A-Helix", "3-Helix", and "5-Helix" together make up the helical structure category.

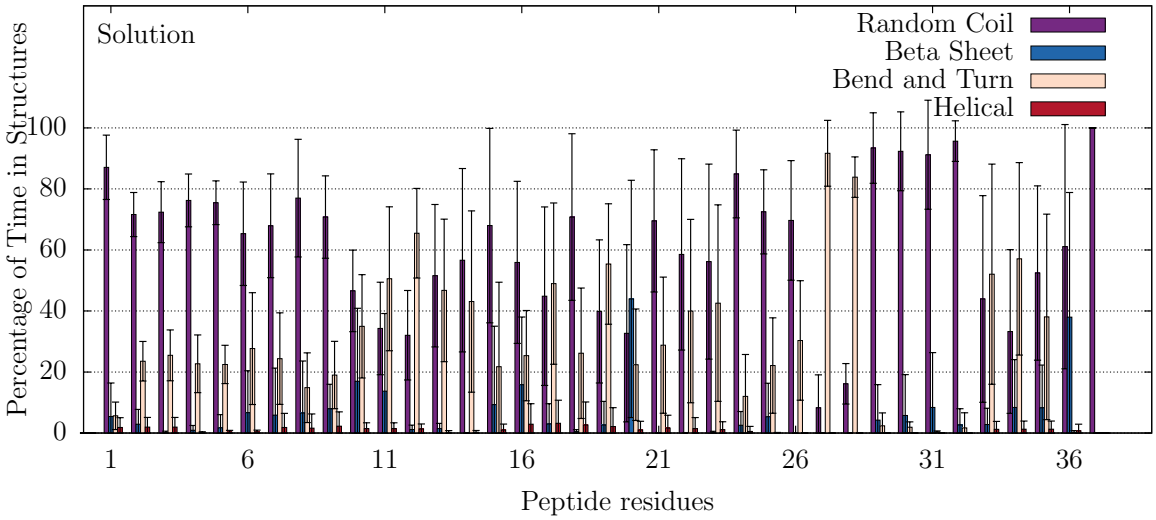

Figure 5.10: Average percent time that each of the 39 residues of $PAP_{248-286}$ (in solution) spends in one of four conformations, including random coil (purple), β-sheet (blue), bend and turn (pink), and helix (red). Results were calculated from 8 non-overlapipng sections of equilibrated portion (last 4 μs) of the trajectory (0.5 μs each).

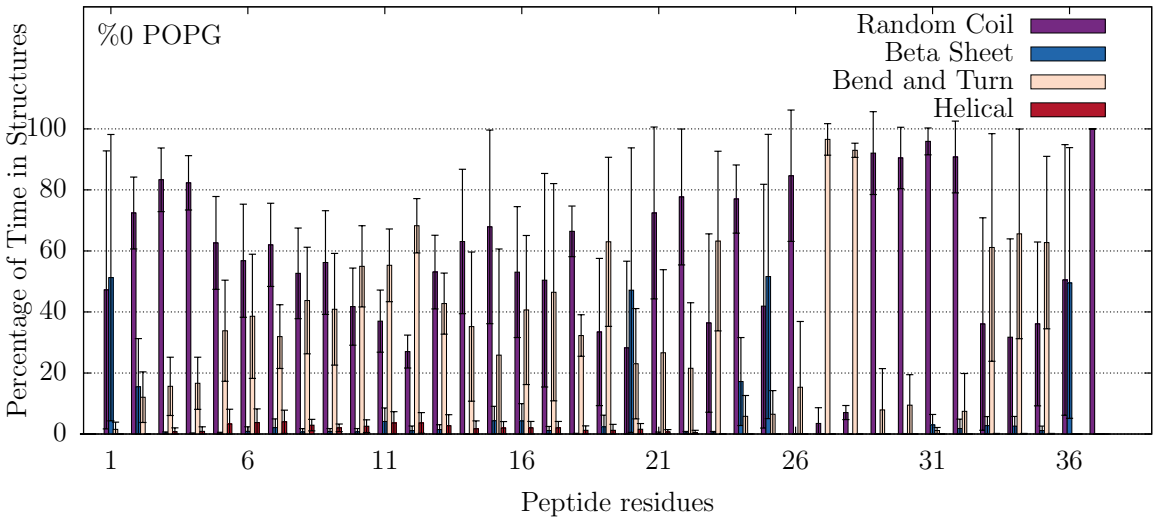

Figure 5.11: Average percent time that each of the 39 residues of $PAP_{248-286}$ (in pure POPC membrane environment) spends in one of four conformations, including random coil (purple), β-sheet (blue), bend and turn (pink), and helix (red). Results were calculated from 8 non-overlapipng sections of equilibrated portion (last 4 μs) of the trajectory (0.5 μs each).

