Non-ideality of mixing in phospholipid bilayer system : molecular dynamics view

Yelyzaveta Berezovska

Université de Strasbourg

UNIVERSITÉ DE STRASBOURG

École doctorale Physique, chimie-physique ED 182 Université de Strasbourg

ÉCOLE DOCTORALE PHYSIQUE ET CHIMIE-PHYSIQUE

Institut Charles Sadron (ICS, UPR22)

THÈSE présentée par :

Yelyzaveta BEREZOVSKA

soutenue le : **12 December 2022**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Physique

Mélanges non-idéaux dans les bicouches phospholipidiques: approche par dynamique moléculaire

THÈSE dirigée par :

Dr. THALMANN Fabrice Maître de Conférences, Université de Strasbourg

RAPPORTEURS :

Dr. LOISON Claire Chargé de recherches, Université de Lyon

Dr. CALLAN-JONES Andrew Maître de Conférences, Université Paris-Diderot

---

AUTRES MEMBRES DU JURY :

Dr. MEYER Hendrik Directeur de recherches, Université de Strasbourg

Dr. MARQUES Carlos Directeur de recherches, Université de Lyon

Acknowledgements

First and foremost, I would like to thank my supervisor, Dr. Fabrice Thalmann, for his invaluable advice, continuous support, and patience during my PhD study. Fabrice introduced me to the world of lipids and made it possible for me to apply my skills to a biologically inspired matter. Thanks to his broad knowledge it was always interesting to discuss with Fabrice regardless of the subject.

I would also like to express my gratitude to all the members in the Mcube team. I will always remember how entertaining and instructive were the lectures of Thierry C. and Wiebke D., Carlos M. and Antonio S. during my Master study. They were mainly the reason why I wanted to join Institut Charles Sadron and MCube afterwards as they showed me that soft matter science can be both fun and complex. I am grateful to Tatiana S. for our passionate discussions about cats and her help with them. It is great having someone to share such love. I would like to thank Pierre M. for his support during my thesis and patience for my not always straight forward way of explaining things.

I am thankful to all PhD students of MCube: Eulalie L., Ann G., Swen H., Florent F., Friedrich W. and Clement M. They showed me how welcoming and open a team can be, especially during our lunch discussions when I discovered a lot of new things and, specifically, definitions. I am glad we could all become friends during my time in Strasbourg. I am especially grateful to Tatiana S., Eulalie L. and Ann G. for our small sports community where I finally had a chance to try tennis.

I would like to thank the postdocs I met during my study: Celine R., William F. and Fedir D., for always supporting me and giving nice advice. The example they showed me kept me going during hard times.

I would also like to express my gratitude to the TSP group for having me in their team. I would like to thank Hendrik M. for accepting me for the Master internship and introducing me to LAMMPS and Molecular Dynamics, Joerg B. for having patience during his lectures on Statistical Physics and Olivier B. for great discussions, technical support and tasty mint tea.

I am grateful to all Institut Charles Sadron for the opportunity to work there and meet numerous interesting and inspiring people. The atmosphere of ICS will always be in my heart. I would like to thank the BJS student organisation for an opportunity to socialise with other students and tasty snacks during their coffee afternoon.

I would like to thank all of my friends. Some of them I knew for a long time already, Maria A., Katya B., Anastasia H., and some I met during my stay in Strasbourg, Yuliana Z., Iryna M., Anastasia S. They always supported me and gave their advice. I am grateful for making my time in France enjoyable and fun. Specially, I would like to thank my close friend Liudmyla K. for always being there for me and sharing my passion for stupid series and junk food. Liudmyla was the first person I became friends with in Strasbourg and she helped me a lot during my thesis journey.

I would like to express my gratitude to my family. Without their tremendous understanding and encouragement in the past few years, it would be impossible for me to complete my study. I am thankful to my father, Oleksii B., who supported me to his best ability. And I am especially thankful to my mother, Svitlana K., who gave her time and love to me regardless of my response, and who patiently made me the person I am today. I will forever be grateful to her.

And of course, I should not forget to thank anyone reading this thesis. I hope you find something enjoyable on the following pages.

Contents

List of Figures

List of Tables

Acronyms

2d two dimensions.

APL Area per lipid.

CG Coarse-grained.
COM Center of mass.

DLiPC 1,2-dilinoleoyl-sn-glycero-3-phosphocholine.
DOPC 1,2-dioleoyl-sn-glycero-3-phosphocholine.
DOPE 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine.
DPPC 1,2-dipalmitoyl-sn-glycero-3-phosphatidylcholine.

FF Force field.

LAMMPS Large-scale Atomic/Molecular Massively Parallel Simulator.
LJ Lennard-Jones.

MD Molecular Dynamics.

NPT Isobaric-Isothermal ensemble.
NVT Canonical ensemble.

PBC Periodic boundary conditions.
pCOM Pseudo center of mass.
PDB Protein Data Bank.

PSM Palmitoyl sphingomyelin.

SSM Stearoyl sphingomyelin.

VMD Visual Molecular Dynamics.

To my grandmother, Valentyna, who inspired me to explore life

Introduction

There are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns — the ones we don't know we don't know.

Donald Rumsfeld

EVERY organism in our world consists of all kinds of cells. These organisms vary in size and complexity, starting with some fascinating one-cell species such as bacterias or yeasts, and going up to more complex ones with animals, mushrooms, plants, etc. They have different cells with various functions and organelles inside. There is a barrier surrounding the insides of each cell and that exactly is a membrane. The fundamental architecture of any biological membrane is a phospholipid bilayer with proteins, sterols, etc. embedded in it or surrounding it. Lipid molecules can self-assemble and form bilayer structures in aqueous media so that their physical and chemical properties determine many of the cell membrane properties. In biological membranes, thousands of lipid species are found while the function of many of them remains a mystery to modern science [46, 89]. The lipids' functions and importance in biological membranes took a long time to be properly recognized. It's only by 1935 that Dannielli and Davson managed to propose the basic elements to model and understand realistic membranes in their "Pauci-Molecular" model [11]. We will discuss the history of the evolution of the membrane later in this chapter.

Model membrane properties are studied in both experiments and simulations. Each of these two approaches has its pros and cons. Firstly, let us describe the experimental possibilities [90] and then move on to the simulation part. One of the most commonly used model membrane systems is the multilamellar liposome. One can imagine it as an onion with multiple bilayers forming a skin. A second popular approach is to work with unilamellar vesicles of different sizes. Small ones are useful because of their homogeneity and high lipid concentration while large ones are good for transport experiments and they can mimic biological cells. More generally, vesicles are suitable for drug delivery:

large ones due to their ability to encapsulate materials and small ones because they can be transported in thin capillary while avoiding the negative reaction from the body. Last but not least to be mentioned here is the lipid monolayer in the Langmuir trough, which is a technique to create and characterize single molecule thick films by controlling the packing density of the molecules.

By contrast the study of membrane model systems with computational methods often permits the examination of questions for which the biological membranes present too complex structures for the experimental techniques. The main goal of the present work was to investigate some of the available model lipids using Molecular Dynamics (MD) simulations of a free lipid bilayer in water solution. Although these membrane model systems do not exhibit all the complexity of a biological membrane, they effectively mimic some of the properties conferred on membranes by the lipid bilayer. These pure lipid systems do not contain membrane proteins and the lipid composition is much simpler than found in biological membranes. Despite the many limitations present in these systems, our study gives a quantitative and qualitative insight into the complex mixing phenomena of model lipid mixtures.

We present a theoretical framework to treat symmetric lipid bilayers with Molecular Dynamics simulations. We focus on the thermodynamic properties of membranes, and we use for this purpose a Coarse-grained (CG) approach. We study the mixing of a few model lipids and lipids-cholesterol mixtures, also looking at their dynamic and mechanical features. We access directly the structure of the studied objects and obtain valuable information relative to the mechanical compressibility and to the lipid interactions within a leaflet, or across opposite leaflets. More generally, studies of complex lipid mixtures with the use of Molecular Dynamics broaden our knowledge of lipid-lipid interactions.