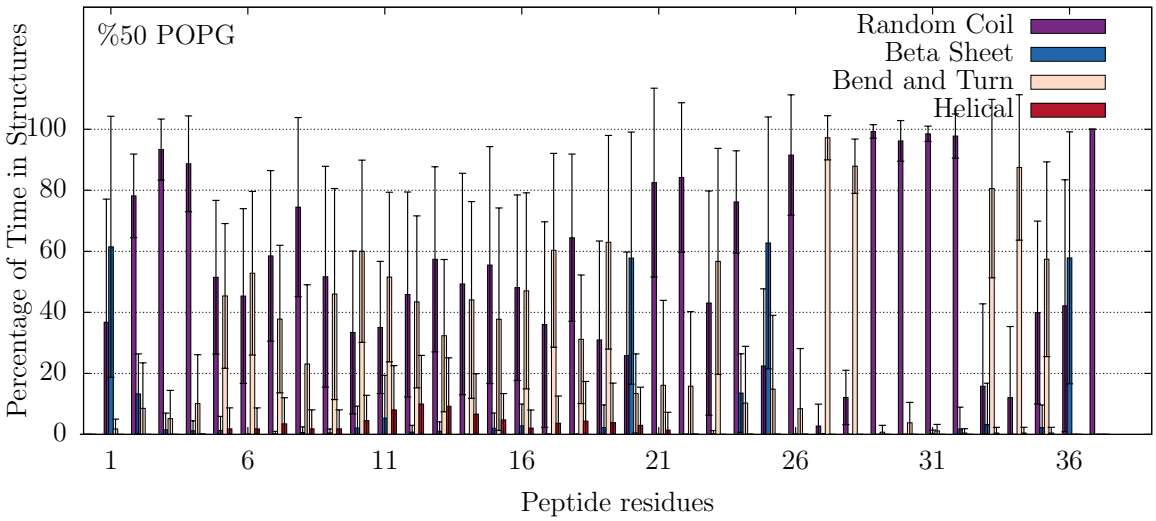

Figure 5.12: Average percent time that each of the 39 residues of $PAP_{248-286}$ (in 50% POPG membrane environment) spends in one of four conformations, including random coil (purple), β-sheet (blue), bend and turn (pink), and helix (red). Results were calculated from of equilibrated portion (last 0.5 μs) of the 20 independent trajectories with a membrane containing 50% POPG.

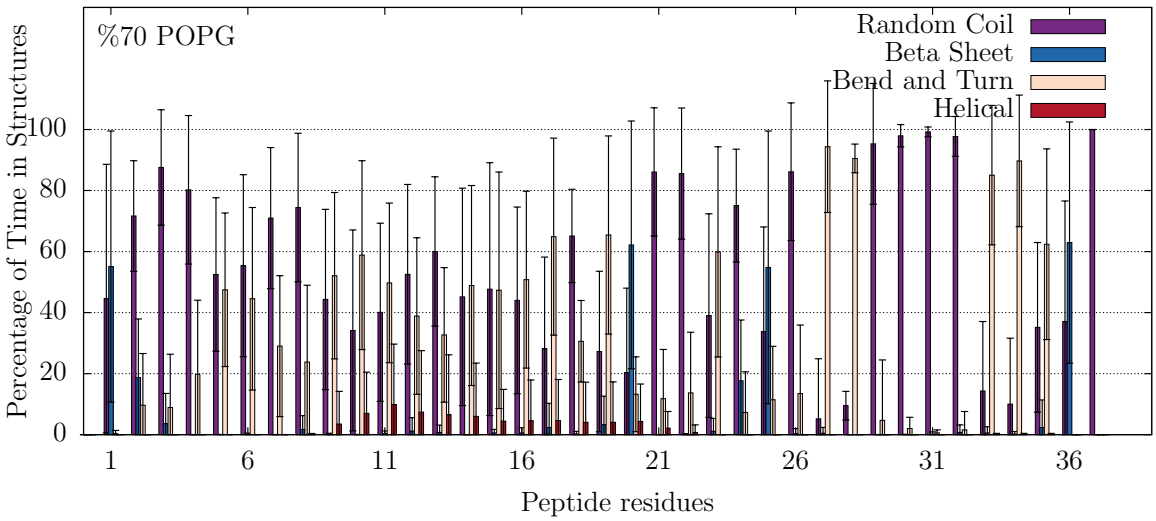

Figure 5.13: Average percent time that each of the 39 residues of $PAP_{248-286}$ (in 70% POPG membrane environment) spends in one of four conformations, including random coil (purple), β-sheet (blue), bend and turn (pink), and helix (red). Results were calculated from of equilibrated portion (last 0.5 μs) of the 20 independent trajectories with a membrane containing 70% POPG.

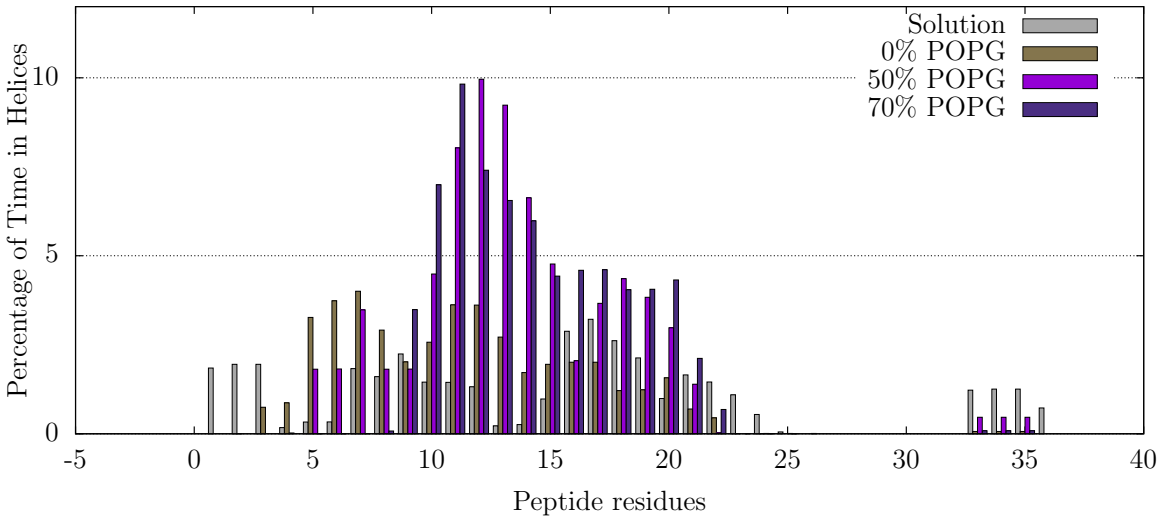

Figure 5.14: Average percent time that each of the 39 residues of $PAP_{248-286}$ spends in helical conformation for the peptide in solution (silver), on a membrane with 100% POPC (gold), on a membrane containing 50% POPGv(light purple), and on a membrane containing 70% POPG.

Chapter 6

SUMMARY AND FUTURE DIRECTIONS

In Chapter 2, our continuum and discretiized model for strongly seggregated systems describe the key structural features of spatial heterogeneity within a two-dimensional membrane systems, which allows us to identify the signature properties of membranes with large, micron-sized domains or with small, nano-scale domains.

In Chapter 3, using the tools we presented in Chapter 2, as well as statistical methods such as Gaussian cluster model and Bayesian information criterion, we found that the degree of heterogeneity within multi-component bilayers depends on the global composition of multiply-unsaturated phospholipids. We also produced multi-component phase diagram for an array of four-component lipid bilayers systems that were previously studied by many experimental researchers. More importantly, we developed a statistical method that can be used to detect the onset of phase separation of any simulated model membrane system. Some huge advantages of our method are that 1) it does not rely on pre-defined L_d or L_o identities of lipids, and that 2) it can still provide insight when the simulated system is not big enough for researchers to distinguish phase separation from large length scale density fluctuations, which prompts the researchers to realize that the current system is size-limited, and that a larger system should be simulated. Furthermore, the results presented in this work contribute to the quest to identify membrane inhomogeneities on small length scales by using novel approaches to accurately determine the boundaries of the phase coexistence region in quaternary lipid bilayers.

In Chapter 4, we presented a project conducted in collaboration with the Keller and Levental

labs. We investigated heterogeneous vesicle membranes with characteristic length scales to test predictions arising from various proposed mechanisms of how modulated phases and microemulsions arise in membranes. We found that the characteristic length scales of both the GUVs and GPMVs increase as temperature decreases and as tension increases. Our approach and results will provide insight into future experimental and computational studies to evaluate the theoretical predictions associated with mechanisms of the formation of small domains.

In Chapter 5, we showcased a toolkit that we developed to examine and characterize the secondary structure of a small, intrinsically disordered, charged peptide that binds to a membrane. Our goal was to understand the mechanism behind the membrane binding and structural change observed by our experimental collaborators, and to cast light on the mechanism of co-aggregate formation observed by our collaborators. We performed atomistic simulations of $PAP_{248-286}$ on 70:30 POPG:POPC bilayers. We quantified changes in secondary structure and interaction patterns of the peptides. In order to investigate the mechanism of the formation of messicles on the anionic bilayer, as observed by our collaborators, future super-computer facilitated computational studies will be required. A logical next step would be to simulate and sample the structural properties of co-aggregate formation involving multiple $PAP_{248-286}$ molecules, for which the structural characterization analysis for single $PAP_{248-286}$ described in this work will be immensely helpful.

Overall, we contributed to the field of computational biophysics by developing a set of noval statistical methods to systematically study the length scale of membrane inhomogeneity, the possible mechanisms controlling these characteristic length scales, and mechanisms by which membrane systems interact with physiologically relevant small peptide molecules. These methods provide new insight and perspetives into many long-standing open questions, as well as noval and uncharted territories in the field of biophysics.

BIBLIOGRAPHY

[1] David G. Ackerman and Gerald W. Feigenson. Multiscale Modeling of Four-Component Lipid Mixtures: Domain Composition, Size, Alignment, and Properties of the Phase Interface. *J. Phys. Chem. B*, 119:4240–4250, 2015.

[2] Jonathan J. Amazon, Shih Lin Goh, and Gerald W. Feigenson. Competition between line tension and curvature stabilizes modulated phase patterns on the surface of giant unilamellar vesicles: A simulation study. *Phys. Rev. E.*, 022708:1–10, 2013.

[3] R. G. W. Anderson and K. Jacobson. A Role for Lipid Shells in Targeting Proteins to Caveolae, Rafts, and Other Lipid Domains. *Science*, 296(5574):1821–1825, 2002.

[4] Jun Ando, Masanao Kinoshita, Jin Cui, Hiroyuki Yamakoshi, Kosuke Dodo, and Katsumasa Fujita. Sphingomyelin Distribution in Lipid Rafts of Artificial Monolayer Membranes Visualized by Raman Microscopy. *Proc. Natl. Acad. Sci. U. S. A.*, 112(15):4558–4563, 2015.

[5] Clare L. Armstrong, Drew Marquardt, Hannah Dies, Norbert Kučerka, Zahra Yamani, Thad A. Harroun, John Katsaras, An-Chang Shi, and Maikel C. Rheinstädter. The observation of highly ordered domains in membranes with cholesterol. *PLOS ONE*, 8(6):1–10, 06 2013.

[6] Svetlana Baoukina, Eduardo Mendez-Villuendas, W. F. Drew Bennett, and D. Peter Tieleman. Computer Simulations of the Phase Separation in Model Membranes. *Faraday Discuss.*, 161:63–75, 2013.

[7] H. E. R. Baughman, A. F. Clouser, R. E. Klevit, and A. Nath. HspB1 and Hsc70 chaperones engage distinct tau species and have different inhibitory effects on amyloid formation. *J. Biol. Chem.*, 293:2687–2700, 2018.

[8] T. Baumgart, S. T. Hess, and W. W. Webb. Imaging coexisting fluid domains in biomembrane models coupling curvature and line tension. *Nature*, 425:821–824, 2003.

[9] Tobias Baumgart, Adam T Hammond, Prabuddha Sengupta, Samuel T Hess, David A Holowka, Barbara A Baird, and Watt W Webb. Large-Scale Fluid/Fluid Phase Separation of Proteins and Lipids in Giant Plasma Membrane Vesicles. *Proc. Natl. Acad. Sci. U. S. A.*, 104(9):3165–3170, 2007.