Outline of the work

The thesis is organized as follows. Firstly, we introduce some general facts about cell membranes and how we can work with these systems from a physical point of view. We discuss a few membrane models and introduce the basics of lipid molecules in Chapter 1. This will help the reader to understand more about the background and biological inspiration of this work.

Chapter 2 outlines the main theoretical concepts that are needed to interpret our simulations, including the theory of regular solutions and the Kirkwood-Buff approaches to the thermodynamics of mixing.

In Chapter 3 we discuss the computational details and introduce the reader to the basic concepts of Molecular Dynamics and the Coarse-Graining model SPICA that is used in this work.

In Chapters 4 and 5 we present the results of this work and discuss the possible interpretations of the obtained results. The reader can follow our steps in the discovery of the lipid membrane structural and dynamic characteristics. We start our discussion in Chapter 4 with a pure DPPC bilayer, establish the simulations and analysis framework and generalize it to the other available lipids: DOPC, DOPE, DLiPC, PSM and SSM. We describe membrane characteristics such as the structure factors and area compressibility. We discuss in detail the influence of the choice of the lipid representation on the obtained results. In Chapter 5 we study a binary mixture of phospholipids: DPPC and DLiPC. Firstly, we develop a "fake" mixture procedure to test our predictions regarding the Kirkwood-Buff and regular solutions theories. Next we probe the consistency of the area compressibility and mixing parameters that are obtained. Finally, we present lipid mixtures containing Cholesterol and deliver our interpretation of the mixing properties of such systems.

In the final Chapter 5.2, we draw our conclusions on this thesis work and provide a few insights into possible future developments that could be done with profit.

CHAPTER 1

Membranes, lipids, and their models

It is common to make some simplifications when it comes to the study of such complex systems as a cell membrane. When we are talking about the sun we usually forget details like sunspots or prominence. When we look at the lake we most likely think of the water surface and not the molecules from which it is made. And similarly here we concentrate on the phospholipids that are the main blocks of the membrane and ignore the inner organelles or proteins present in biological cells. Of course, for someone who is focused on biological or chemical aspects that would sound oversimplified but from the physical point of view it is a fairly good assumption.

Let us start a discussion in this chapter with a part of the history of the model biological membranes. Afterwards we will discover the complexity of lipid molecules from chemical, physical and biological point of view. We will describe some simple cell membranes and move on to the importance of the lipid-lipid interactions. Finally, we will conclude with a brief overview of the model bilayers in simulations.

1.1 Short history of a model membrane

Model membranes were developed not so long ago with the first attempt in 1935 by Danielli and Davson with their **"Pauci-Molecular" model** [11, 77]. This happened after 10 years of the first lipid membrane observation during the experiment by Gorter and

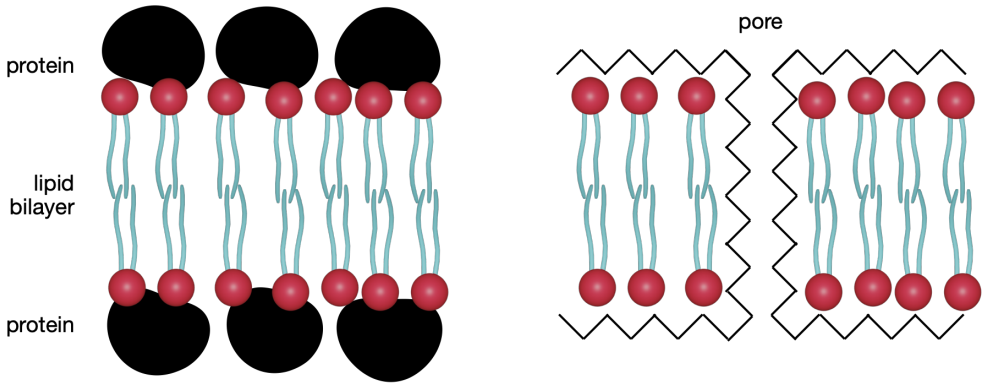


FIGURE 1.1: A schematic representation of the Pauci-Molecular model by Danielli and Davson. The left-hand side shows the original model consisting of the lipid bilayer and globular proteins in a "sandwich". The right-hand side shows an improved version with trans-membrane pores present, which could explain protein transfer through the membrane.

Grendel [18]. They did try to propose a pure lipid membrane model by measuring microscopically the surface area of the bilayer and corresponding monolayer and comparing them. They stated that if the area of the bilayer is exactly twice smaller than the area of a monolayer then the membrane can be made out of a lipid bilayer. And they got the results to confirm their theory, with some errors of course. It is known nowadays that membrane has some proteins embedded in it and so depending on the concentration of both lipids and proteins the surface area will also change.

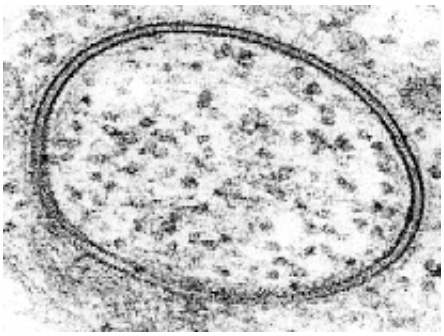


FIGURE 1.2: A typical example of an electron micrograph of a plasma membrane with "railroad tracks" is shown. Robertson based the Unit membrane model on the idea of plasma membrane wrapping the cell. The membrane is seen as dark lines with a light interior area [87]. Thickness of the membrane is around 7 nm.

The first theoretical model was not a pure phospholipid bilayer but a structure with proteins involved. At the time only the globular proteins were known and so the first realistic membrane model looked like a protein-bilayer-protein sandwich (Fig. 1.1). Later on, Danielli added some thin peptide-lined trans-membrane pores to support the idea of membrane permeability but the model remained static with no dynamics included.

The next step was done by Robertson in 1957 when he proposed his **Unit membrane model** [66]. He obtained electron micrographs (EMs), which showed unusual structure as a barrier of the cell (Fig. 1.2). Robertson saw a black contour with a lighter area inside as a membrane after staining it with $KMnO_4$ that he called "railroad tracks". Now

we know that the black area would correspond to the polar head group of the lipid inside the membrane but at the time Robertson assumed that was proteins floating on the lighter membrane (as in the Danielli-Davson model). The only advantage was that he could estimate the thickness of the membrane using EM. Robertson defined his Unit membrane as a bilayer of $\sim 65 - 100$ Å wide that includes two protein layers of around $20 - 30$ Å and phospholipid bilayer of around $35 - 40$ Å. This model had the same uncertainties as the previous one and further developments were needed.

Further development to define the cell membrane model was done in 1972 by Singer and Nicolson when they proposed a **Fluid Mosaic model** that viewed cell membranes as a 2d solution of oriented globular proteins and lipids [73]. In Fig. 1.3 one can see a schematic structure of the Singer-Nicolson model. They added different proteins into the membrane body and made a "mosaic structure" where both proteins and lipids were dispersed inhomogeneously. This allowed them to create a more realistic model that takes into account lateral and trans-membrane protein placement while maintaining the lipid bilayer as a building block of the membrane. Finally, the Fluid Mosaic model included information about the dynamics of the membrane, for the first time modeling a biological membrane as a fluid and changing environment.