[10] H. J C Berendsen, D. van der Spoel, and R. van Drunen. GROMACS: A Message-

Passing Parallel Molecular Dynamics Implementation. *Comput. Phys. Commun.*, 91(1-3):43–56, 1995.

[11] K. Binder. Critical properties from monte carlo coarse graining and renormalization. *Phys. Rev. Lett.*, 47:693–696, 1981.

[12] N. F. Breen, T. Weidner, K. Li, D. G. Castner, and G. P. Drobny. A solid-state deuterium NMR and SFG study of the side chain dynamics of peptides adsorbed onto surfaces. *J. Am. Chem. Soc.*, 131(40):14148–14149, 2009.

[13] A. D. Bruce. Probability density functions for collective coordinates in Ising-like systems. *J. Phys. C*, 14:3667–3688, 1981.

[14] Giovanni Bussi, Davide Donadio, and Michele Parrinello. Canonical sampling through velocity-rescaling. *J. Chem. Phys.*, 126(1):014101, 2007.

[15] Caitlin E. Cornell, Allison D. Skinkle, Shushan He, Ilya Levental, Kandice R. Levental, and Sarah L. Keller. Tuning length scales of small domains in cell-derived membranes and synthetic model membranes. *Biophys. J.*, 115:690–701, 2018.

[16] Subir K. Das, Jrgen Horbach, and Kurt Binder. Transport phenomena and microscopic structure in partially miscible binary fluids: A simulation study of the symmetrical lennard-jones mixture. *J. Chem. Phys.*, 119(3):1547–1558, 2003.

[17] Ryan S. Davis, P. B. Sunil Kumar, Maria Maddalena Sperotto, and Mohamed Laradji. Predictions of Phase Separation in Three-Component Lipid Membranes by the MARTINI Force Field. *J. Phys. Chem. B*, 117(15):4072–4080, 2013.

[18] Blanca B. Diaz-Rohrer, Kandice R. Levental, and Ilya Levental. Rafting through Traffic: Membrane Domains in Cellular Logistics. *Biochim. Biophys. Acta - Biomembranes*, 1838(12):3003–3013, 2014.

[19] Blanca B. Diaz-Rohrer, Kandice R. Levental, Kai Simons, and Ilya Levental. Membrane Raft Association Is a Determinant of Plasma Membrane Localization. *Proc. Natl. Acad. Sci. U.S.A.*, 111(23):8500–8505, 2014.

[20] Celsa Díaz-Tejada, Igor Ariz-Extreme, Neha Awasthi, and Jochen S. Hub. Quantifying Lateral Inhomogeneity of Cholesterol-Containing Membranes. *J. Phys. Chem. Lett.*, 6:4799–4803, 2015.

[21] David Easterhoff, Fernando Ontiveros, Lauren R. Brooks, Yoel Kim, Brittany Ross, Jharon N. Silva, Joanna S. Olsen, Changyong Feng, Dwight J. Hardy, Paul M. Dunman, and Stephen Dewhurst. Semen-derived enhancer of viral infection (sevi) binds bacteria, enhances bacterial phagocytosis by macrophages, and can protect against vaginal infection by a sexually transmitted bacterial pathogen. *Antimicrobial Agents and Chemotherapy*, 57(6):2443–2450, 2013.

[22] Michael Edidin. The State of Lipid Rafts: from Model Membranes to Cells. *Annu. Rev. Biophys.*, 32:257–283, 2003.

[23] Shih Lin Goh, Jonathan J. Amazon, and Gerald W. Feigenson. Toward a Better Raft Model: Modulated Phases in the Four-Component Bilayer, DSPC/DOPC/POPC/-CHOL. *Biophys. J.*, 104(4):853–862, 2013.

[24] Richard J Gowers, Max Linke, Jonathan Barnoud, Tyler J E Reddy, Manuel N Melo, Sean L Seyler, Jan Domański, David L Dotson, Sébastien Buchoux, Ian M Kenney, and Oliver Beckstein. MDAnalysis: A Python Package for The Rapid Analysis of Molecular Dynamics Simulations. *Proceedings of the 15th Python in Science Conference*, pages 98–105, 2016.

[25] Axelle Grélard, Cécile Loudet, Anna Diller, and Erick J Dufourc. Membrane Protein Structure Determination. In Jean-Jacques Lacapere, editor, *Methods in molecular biology*, volume 654, chapter 18 NMR Spe, pages 341–359. Humana Press, 2010.

[26] Davit Hakobyan and Andreas Heuer. Phase Separation in a Lipid/Cholesterol System: Comparison of Coarse-Grained and United-Atom Simulations. *J. of Phys. Chem. B*, 117(14):3841–3851, 2013.

[27] J.-P. Hansen and I. R. McDonald. *Theory of Simple Liquids (Fourth Edition).* Academic Press, Oxford, 4th edition, 2013.

[28] J. L. Harden, F. C. MacKintosh, and P. D. Olmsted. Budding and domain shape transformations in mixed lipid films and bilayer membranes. *Phys. Rev. E*, 72:011903, Jul 2005.

[29] Shushan He and Lutz Maibaum. Identifying the onset of phase separation in quaternary lipid bilayer systems from coarse-grained simulations. *J. Phys.Chem. B*, 122(14):3961–3973, 2018. PMID: 29547692.

[30] Frederick A. Heberle, Milka Doktorova, Shih Lin Goh, Robert F. Standaert, John Katsaras, and Gerald W. Feigenson. Hybrid and Nonhybrid Lipids Exert Common Effects on Membrane Raft Size and Morphology. *J. Am. Chem. Soc.*, 135(40):14932–14935, 2013.

[31] Frederick A. Heberle, Robin S. Petruzielo, Jianjun Pan, Paul Drazba, Norbert Kučerka, Robert F. Standaert, Gerald W. Feigenson, and John Katsaras. Bilayer Thickness Mismatch Controls Domain Size in Model Membranes. *J. Am. Chem. Soc.*, 135(18):6853–6859, 2013.

[32] Frederick A. Heberle, Jing Wu, Shih Lin Goh, Robin S. Petruzielo, and Gerald W. Feigenson. Comparison of Three Ternary Lipid Bilayer Mixtures: FRET and ESR Reveal Nanodomains. *Biophys. J.*, 99(10):3309–3318, 2010.

[33] E. Hellstrand, A. Nowacka, D. Topgaard, S. Linse, and E. Sparr. Membrane lipid co-aggregation with α-synuclein fibrils. *PLoS One*, 8(10):1–10, 10 2013.

[34] Erik Hellstrand, Agnieszka Nowacka, Daniel Topgaard, Sara Linse, and Emma Sparr. Membrane lipid co-aggregation with -synuclein fibrils. *PLOS ONE*, 8(10):1–10, 10 2013.

[35] David Holowka, Julie A. Gosse, Adam T. Hammond, Xuemei Han, Prabuddha Sengupta, Norah L. Smith, Alice Wagenknecht-Wiesner, Min Wu, Ryan M. Young, and Barbara Baird. Lipid Segregation and IgE Receptor Signaling: A Decade of Progress. *Biochim. Biophys. Acta - Molecular Cell Research*, 1746(3):252–259, 2005.

[36] Aurelia R. Honerkamp-Smith, Sarah L. Veatch, and Sarah L. Keller. An Introduction to Critical Points for Biophysicists; Observations of Compositional Heterogeneity in Lipid Membranes. *Biochim. Biophys. Acta - Biomembranes*, 1788(1):53–63, 2009.

[37] Jing Huang, Sarah Rauscher, Grzegorz Nawrocki, Ting Ran, Michael Feig, Bert L. de Groot, Helmut Grubmuller, and Alexander D. MacKerell Jr. Charmm36m: an improved force field for folded and intrinsically disordered proteins. *Nature Methods*, 14(1):71–73, 2017.

[38] Helgi I. Ingólfsson, Manuel N. Melo, Floris J. Van Eerden, Clément Arnarez, Cesar A. Lopez, Tsjerk A. Wassenaar, Xavier Periole, Alex H. De Vries, D. Peter Tieleman, and Siewert J. Marrink. Lipid Organization of the Plasma Membrane. *J. Am. Chem. Soc.*, 136(41):14554–14559, 2014.

[39] et al. J. Mnch. Semen-derived amyloid fibrils drastically enhance hiv infection. *Cell*, 131:1059–1071, 2007.

[40] Y. Yamamura J. R. Brender, N. R. Roan. Helical conformation of the sevi precursor peptide pap248-286, a dramatic enhancer of hiv infectivity, promotes lipid aggregation and fusion. *Biophys. J.*, 97:2474–2483, 2009.

[41] Eric Jones, Travis Oliphant, Pearu Peterson, and et al. SciPy: Open source scientific tools for Python, 2001–. [Online; accessed 2018-01-22].

[42] Robbie P. Joosten, Tim A. H. te Beek, Elmar Krieger, Maarten L. Hekkelman, Rob W. W. Hooft, Reinhard Schneider, Chris Sander, and Gert Vriend. A series of pdb related databases for everyday needs. *Nucleic Acids Res.*, 39:D411–D419, 2011.

[43] Wolfgang Kabsch and Christian Sander. Dictionary of protein secondary structure: Pattern recognition of hydrogenbonded and geometrical features. *Biopolymers*, 22(12):2577–2637, 1983.

[44] B. L. Kagan, H. Jang, R. Chapone, F. Teran Arce, S. Ramachandran, R. Lai, and R. Nussinov. Antimicrobial properties of amyloid peptides. *Mol. Pharm.*, 9(4):708–717, 2012.

[45] Bruce L. Kagan, Hyunbum Jang, Ricardo Capone, Fernando Teran Arce, Srinivasan Ramachandran, Ratnesh Lal, and Ruth Nussinov. Antimicrobial properties of amyloid peptides. *Mol. Pharm.*, 9(4):708–717, 2012.

[46] Robert E Kass and Adrian E Raftery. Bayes Factors. *J. Am. Stat. Assoc.*, 90(430):773–795, 1995.

[47] Jeffery B Klauda, Norbert Kučerka, Bernard R Brooks, Richard W Pastor, and John F Nagle. Simulation-Based Methods for Interpreting X-Ray Data from Lipid Bilayers. *Biophys. J*, 90(8):2796–2807, 2006.

[48] Tatyana M. Konyakhina, Shih Lin Goh, Jonathan Amazon, Frederick A. Heberle, Jing Wu, and Gerald W. Feigenson. Control of a Nanoscopic-to-Macroscopic Transition: Modulated Phases in Four-Component DSPC/DOPC/POPC/Chol Giant Unilamellar Vesicles. *Biophys. J.*, 101(2):L8–L10, 2011.

[49] Tatyana M. Konyakhina, Jing Wu, James D. Mastroianni, Frederick A. Heberle, and Gerald W. Feigenson. Phase Diagram of a 4-Component Lipid Mixture: DSPC/-DOPC/POPC/Chol. *Biochim. Biophys. Acta - Biomembranes*, 1828(9):2204–2214, 2013.

[50] Norbert Kučerka, John F Nagle, Jonathan N Sachs, Scott E Feller, Jeremy Pencer, Andrew Jackson, and John Katsaras. Lipid Bilayer Structure Determined by the Simultaneous Analysis of Neutron and X-Ray Scattering Data. *Biophys. J.*, 95(5):2356–2367, 2008.