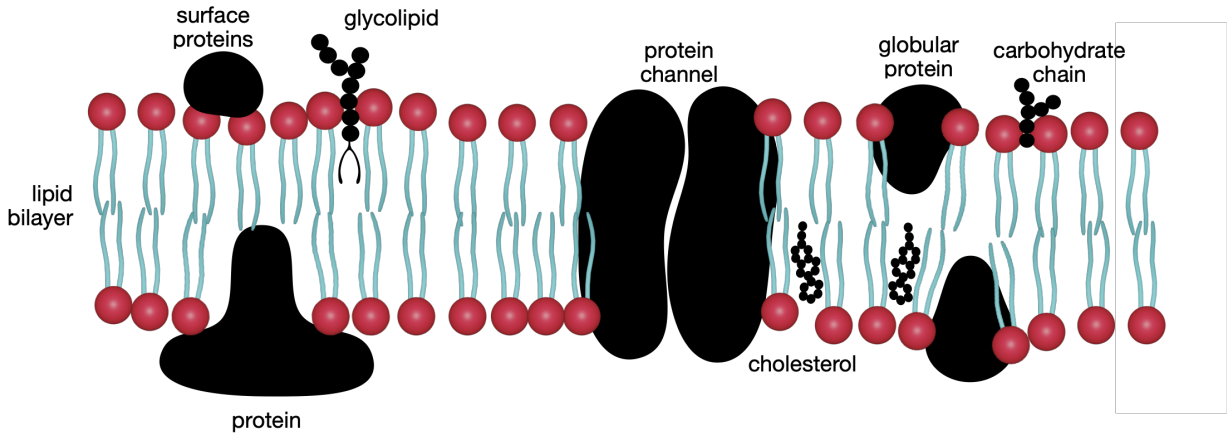


FIGURE 1.3: A schematic representation of the Fluid Mosaic model of the membrane by Singer and Nicolson. Some proteins are presented as an example of a possible range of included molecules.

On the other hand, all previous models did not include or included very little information on the lipid heterogeneity of the bilayers. Already in 1970s some early proof of the domain formations in lipid membranes was obtained and so it was assumed that composition of the compounds in biological membranes may play role in the regulation of relevant physiological processes [53]. Thus, a few alternative models for cell membrane were proposed. One of them is the so-called "plate model" by Jain and White [28]. This approach simplified membrane to a plane with ordered and disordered domains that

could occur in the biological membranes spontaneously as a consequence of specific inter molecular interactions, external environment or lattice deformation. The other idea of a model for cell membranes was presented in 1984 by Mouritsen and Bloom [52] in their "mattress model". This representation of a bilayer membrane associates interactions between lipids and proteins with a a hydrophobicity in the system. Such phenomena gives rise to the interfacial tension between lipids and proteins, and so specific lipid type can cluster around proteins to form phase separated regions.

Another development was done by Karnovsky and Klausner in 1982 when they formalized lipid clusters in membranes [29] and by Kai Simons when he gave the name "lipid raft" to these clusters [72]. The catchy name gave a burst to the research in the area of lipid rafts and since then it was one of the most studied membrane topics. However, it was still not obvious which biological properties were influenced by these formations. Fortunately, a bloom of studies of sphingolipids and Cholesterol mixtures happened not so long after and new features were discovered. Now a lot of research is done using both computational and experimental methods [67]. It is currently well-known that mixtures of Cholesterol with phospholipids, sphingolipids, and proteins play an important role in membrane activity in living organisms. The concentration of Cholesterol in such systems can influence the mechanical properties of the membrane as the so-called lipid rafts tend to be formed and so the study of lipid-lipid interactions has an important role to understand better mechanisms behind this effect. Although some development has already been made [3] there was still no quantitative proof of the interactions involved yet.

1.2 Lipids in nature

Let us discuss in more detail the variety of lipids that exist in nature and define lipids as molecules. A common feature of most lipid molecules of biological membranes is their amphipathic structure. This means that part of the lipid is hydrophobic and another part is hydrophilic. Often these parts are referred to as a "tail group" and a "head group" of the molecule. The tail group usually includes long hydrocarbon chains of covalently bonded carbon atoms while more polar atoms such as phosphorus, oxygen or nitrogen can be present in the head group. The latter can also be polar and contain charged atoms. There are a wide variety of lipid head groups that have different chemical structures which means different sizes and weights. Lipids can have various tail lengths and unsaturation level, which alter the overall size and shape of the molecule.

A variety of interactions exists between different and same lipid types. Overall, a variety of molecules can be included into the lipids group. Some of the examples are fats, waxes, sterols, fat-soluble molecules, glycerides, and others. More specifically, plasma

membrane can include phospholipids, sphingolipids, Cholesterol, etc. Here we will focus on lipids that were used in the present study and discuss their main features and characteristics (Fig. 1.4) [89].

It is interesting to mention here that is the biological membranes lipids are mostly present in the fluid phase. Overall, lipid systems can undergo a phase transition when the system temperature crosses the critical temperature, the so-called melting temperature. Lipids can be in liquid disordered (L_d) phase, liquid ordered (L_o) phase, and gel (L_β) phase. Both L_d and L_o phases are fluid phases, but the liquid ordered phase is a tad more complex to define. This state has similar properties to the liquid crystalline phase and usually occurs for mixtures of phospholipids or sphingolipid with Cholesterol. Unfortunately, the scope of this work does not include the investigation of the phase transition phenomena in the lipid bilayers but is indeed an interesting topic to extend for the future research.

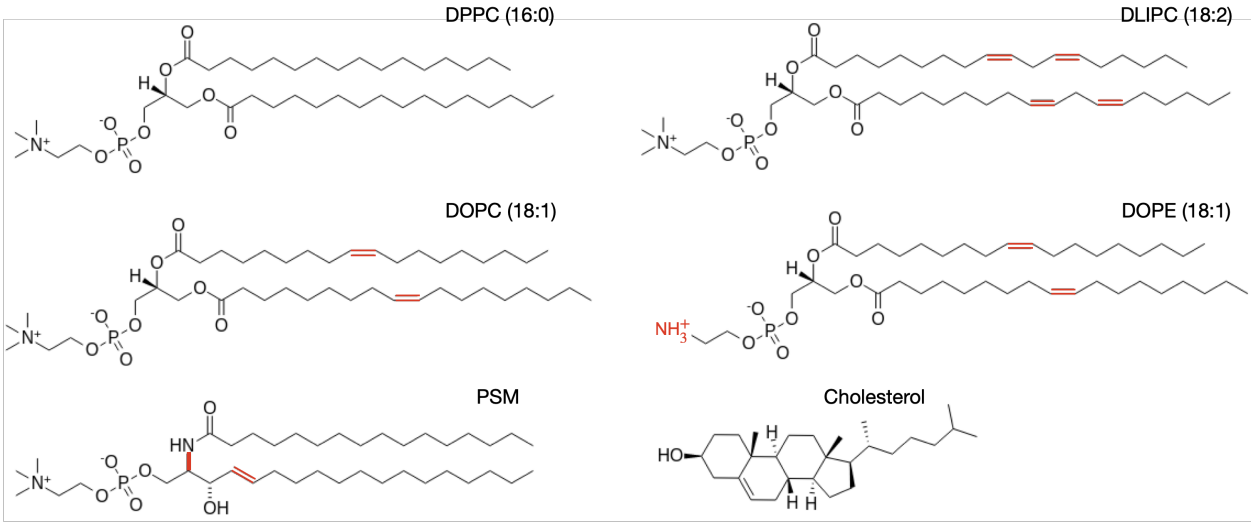


FIGURE 1.4: Chemical formulas of molecules used in the present work: DPPC, DOPC, DOPE, DLiPC, PSM, and Cholesterol.