[51] D. K. V. Kumar, S. H. Choi, K. J. Washicosky, W. A. Eimer, S. Tucker, J. Ghofrani, A. Lefkowitz, G. McColl, L. E. Goldstein, R. E. Tanzi, and R. D. Moir. Amyloid-β peptide protects against microbial infection in mouse and worm models of Alzheimer's disease. *Science Transl. Med.*, 8(340):340ra72, 2016.

[52] Deepak Kumar Vijaya Kumar, Se Hoon Choi, Kevin J. Washicosky, William A. Eimer, Stephanie Tucker, Jessica Ghofrani, Aaron Lefkowitz, Gawain McColl, Lee E. Goldstein, Rudolph E. Tanzi, and Robert D. Moir. Amyloid-β peptide protects against microbial infection in mouse and worm models of alzheimer's disease. *Science Translational Medicine*, 8(340):340ra72–340ra72, 2016.

[53] J. Lee, X. Cheng, J. M. Swails, M. S. Yeom, P. K. Eastman, J. A. Lemkul, S. Wei, J. Buckner, J. C. Jeong, Y. Qi, S. Jo, V. S. Pande, D. A. Case, C. L. Brooks, A. D. MacKerell, J. B. Klauda, and W. Im. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theo. Comp.*, 12(1):405–413, 2016.

[54] Avigdor Leftin, Trivikram R Molugu, Constantin Job, Klaus Beyer, and Michael F Brown. Article Area per Lipid and Cholesterol Interactions in Membranes from Separated Local-Field 13C NMR Spectroscopy. *Biophys. J.*, 107(10):2274–2286, 2014.

[55] Ilya Levental and Sarah L. Veatch. The Continuing Mystery of Lipid Rafts. *J. Mol. Biol.*, 428(24):4749–4764, 2016.

[56] Göran Lindblom and Greger Orädd. Lipid Lateral Diffusion and Membrane Heterogeneity. *Biochim. Biophys. Acta - Biomembranes*, 1788(1):234–244, 2009.

[57] Daniel Lingwood and Kai Simons. Lipid Rafts as a Membrane-Organizing Principle. *Science*, 327(5961):46–50, 2010.

[58] J. R. Long, N. Oyler, G. P. Drobny, and P. S. Stayton. Assembly of α-helical peptide coatings on hydrophobic surfaces. *J. Am. Chem. Soc.*, 124(22):6297–6303, 2002.

[59] Siewert J. Marrink, H. Jelger Risselada, Serge Yefimov, D. Peter Tieleman, and Alex H. De Vries. The MARTINI Force Field: Coarse Grained Model for Niomolecular Simulations. *J. Phys. Chem. B*, 111(27):7812–7824, 2007.

[60] Siewert J. Marrink, Jelger Risselada, and Alan E. Mark. Simulation of Gel Phase Formation and Melting in Lipid Bilayers Using a Coarse Grained Model. *Chem. Phys. Lipids*, 135(2):223–244, 2005.

[61] Siewert J. Marrink, Ah De Vries, and Ae Mark. Coarse Grained Model for Semiquantitative Lipid Simulations. *J. Phys. Chem. B*, 108:750–760, 2004.

[62] D. Marsh. Cholesterol-induced fluid membrane domains: A compendium of lipid-raft ternary phase diagrams. *Biochim. Biophys. Acta – Biomembranes*, 1788(10):2114–2123, 2009.

[63] S. Materniak, A. Patrykiejew, and S. Sokoowski. The Phase Behavior of Two-Dimensional Symmetrical Mixtures. *J. Chem. Phys.*, 133(24), 2010.

[64] K K Mon and Kurt Binder. Finite Size Effects for the Simulation of Phase Coexistence in the Gibbs Ensemble Near the Critical Point. *J. Chem. Phys.*, 96(1992):6989–6995, 1992.

[65] M. F. Moody. Diffraction by dispersion of spherical membrane vesicles. I. The basic equations. *Acta Crystallogr. A*, 31(1):8–15, 1975.

[66] A. D. Miranker N. Last, D. E. Schlamadinger. A common landscape for membrane-active peptides. *Protein Sci.*, 22:870–882, 2013.

[67] et al. N. R. Roan. Semen amyloids participate in spermatozoa selection and clearance. *Elife.*, 6:24888, 2017.

[68] E. Nakatani-Webster and A. Nath. Inferring mechanistic parameters from amyloid formation kinetics by approximate Bayesian computation. *Biophys. J.*, 112(5):868–880, 2017.

[69] Ravi P. R. Nanga, Jeffrey R. Brender, Subramanian Vivekanandan, Nataliya Popovych, and Ayyalusamy Ramamoorthy. Nmr structure in a membrane environment reveals putative amyloidogenic regions of the sevi precursor peptide pap248286. *J. Am. Chem. Soc.*, 131(49):17972–17979, 2009. PMID: 29547692.

[70] Chiara Nicolini, Julia Kraineva, Monika Khurana, Nagarajan Periasamy, Sérgio S. Funari, and Roland Winter. Temperature and Pressure Effects on Structural and Conformational Properties of POPC/SM/Cholesterol Model Raft Mixtures-a FT-IR, SAXS, DSC, PPC and Laurdan Fluorescence Spectroscopy Study. *Biochim. Biophys. Acta - Biomembranes*, 1758(2):248–258, 2006.

[71] Perttu S. Niemela, Samuli Ollila, Marja T. Hyvonen, Mikko Karttunen, and Ilpo Vattulainen. Assessing the Nature of Lipid Raft Membranes. *PLOS Comput. Biol.*, 3(2):0304–0312, 2007.

[72] T. E. Oliphant. Python for scientific computing. *Comput. Sci. Eng.*, 9(3):10–20, May 2007.

[73] Sagar A. Pandit, S. Vasudevan, S.W. Chiu, R. Jay Mashl, Eric Jakobsson, and H.L. Scott. Sphingomyelin-Cholesterol Domains in Phospholipid Membranes: Atomistic Simulation. *Biophys. J.*, 87(2):1092–1100, 2004.