Phospholipids

The best-known and most studied lipids of biological membranes are phospholipids. They take a big part in my thesis work as the most experimental and computational results are present for phospholipids making it easier to compare and check consistency in the results. One can see in Fig. 1.4 four phospholipids that were used in this work: 1,2-dipalmitoyl-sn-glycero-3-phosphatidylcholine (DPPC), 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE) and 1,2-dilinoleoyl-sn-glycero-3-phosphocholine (DLiPC). Phospholipids fit perfectly with the description of lipids given above: they have a polar hydrophilic head group and a non-polar hydrophobic tail group. All of them are amphipathic lipids that self-assemble into a bilayer structure. One can guess that phospholipids got their name from the phosphate

group found in the head group. Two fatty acids are connected to glycerol and form the hydrocarbon chains. The phosphate is esterified to alcohol which then gives a further classification to the nomenclature of the phospholipid. Here you can notice PC (phosphatidylcholine) and PE (phosphatidylethanolamine) lipids.

Let us continue with the phospholipid classification. Now as a basic idea of the head group nomenclature is clear, we can move on to the classification based on the tail group of the lipid. Even more complexity is present in the hydrocarbon chain composition depending on the phospholipid class. There are different possibilities for the composition of the lipid chain but most phospholipids are composed of 1, 2, or 4 hydrocarbon chains. Fatty acids determine the non-polar hydrophobic part of the molecule. Glycerol is a binding group for both saturated and unsaturated fatty acids. Saturated fats are "saturated" with hydrogen atoms, which means they have the greatest number of hydrogen atoms possible and no double bonds in their chemical structure. As a consequence of unsaturated lipids here, we mean those that have one or more double bond(s) in the hydrocarbon chain. It is known that usually saturated lipids have higher melting temperatures than unsaturated ones. In biological membranes, it is common to have lipids with various tail lengths. For example, they can have fatty acids from as few as 12 carbons to as many as 26 carbons, and even longer. The number of double bonds also varies from zero to six depending on the lipid type. It is interesting to mention here that double bonds are usually not conjugated in the carbon tail but some exceptions can be present. In this work, we will focus on phospholipids with two symmetrical chains. Typically, lipid names starting with di- (D) indicate identical tail group. Another important group would be the lipids with different tails present in the molecule, such as 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC (16:0-18:1)) or 1-stearoyl-2-oleoyl-sn-glycero-3-phosphocholine (SOPC (18:0-18:1)). The POPC lipid had chains of different saturation and length, while SOPC had two chains of the same lenght, but one with unsaturation. Most bio-membranes consist of lipids with non identical tails.

There is a lot of discussions in medical and medically related fields about the impact of saturated and unsaturated fatty acids on the human body. Most health organizations advise limiting the consumption of saturated molecules (animal meat, palm oil, dairy products, fast food, etc.) and including more unsaturated fats (nuts, fish, olive oil, etc.) in the diet. It is believed that this can lower the risk of various heart diseases but the relationship between that and the consumption of saturated fats remains complex and new gatherings are made each day [79].

Let us introduce the nomenclature for lipids based on their level of unsaturation which is common for lipid sciences and will be used throughout the manuscript. As an example, we will look at the palmitic acid with 16 carbon atoms and no double bonds, and oleic

acids with 18 carbon atoms and one double bond (see DPPC and DOPC tail group at Fig. 1.4). In the given classification two numbers are used to represent a lipid tail group: X:Y, where X stands for the length of the fatty acid in carbon atoms and Y is the number of carbon-carbon double bonds present. Thus, in the case of DPPC we have 16:0 and for DOPC it is 18:1. It is interesting to mention that DOPE would also be labeled as 18:1 because no information on the head group of the lipid is included here so usually one can mark lipids as (18:1) PE to give full information. It might also be important to know at which position exactly a double bond is in the tail. The indication for that would be as follows: if the double bond is located between 5 and 6 carbons of the hydrocarbon chain we will see Δ_5 after the two numbers explained before. In the case of DOPC lipid that would mean 18:1 Δ_9. In the analogy, for the double unsaturated DLiPC one would have 18:2 $\Delta_{9,12}$. In fact, in mammalian cell membranes, the last two double bond positions are met the most often. More complex structures can be found in nature or created in the laboratory, for example, some bacteria cells can have cyclopropane rings in the fatty acid chains.

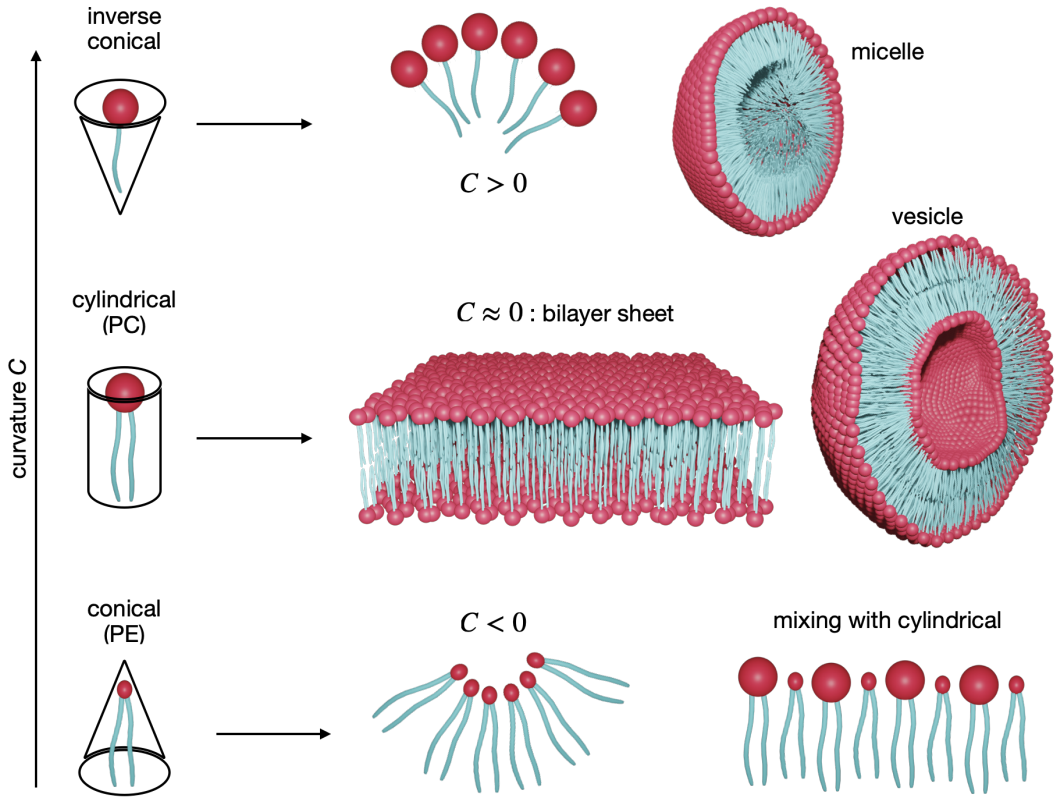


FIGURE 1.5: Schematic representation of different shapes that can be used to approximate lipids of various chemical structures.

As a physicist and not a chemist I would like to introduce here volume models for different lipid types [10]. In Fig. 1.5 one can see a schematic mapping of approximated volume to some lipids. For example, lipids with a symmetrical tail group and similar spacial

size of the head group can be approximated by a cylinder. It is natural to assume that the unsaturated lipids like DOPC (18:1) and DLiPC (18:2) are less rigid than saturated ones because the tails of such lipids are more flexible due to the presence of the double bonds. Then, the space required for the tail group of the unsaturated lipids would be a bit larger than for the saturated ones. The reason behind this is due to the possibility of kink formation in the carbon chain of the lipids with double bonds. On the other hand, lipids like DOPE, that have smaller head groups, can be approximated by a conical structure. Thus, we can assume that pure DOPE will give a negative curvature to the lipid monolayer. However, if mixed with lipids from the cylindrical group they might prefer to mix with each other and form a stable monolayer plane. Another interesting structure is a one-chain phospholipid, for example, a lysophospholipid. Due to the specific shape, such lipids form micelles. So it can be concluded that the balance of size between the head and tail groups has a strong influence on the membrane's spontaneous curvature and packing of the bilayer components. Though typically in the case of the bilayer system the overall curvature is balanced out by the effects in both monolayers, the influence of the lipid molecule shape is important when dealing with monolayer structures.

Let us describe each phospholipid that was involved in this work in more detail. 1,2-dipalmitoyl-sn-glycero-3-phosphatidylcholine, also known as DPPC, is a phospholipid containing the 16 carbon fatty acid called palmitic acid (16:0). Typically, this lipid is used to form various bilayers, liposomes, and other artificial membranes to study their properties. DPPC self-organizes into thin membranes due to its relatively short chains length. This lipids has a higher melting temperature compared to the lipids with unsaturation in their hydrocarbon chains.

The next group to be introduced here is the helper lipids. They tend to increase particle stability and fluidity of the lipid bilayer. In this work, we have used DOPC (18:1) and DOPE (18:1). DOPC, also known as 18:1 (Δ_9-Cis) PC, is a phospholipid containing the unsaturated long-chain (18:1) oleic acid in the 9-10 bond on both carbon chains of the tail group. As in the case of DPPC, it is used to form various artificial membranes in different forms in combination with other components or alone. DOPE, also known as 1,2-dioleoyl-sn-glycero-3-PE, is PE containing 18:1 fatty acids. The only difference between DOPC and DOPE is the head group of the lipid. DOPE can be used as an emulsifier to facilitate DNA-liposome complex transport across membranes. It is used in combination with cationic phospholipids to increase efficiency during DNA transfection studies as a non-viral method of gene delivery.

The last one to be discussed here is the 1,2-dilinoleoyl-sn-glycero-3-phosphatidylcholine (DLiPC 18:2). DLiPC is a phospholipid containing the unsaturated (18:2) linoleic acid at 9-10 and 12-13 bonds. The double bonds are placed consequently in the structure of this

phospholipid. By consequently here we mean that the unsaturation is still separated by one carbon atom. It also self-assembles in water and can be used to generate micelles, liposomes, and other types of artificial membranes.

As can be seen from Tables 1.1, 1.2 these lipids are ones of the most occurring in most biological cell membranes. That is why it is important to continue studying them both in pure systems and in combination with each other.

Sphingolipids

Another important group of lipids is the sphingolipids. They include phosphosphingolipids and glycosphingolipids. In the animal cell, their metabolism plays role in such diseases as Neimann-Pick's and Gaucher's disease both of which affect cell ability to generate fat. Furthermore, sphingolipids are known to take part in the signaling and cell recognition functions.

The phosphosphingolipids exhibit the same amphiphilic characteristics as the phospholipids but have different chemical structures. One can see an example of such a molecule in Fig. 1.4 (palmitoyl sphingomyelin, PSM). PSM has a phosphorylcholine head group, one fatty acid tail and sphingosine tail. Looking briefly at the chemical structure of this molecule one can assume that it is very similar to the DPPC lipid. However, a few differences can be noticed if we look closer. The most common form of the sphingomyelin core structure - sphingosine - is an 18 carbon chain, although other variations are also possible. The overall asymmetry of the two tails in this type of lipids is what makes them very distinguishable from the phospholipids. One of the two carbon chains is linked to the head group with an amide instead of an ester in case of PSM compared to DPPC. Another difference is the free hydroxyl group on the sphingosine chain which is a part of ceramide. The latter is in turn a core structure for both phosphosphingolipids and glycosphingolipids. It is not common for sphingolipids to have multiple unsaturation though there is a considerable variety of fatty acids in such molecules. Chain length is typically bigger than for phospholipids. Due to the amphipathic feature sphingomyelin also forms bilayers but has somewhat different behavior than regular phospholipids (higher compressibility as an example).

Sterols

The most common sterol found in mammalian membrane cells is Cholesterol. In Fig. 1.4 one can see the chemical structure of the Cholesterol molecule and notice how much more complex it is. This lipid consists of a hydrophilic polar hydroxyl group, steroid nucleus with four carbon rings, and a hydrocarbon side chain.

Cholesterol's role in the cell membrane is well-studied [36, 56] and it is known to influence membrane properties in various ways. The most prominent one is the creation of lipid rafts, meaning areas with higher packing inside fluid membrane areas. This process was not known until recently and is being studied actively now. It is believed that such clustering affects the membrane's ability to attach proteins and interact with a different type of stimulus. In mammalian cells, Cholesterol helps with cell growth and is necessary for the normal biological operation of the membranes.

Sterols are very specific to the type of cells and still, their functions are widespread in each. They are unique in their difference. For example, in yeast only ergosterol can be found, while the plant cell membrane is composed of sitosterol which is specific only to plants. Similarly, Cholesterol is specific to mammalian cells. In addition, no Cholesterol can be found in the plant cell membrane. Sitosterol and Cholesterol have both similar functions even though one is specific to plant and the other - to mammalian cells.

Membrane composition

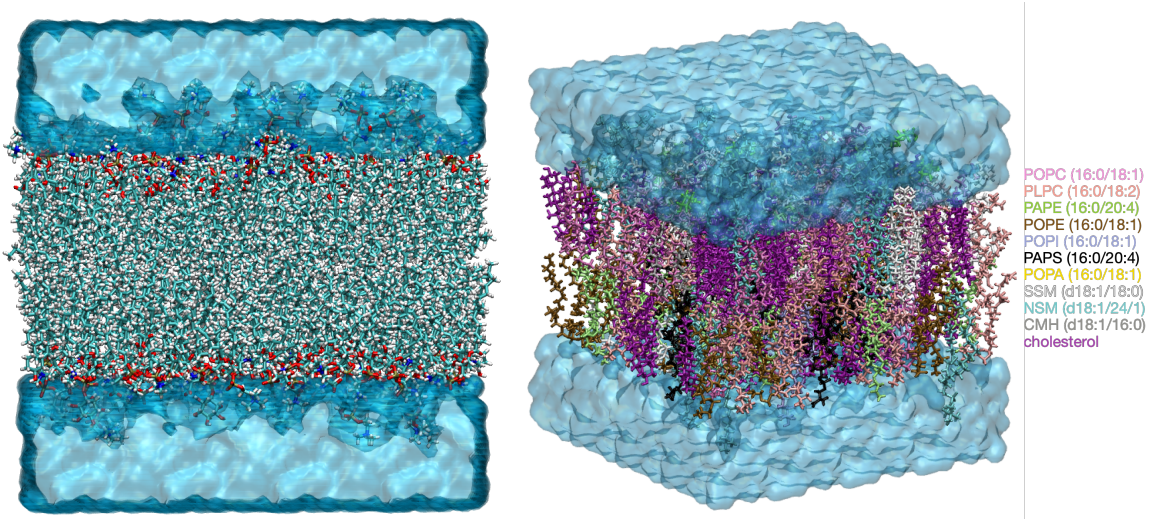


FIGURE 1.6: An example of an asymmetric mammalian plasma membrane (PMm) at $T = 310.15K$ in the water solvent. On the left snapshot, one can see a lipid membrane with a tail group shown in cyan and a head group in red (phosphorus) and blue (nitrogen). The right-hand side figure shows each lipid type present in the membrane in a different color. Water is shown as a blue surface. The snapshot was made with VMD using data from [60].

Even though phospholipids are only one class of the many lipids present in the cell membrane, it is already obvious how complex biological membranes can be considering all possible compositions. It is also interesting to mention here, that most biological membranes are asymmetric, meaning the composition of inner and outer monolayers can differ. In Tabs. 1.1, 1.2, 1.3 one can see different examples of the composition of lipids in the biological membranes. Fig. 1.6 illustrates how rich in different lipid types a mammalian

TABLE 1.1: The lipid composition as a molar percentage of the total lipids present in some biological membranes [63].