[74] G. A. Pantelopulos, T. Nagai, A. Bandara, A. Panahi, and J. E. Straub. Critical size dependence of domain formation observed in coarse-grained simulations of bilayers composed of ternary lipid mixtures. *J. Chem. Phys.*, 147:095101, 2017.

[75] Fabian Pedregosa, Ron Weiss, and Matthieu Brucher. Scikit-Learn : Machine Learning in Python. 12:2825–2830, 2011.

[76] Jason D Perlmutter and Jonathan N Sachs. Inhibiting Lateral Domain Formation in Lipid Bilayers: Simulations of Alternative Steroid Headgroup Chemistries. *J. Am. Chem. Soc.*, 131(45):16362–16363, 2009.

[77] Robin S. Petruzielo, Frederick A. Heberle, Paul Drazba, John Katsaras, and Gerald W. Feigenson. Phase Behavior and Domain Size in Sphingomyelin-Containing Lipid Bilayers. *Biochim. Biophys. Acta - Biomembranes*, 1828(4):1302–1313, 2013.

[78] Sander Pronk, Szilárd Páll, Roland Schulz, Per Larsson, Pär Bjelkmar, Rossen Apostolov, Michael R. Shirts, Jeremy C. Smith, Peter M. Kasson, David Van Der Spoel, and et al. GROMACS 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit. *Bioinformatics*, 29(7):845–854, 2013.

[79] L. Rajendran. Lipid Rafts and Membrane Dynamics. *J. Cell Sci.*, 118(6):1099–1102, 2005.

[80] M. C. Rheinstädter and O. G. Mouritsen. Small-scale structure in fluid cholesterol–lipid bilayers. *Curr. Op. in Coll. & Interf. Sci.*, 18(5):440–447, 2013.

[81] H Jelger Risselada and Siewert J. Marrink. The Molecular Face of Lipid Rafts in Model Membranes. *Proc. Natl. Acad. Sci. U.S.A.*, 105(45):17367–17372, 2008.

[82] H. Jelger Risselada and Siewert J. Marrink. The Molecular Face of Lipid Rafts in Model Membranes. *Proc. Natl. Acad. Sci. U. S. A.*, 105(45):17367–17372, 2008.

[83] N. R. Roan, N. Sandi-Monroy, N. Kohgadai, S. M. Usmani, K. G. Hamil, J. Neidleman, M. Montano, L. Ständker, and A. Röcker. Semen amyloids participate in spermatozoa selection and clearance. *eLife*, 6:e24888, 2017.

[84] Nadia R. Roan, Jan Münch, Nathalie Arhel, Walther Mothes, Jason Neidleman, Akiko Kobayashi, Karen Smith-McCune, Frank Kirchhoff, and Warner C. Greene. The cationic properties of sevi underlie its ability to enhance human immunodeficiency virus infection. *Journal of Virology*, 83(1):73–80, 2009.

[85] Tomasz Róg, Marta Pasenkiewicz-Gierula, Ilpo Vattulainen, and Mikko Karttunen. Ordering Effects of Cholesterol and Its Analogues. *Biochim. Biophys. Acta - Biomembranes*, 1788(1):97–121, 2009.

[86] C. Rosetti and C. Pastorino. Comparison of Ternary Bilayer Mixtures with Asymmetric or Symmetric Unsaturated Phosphatidylcholine Lipids by Coarse Grained Molecular Dynamics Simulations. *J. Phys. Chem. B*, 116(11):3525–3537, 2012.

[87] E. Rhoades S. Elbaum-Garfinkle, T. Ramlall. Sthe role of the lipid bilayer in tau aggregation. *Biophys. J.*, 98:2722–2730, 2010.

[88] K. Sapp and L. Maibaum. Suppressing membrane height fluctuations leads to a membrane-mediated interaction among proteins. *Phys. Rev. E*, 94:052424, 2016.

[89] Kayla Sapp, Roie Shlomovitz, and Lutz Maibaum. Seeing the Forest in Lieu of the Trees: Continuum Simulations of Cell Membranes at Large Length Scales. *Annu. Rep. Comput. Chem.*, 10(4):47–76, 2014.

[90] Michael Schick. Membrane heterogeneity: Manifestation of a curvature-induced microemulsion. *Phys. Rev. E*, 85:031902, 2012.

[91] Michael Schick. Strongly correlated rafts in both leaves of an asymmetric bilayer. *J. Phys. Chem. B*, 122(13):3251–3258, 2018.

[92] Prabuddha Sengupta, Barbara Baird, and David Holowka. Lipid Rafts, Fluid/Fluid Phase Separation, and Their Relevance to Plasma Membrane Structure and Function. *Semin. Cell Dev. Biol.*, 18(5):583–590, 2007.

[93] Prabuddha Sengupta, David Holowka, and Barbara Baird. Fluorescence Resonance Energy Transfer between Lipid Probes Detects Nanoscopic Heterogeneity in the Plasma Membrane of Live Cells. *Biophys. J.*, 92(10):3564–3574, 2007.

[94] Daniel A. Sheik, Jeffrey M. Chamberlain, Lauren Brooks, Melissa Clark, Young Hun Kim, Geoffray Leriche, Clifford P. Kubiak, Stephen Dewhurst, and Jerry Yang. Hydrophobic nanoparticles reduce the -sheet content of sevi amyloid fibrils and inhibit sevi-enhanced hiv infectivity. *Langmuir*, 33(10):2596–2602, 2017. PMID: 28207276.

[95] Roie Shlomovitz, Lutz Maibaum, and Michael Schick. Macroscopic phase separation, modulated phases, and microemulsions: A unified picture of rafts. *Biophys. J.*, 106(9):1979–1985, 2014.