Membrane type | Component	Chol	PC	PE	SM	PS^-	PG^-
Bacterial inner membrane	0	0	80	0	0	20
Mammalian plasma membrane	50	20	11	13	6	0
Mammalian nuclear membrane	15	51	20	9	5	0

TABLE 1.2: Fatty acids found in egg phosphatidylcholine (PC) and human erythrocyte phosphatidylethanolamine (PE) [10].

Fatty acid	Mol% in egg PC	Mol% in erythrocyte PE
16:0	33	19
16:1	2	-
18:0	15	13
18:1	32	22
18:2	17.8	7
20:4	4.3	19
22:4	-	5

plasma membrane can be. As was mentioned before Cholesterol is present very little or not present at all in most membranes other than mammalian. In Tab. 1.1 and at Fig. 1.6 one can see that Cholesterol takes up to 50% of a mammalian plasma membrane while in a nuclei membrane only 15% of Cholesterol is typically found. This again is proving a fact that studying different Cholesterol concentrations and bilayer compositions is extremely important to understanding the mechanisms behind membrane dynamics and mechanics. One can also see from Tab. 1.1 that bacterial cells have no PC but are mostly a mix of PE and PG lipids. In Tab. 1.2 it is shown how diverse the composition of egg and erythrocyte membrane is. This yet again illustrates how complex membrane compositions are as they also depend on the environment and various internal and external factors. The egg membrane is mostly composed of mixtures of (16:0) and (18:1) PCs while the erythrocyte membrane has most of the (18:1), (16:0), and (20:4) PE lipids. It should be noted here that the composition of the chicken egg also highly depends on the species

TABLE 1.3: Fatty acid combinations of PC lipids found in egg [61].

Fatty acid 1	Fatty acid 2	Mol%
16:0	18:1	45
16:0	18:2	31
18:0	18:2	12
18:0	18:1	10
18:0	20:4	8

and nutritional state. From Tab. 1.3 one can see that the most observant mix of PC lipids in egg membrane is indeed (16:0) and (18:1), followed by (16:0) and (18:2) combination. Thus, one can conclude that those phospholipids of (16:0), (18:1) and (18:2) PC type are among the most widely spread in natural membranes.

It is possible to conclude that the phospholipid composition in a membrane is very specific to the class of lipids, the membrane function, the cell type, and many other factors. Based on the above examples we can generalize that the composition of membrane lipids is highly complex. Even though there is yet a lot to discover about each specific composition and lipid type involved in the functioning of a cell membrane, we can conclude that the maintenance of a suitable membrane environment is important and influenced by the fatty acid combinations of membrane lipids.

1.3 The importance of lipid interactions

In their monograph on lipid membrane biophysics, Mouritsen and Bagatolli devote a chapter to "Social lipids" [53]. This underlines how diverse, subtle and intricate the interactions between lipid molecules can be. As the Gibbs phase rule states, binary lipid membranes can display phase coexistence at room temperature and ambient pressure. The number of possible coexisting phases increases with the number of lipid species, and this number is quite large. Binary coexistence in binary lipid system is seen when one of the lipid is in the gel state and the other in the fluid state at a given temperature [27]. When the two lipids are in the same phase, phase separation originate from unfavorable mutual interactions, including chain length mismatch, difference in shapes, electrostatics, etc. On the other hand, Cholesterol molecule is very special. It cannot form bilayers alone but has a strong affinity for phospholipids with saturated lipid chains. Under appropriate conditions it triggers a fluid-fluid coexistence [83] referred as liquid disordered (L_d) -liquid ordered (L_o) coexistence.

The latter also true for ternary lipid systems associating low melting point unsaturated lipids, high melting point saturated lipids and cholesterol. Such systems are believed to show criticality for some selected compositions [24]. These ternary systems have become popular in relation with the raft hypothesis which claims that the plasma membrane is simultaneously asymmetric and heterogeneous, so it has different and non uniform leaflet composition. Furthermore, the outer membrane of eukaryotic cells is polarized due to the fact that apical (free upper part of the cell) and basal (attached lower part) surfaces are different, which results from the differential distribution of phospholipids, protein complexes, etc. between the various plasma membrane domains, and is specific to their

functions. These biological rafts are elusive, contrary to model lipid bilayers which display large binary domains. Such kind of domains are dynamic and not permanent in time or shape. The explanation for this kind of a short ranged-short lived heterogeneity is still missing, but for sure understanding mutual lipid interactions is a key ingredient. It is known by now that the chain length is important for domain formation in lipid binary mixtures.

A possible explanation is that membrane inclusions, especially membrane proteins can cluster together and recruit lipids that have some affinity with them. In other words domains only exist when proteins are present in sufficient amount. Building on this idea Gil et al [17] performed MC simulations demonstrating that under favorable conditions proteins can nucleate a phase whose composition differs from the average membrane and wets the protein rim (Fig. 1.7). This new phase can then bridge between neighboring proteins and nucleate a larger protein-lipid complex.

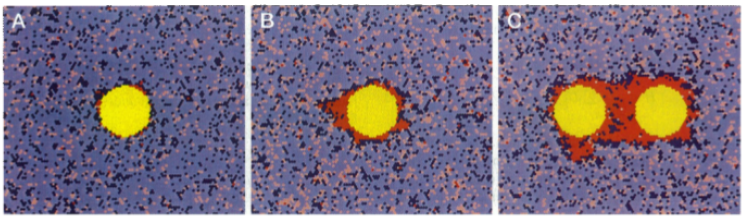


FIGURE 1.7: Bridging between two inclusions driven by the nucleation of an intermediate phase, MC simulations, Fig. 2 from Gil et al. [17].

The lateral bilayer structure is hard to characterize on the small scales in the range of nm and μm. As was said before, lipid membranes are not static objects with completely random organization and this also makes any direct determination of the lipid-lipid interactions more difficult. Furthermore, such kind of interaction can be considered slow as in experiments complete phase separation of the lipids is considered to be seen on the scales of hours. On the other hand, domain coexistence is observed for vesicles of tenth of micrometer diameter, while this can be considered large systems from the simulation point of view.

There are therefore many interesting phenomena taking place in a lipid membrane that requires a fine understanding of mutual lipid interactions. Experimental evidence are difficult to obtain and subject to large uncertainties [3]. Numerical simulations can help a lot with answering these questions. For this it is necessary to go beyond the simple approach consisting in counting the number of neighbors surrounding each lipid type, and to take into account the full thermodynamic complexity of lipid mixtures. The motivation of the current work mainly comes from the idea of understanding this phenomena in binary lipid mixtures.

1.4 Lipids in simulation

With all the variety available it is always good to start with a simple model. Even though a few model membranes were developed so far, we go back to the basics - the simplest phospholipid bilayer sheet. A building block of such a system would naturally be a lipid, meaning we have to model lipids and then let them organize into a bilayer structure. There are different ways to model a molecule. But first, let us understand why we need to create any models at all.

Computer simulations allow us to gain access to properties of a system of interest that are typically not or hardly accessible by real-life experiments. This can be the structural behavior of complex systems, surface tension, crystallization propagation, etc. Simulations grant us the possibility to study different phenomena from the inside of it. We can model an effect and predict the behavior of some system without performing an actual experiment to check if some theory is applicable or not. At the same time, computer studies can lead to new questions that can be checked by the experimentalist. The best part of working in a scientific environment is when one can organically connect both experimental and theoretical views on the problem and computer simulations aid a lot in that.

Let us focus on lipid simulations now. There are a few ways to model a molecule: all-atom (AA), united-atom (UA), and coarse-grained (CG). Each of them requires interaction potential to correctly describe the behavior of the molecules. AA model means each atom of the actual chemical structure is included in the simulation and all interactions should be taken into account. Such kinds of models have the highest resolution in the atom-atom interactions access but they require a lot of computational power to evaluate the interactions for thousands of particles at each time step. This means that in the case of AA systems we can only go up to some small time scales or very small systems, typically of the order of hundreds of nanoseconds and several nanometers. These limitations can be reduced with the use of coarse-graining. The idea of the CG procedure is rather simple - one has to generate simpler models that do not take into account specific atomistic details but still have enough description to keep crucial features of studied molecules realistically close to the reference data. Thus, CG gives an advantage of simplicity and allows to access longer time and length scales during a computer simulation.

Coarse-grained models are usually built on basis of the AA models with the reduction of the degrees of freedom. The first step of CG is to merge non-polar hydrogen atoms into their heavier neighbors and so UA models are built. However, it is not in the scope of current work and so further details will not be discussed here. The second step typically is to map 3-6 atoms into one heavier bead that interacts with neighboring beads via some simplified potential. Such interactions are modeled with specific force fields (FF) that try

to mimic real biological, chemical or physical characteristics of the given molecule. CG models provide information about the collective behavior of the molecules instead of detailed structural characteristics which are accessible with atomistic simulations [6]. It is important to mention here that in biological membranes collective phenomena of lipids are believed to have a stronger effect on the longer time scales than the detailed atomistic structure of lipids. Therefore, it is possible to reach longer time scales and larger dimensions using CG models which benefits the understanding of the processes and phenomena in biological systems.

CHAPTER 2

Theoretical background

We discuss in this chapter different theoretic approaches and concepts that are needed to study our bilayer systems. We review the concept of density in real and Fourier space, and density correlation functions when the bilayer is treated as a two dimensional fluid. We introduce the Kirkwood-Buff approach for mixtures and the theory of regular solutions. We mention the connection between lipid bilayers and a special instance of replicated fluid approach. Then we expose the basics of density linear response of a fluid to an externally applied bias potential, while membrane area compressibility is described on its own. Finally we expose some ways of analysing the main features of the 2d structure factors and attempt to model the effect of tilt and undulations on them.

The goal of the present approach is to derive the thermodynamical properties from the analysis of the projected lipid number density fluctuations of the bilayer sheets. The mixing properties of such systems are of particular interest to us. Let us discuss here the differences between our idea and more standard methods used to treat mixtures in chemical physics. Firstly, lipid molecules are comparably big objects (typically, 900 Daltons [1]) with many internal degrees of freedom (rotation of chemical groups, chain isomerization, etc.). Secondly, lipid bilayers are non-planar bidimensional structures fluctuating in their transverse directions. The bending modulus of such systems is large enough to maintain apparent planarity. However, it is still not enough to fully suppress all out-of-plane

[1] 1 Dalton $\approx 1.660 \cdot 10^{-27}$ kg

Helfrich fluctuations [68]. Our approach consists of neglecting the corrugation of the bilayer in the first place, thus representing a lipid bilayer sheet as an effective 2d fluid. This implies a "projection and reduction" step where a lipid molecule is reduced to an (x, y) pair of coordinates. However, information on the internal conformations and altitude z is lost in the process. On the other hand, we do check for and analyse the effect of the out-of-plane fluctuations. Also, we try to find an efficient protocol that would be useful to minimize the effects of the latter phenomena.

Another peculiar thing of interest is the bilayer structure of the membrane. Here we mean that the bilayer is made of two monolayers, and if the two have the same composition, one could assume them to be independent replicas. Although there is nowadays a strong interest in studying asymmetric membranes, the symmetric situation is also important to establish and define a methodology of the approach. Technically, the two leaflets can be considered as a mixture of two interacting leaflets. Thus, if their composition is identical, one can make use of the formalism of coupled replicated fluids to treat a bilayer. Such an approach simplifies bilayer systems' analysis, allowing twice as much statistics. A p component membrane is then described by a $2p$ components lipid mixture. To simplify even more, one can assume that the inter-leaflet interactions are weak. This can be tested locally utilising the pair correlation functions in real space. It is interesting to note here, that in the opposite limit, large-scale inter-leaflet correlations are likely to be caused by the undulations.

2.1 Density correlations in real and reciprocal space of a planar fluid

Let us assume that it is clear how to represent a lipid molecule by a pair of (x, y) coordinates. In Chapter 4 one can find various possibilities for this and their comparison. If we designate by $\mathbf{r}$ the real space coordinate, and $\mathbf{q}$ the reciprocal space coordinate, then thermodynamic properties emerge in the low $\mathbf{q}$ "hydrodynamic" regime. Those properties are expected to become insensitive to the details of the projection procedure [33]. Let us consider here a single leaflet, taken as a 2d fluid consisting of p species particles. To each species $\alpha = 1 \dots p$ one associates a density operator

$$n_\alpha(\mathbf{r}) = \sum_{j=1}^{N_\alpha} \delta(\mathbf{r} - \mathbf{r}_j) \tag{2.1}$$

built from a 2d Dirac delta distribution and the coordinates $\mathbf{r}_i$ of the N_α corresponding lipid positions. On the other hand, one can have a reciprocal space density operator as

follows:

$$\hat{n}_{\alpha,\mathbf{q}} = \sum_{j=1}^{N_\alpha} \exp\left(i\mathbf{q}\cdot\mathbf{r}_j\right) \tag{2.2}$$

The observable $\hat{n}_{\alpha,\mathbf{q}}$ can be directly computed from the simulations, provided $\mathbf{q}$ belongs to a set of vectors $(2\pi n_x/L_x, 2\pi n_y/L_y)$ commensurate with the simulation box size (L_x, L_y) and (n_x, n_y) integer numbers, where periodic boundary conditions are assumed.

Denoting with angular brackets $\langle\ \cdot\ \rangle$ the canonical average of the fluid leads to the reciprocal space correlations in the following form:

$$\langle \hat{n}_{\alpha,\mathbf{q}} \hat{n}_{\beta,-\mathbf{q}} \rangle \tag{2.3}$$

with $\beta = 1 \ldots p$. In the case of the identical indices $\alpha = \beta$ this correlation is a finite size approximation of $\tilde{h}_{\alpha\alpha}(\mathbf{q})$:

$$\frac{1}{N_\alpha} \langle \hat{n}_{\alpha,\mathbf{q}} \hat{n}_{\alpha,-\mathbf{q}} \rangle = 1 + \rho_\alpha \tilde{h}_{\alpha\alpha}(\mathbf{q}) \tag{2.4}$$

with $\rho_\alpha = \frac{N_\alpha}{L_x L_y}$ and $\tilde{h}_{\alpha\alpha}$ the unbounded Fourier transform of the short range part of the pair correlation function $h_{\alpha\alpha}(\mathbf{r}) = g_{\alpha\alpha}(\mathbf{r}) - 1$, with the convention

$$\tilde{h}(\mathbf{q}) = \int_{-\infty}^{\infty} \mathrm{d}x \int_{-\infty}^{\infty} \mathrm{d}y\, h(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}} \tag{2.5}$$

A similar relation connects the $\alpha \neq \beta$ case

$$\frac{L_x L_y}{N_\alpha N_\beta} \langle \hat{n}_{\alpha,\mathbf{q}} \hat{n}_{\beta,-\mathbf{q}} \rangle = \tilde{h}_{\alpha\beta}(\mathbf{q}) \tag{2.6}$$

In a continuous unbounded space, the $q \to 0$ limit of the $\tilde{h}$ functions coincide with the unbounded integrals

$$G_{\alpha\beta} = \int_{-\infty}^{\infty} \mathrm{d}x \int_{-\infty}^{\infty} \mathrm{d}y\, h_{\alpha\beta}(\mathbf{r}) \tag{2.7}$$

These quantities are the Kirkwood-Buff (KB) integrals, and are related to some fundamental thermodynamic properties of the fluid [5, 30].