[96] Roie Shlomovitz and M. Schick. Model of a Raft in Both Leaves of an Asymmetric Lipid Bilayer. *Biophys. J.*, 105(6):1406–1413, 2013.

[97] Kai Simons and Mathias J. Gerl. Revitalizing Membrane Rafts: New Tools and Insights. *Nat. Rev. Mol. Cell Biol.*, 11(10):688–699, 2010.

[98] Kai Simons and Elina Ikonen. Functional Rafts in Cell Membranes. *Nature*, 387(6633):569–572, 1997.

[99] Kai Simons and Julio L Sampaio. Membrane Organization and Lipid Rafts. *Cold Spring Harbor perspectives in biology*, 3:a004697(10):1–17, 2011.

[100] Kai Simons and Winchil L C Vaz. Model Systems, Lipid Rafts, and Cell Membranes. *Annu. Rev. Biophys.*, 33:269–295, 2004.

[101] Alexander J. Sodt, Richard W. Pastor, and Edward Lyman. Hexagonal Substructure and Hydrogen Bonding in Liquid-Ordered Phases Containing Palmitoyl Sphingomyelin. *Biophys. J.*, 109(5):948–955, 2015.

[102] Matthew B. Stone, Sarah A. Shelby, and Sarah L. Veatch. Super-Resolution Microscopy: Shedding Light on the Cellular Plasma Membrane. *Chem. Rev.*, 117(11):7457–7477, 2017.

[103] Q. Tang, N. R. Roan, and Y. Yamamura. Seminal plasma and semen amyloids enhance cytomegalovirus infection in cell culture. *J. Virol.*, 87(23):12583–12591, 2013.

[104] Qiyi Tang, Nadia R. Roan, and Yasuhiro Yamamura. Seminal plasma and semen amyloids enhance cytomegalovirus infection in cell culture. *J. Virol.*, 87:12583–12591, 2013.

[105] John Toner and David R. Nelson. Smectiv, cholesteric, and Rayleigh-Benard order in two dimensions. *Phys. Rev. B*, 23(1):316–334, 1981.

[106] Alexandre Toulmay and William A. Prinz. A Conserved Membrane-Binding Domain Targets Proteins to Organelle Contact Sites. *J. Cell Sci.*, 125(Pt 1):49–58, 2012.

[107] Alexandre Toulmay and William A. Prinz. Direct Imaging Reveals Stable, Micrometer-Scale Lipid Domains that Segregate Proteins in Live Cells. *J. Cell Biol.*, 202(1):35–44, 2013.

[108] S. Toxvaerd. Phase-Stability Lines, Spinodals and Distribution Functions for Binary Immiscible Mixtures. *Mol. Phys.*, 86(4):845–855, 1995.

[109] David van der Spoel, Erik Lindahl, Berk Hess, Aldert R. van Buuren, Emile Apol, Peiter J. Meulenhoff, Peter D. Tieleman, Alfons L. T. M. Sijbers, Anton K. Feenstra, Rudi van Drunen, and Herman J. C. Berendsen. *GROMACS User Manual Version 4.6.5*. University of Groningen, 2010.

[110] Sarah L. Veatch, Pietro Cicuta, Prabuddha Sengupta, Aurelia Honerkamp-Smith, David Holowka, and Barbara Baird. Critical Fluctuations in Plasma Membrane Vesicles. *ACS Chem. Biol.*, 3(5):287–293, 2008.

[111] Sarah L. Veatch and Sarah L. Keller. Organization in lipid membranes containing cholesterol. *Phys. Rev. Lett.*, 89:268101, Dec 2002.

[112] Sarah L. Veatch and Sarah L. Keller. Separation of Liquid Phases in Giant Vesicles of Ternary Mixtures of Phospholipids and Cholesterol. *Biophys. J.*, 85(5):3074–3083, 2003.

[113] Sarah L. Veatch and Sarah L. Keller. Seeing Spots: Complex Phase Behavior in Simple Membranes. *Biochim. Biophys. Acta - Molecular Cell Research*, 1746(3):172–185, 2005.

[114] Sarah L. Veatch, Olivier Soubias, Sarah L. Keller, and Klaus Gawrisch. Critical Fluctuations in Domain-Forming Lipid Mixtures. *Proc. Natl. Acad. Sci. U.S.A.*, 104:17650–17655, 2013.

[115] Tsjerk A. Wassenaar, Helgi I. Ingólfsson, Rainer A. Bockmann, D. Peter Tieleman, and Siewert J. Marrink. Computational Lipidomics with Insane: A Versatile Tool for Generating Custom Membranes for Molecular Simulations. *J. Chem. Theory Comput.*, 11(5):2144–2155, 2015.

[116] Tobias Weidner, Newton T. Samuel, Keith McCrea, Lara J. Gamble, Robert S. Ward, and David G. Castner. Assembly and structure of α-helical peptide films on hydrophobic fluorocarbon surfaces. *Biointerphases*, 5(1):9–16, 2010.

[117] E. L. Wu, X. Cheng, S. Jo, H. Rui, K. C. Song, E. M. Dávila-Contreras, Y. Qi, J. Lee, V. Monje-Galvan, R. M. Venable, J. B. Klauda, and W. Im. CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. *J. Comp. Chem.*, 35(27):1997–2004, 2014.

[118] Ruitian Zhang, Robert M. Suter, and John F. Nagle. Theory of the Structure Factor of Lipid Bilayers. *Phys. Rev. E*, 50(6):5047–5060, 1994.

[119] Jiang Zhao, Jing Wu, Frederick A. Heberle, Thalia T. Mills, Paul Klawitter, Grace Huang, Greg Costanza, and Gerald W. Feigenson. Phase Studies of Model Biomembranes: Complex Behavior of DSPC/DOPC/Cholesterol. *Biochim. Biophys. Acta - Biomembranes*, 1768(11):2764–2776, 2007.