2.2 Kirkwood-Buff integral and thermodynamics of a pure fluid

The KB integral G_{11} of a pure fluid is related to its area compressibility (or volume compressibility for the 3 dimensional case). The connection appears when dealing with the system in the grand-canonical ensemble. The number of particles N evolves as the chemical potential μ is varied. Considering a finite area system (possibly with periodic boundary conditions) one can derive

$$\left.\frac{\partial\langle N\rangle}{\partial\mu}\right|_{A,T} = \frac{\langle N^2\rangle - \langle N\rangle^2}{k_B T} = -\left.\frac{\partial^2\mathcal{J}}{\partial\mu^2}\right|_{A,T} \tag{2.8}$$

where $\mathcal{J}$ represents the grand potential equal to $-PA$ (pressure times area), $k_B = 1.380649 \cdot 10^{-23}$ J K^{-1} is the Boltzmann constant, T is the temperature. On the other hand, the derivative $\partial\langle N\rangle/\partial\mu$ can be related to the KB integral as follows (see Appendix A):

$$k_B T\frac{\partial\langle N\rangle}{\partial\mu} = \langle N\rangle + \frac{\langle N\rangle^2}{A}G_{11} \tag{2.9}$$

Furthermore, by virtue of the Legendre transform

$$\mathcal{J} + \mu N = \mathcal{F} \tag{2.10}$$

the right hand side of the eq. 2.8 can be related to a derivative of the Helmholtz free-energy $\mathcal{F}(N, A, T)$:

$$\left.\frac{\partial^2\mathcal{J}}{\partial\mu^2}\right|_{A,T} \times \left.\frac{\partial^2\mathcal{F}}{\partial N^2}\right|_{A,T} = -1 \tag{2.11}$$

The Helmholtz function is the potential of the canonical ensemble, meaning a constant N ensemble. $\mathcal{F}$, A and N being the extensive parameters, one has $\mathcal{F}(N, A, T) = A\mathcal{F}(N/A, T)$. The derivatives of $\mathcal{F}$ with respect to N at constant A and with respect to A at constant N are related. As a consequence, one can define a tension σ (a so-called opposite of the 2d "pressure") as follows:

$$\sigma = \left.\frac{\partial\mathcal{F}}{\partial A}\right|_{N,T} \tag{2.12}$$

Next, a compressibility coefficient χ_T can be defined:

$$\begin{aligned}\chi_T &= \frac{1}{A}\left(\frac{\partial A}{\partial \sigma}\right)_{N,T} \\ &= \frac{1}{A(\partial\sigma/\partial A)_{N,T}} \\ &= \frac{A}{N^2}\frac{1}{(\partial^2\mathcal{F}/\partial N^2)} \\ &= -\frac{A}{N^2}\frac{\partial^2\mathcal{J}}{\partial\mu^2} \\ &= \frac{1+\rho G_{11}}{\rho k_B T}\end{aligned} \tag{2.13}$$

leading to the compressibility relation for pure systems:

$$1+\rho G_{11} = \rho k_B T \chi_T \tag{2.14}$$

In membrane biophysics, one often uses the area elastic coefficient $K_A = 1/\chi_T$, homogeneous to a surface tension. Typically, a magnitude of K_A is of the order of 200 mN/m. A few other methods to obtain the area elastic modulus will be discussed later in this chapter.

2.3 Structure factors and density fluctuations

Eq. 2.4 establishes a relation between the second moment of the density fluctuations and a Fourier space observable $1+\rho\tilde{h}$. For pure systems, this expression is the structure factor of the fluid

$$S(\mathbf{q}) = 1+\rho\tilde{h}(\mathbf{q}) \tag{2.15}$$

The structure factor [91] plays a central role in the X-rays or neutron elastic scattering of the 3d fluids and can be measured experimentally.

In the case of the lipid bilayers, $S(\mathbf{q})$ is not directly measurable, as it corresponds to the distribution of the fictitious points representing the lipids. Current scattering experiments on bilayers are only sensitive to the normal z distribution of the lipid molecular subgroups transverse to the bilayer, but do not provide the information on the lateral x, y distribution of the molecules. Nevertheless, in the current project we attempt to define and compute numerically structure factors as they play a central role in many theoretical models of the 2d fluids.

The low $\mathbf{q}$ region corresponds to the hydrodynamic modes of the bilayer. In this regime,

density fluctuations are expected to obey a Gaussian statistics, with small $\delta\rho/\rho$ amplitude and vanishing average. Such fluctuations are entirely controlled by the value of the structure factor $S(\mathbf{q})$.

$$\langle \hat{n}_{\mathbf{q}} \hat{n}_{-\mathbf{q}} \rangle = NS(\mathbf{q}) \tag{2.16}$$

In the thermodynamic limit at given $q = |\mathbf{q}|$, $S(q)$ is well defined and isotropic, as is the pair correlation $g(r)$. It does not depend on the ensemble and can be computed in the canonical ensemble. It controls the magnitude of the fluctuations of $\hat{n}_{\mathbf{q}}$. In the grand canonical ensemble, the $q \to 0$ limit is regular and one can define a zero-mode $\hat{n}_0 = N$, which is equal to the total fluctuating number of molecules. In the canonical ensemble this mode is constant and non fluctuating. However, the $q \to 0$ limit of the structure factor exists and is assumed to be finite. This means that the fluctuation in the number of particles can be inferred from a canonical ensemble calculation, provided one manages to approximate accurately $S(q)$ in the hydrodynamic regime, and extrapolate it to the $q = 0$ limit.

Our current approach has an operational advantage compared to an approach that would rely directly on computing the integral of a pair distribution coefficient. The latter would depend on how fast $g(r) - 1$ decreases to 0, and would in turn require a lengthy convergence of a statistical histogram involving all possible pairs in the system. Here, the observable of interest, $\hat{n}_{\alpha,\mathbf{q}}$, can be computed with a single loop over the particles, without recursion over pairs of neighboring particles. Finite size effects in our case come from the lower boundary of the accessible wave vectors $q_{\min} = 2\pi/L_x \sim 2\pi/A^{1/2}$ and for a usual fluid they decrease as the simulation box size increases.

Our numerical approximation on the left hand side of eq. 2.4 is assumed to approximate well the theoretical structure factor on the right hand side that occurs in the thermodynamic limit, up to finite size and convergence time effects. Finite q hydrodynamic density modes do not modify the number of particles present in the system, and fluctuate in a statistically independent way. This is why the canonical or grand-canonical ensembles should agree in the thermodynamic limit regarding the value of $\langle \hat{n}_{\alpha,\mathbf{q}} \hat{n}_{\beta,-\mathbf{q}} \rangle$.

In the case of mixtures, generalization of the structure factors associated with the Gaussian fluctuations of the density modes $\hat{n}_{\alpha,\mathbf{q}}$ can be related in a similar way to the KB integrals $G_{\alpha\beta}$ and further thermodynamic quantities of interest.

2.4 Multicomponents Kirkwood-Buff relations

Let us consider a p-component fluid in the grand-canonical ensemble, N_α and μ_α, $\alpha = 1 \dots p$, are the number of molecules and chemical potentials of each species correspondingly. Then the eqs. 2.8 and 2.9 become

$$\left.\frac{\partial \langle N_\alpha \rangle}{\partial \mu_\beta}\right|_{A,T} = \frac{\langle N_\alpha N_\beta \rangle - \langle N_\alpha \rangle \langle N_\beta \rangle}{k_B T} = -\left.\frac{\partial^2 \mathcal{J}}{\partial \mu_\alpha \partial \mu_\beta}\right|_{A,T} \tag{2.17}$$

$$k_B T \frac{\partial \langle N_\alpha \rangle}{\partial \mu_\beta} = \langle N_\alpha \rangle \delta_{\alpha\beta} + \frac{\langle N_\alpha \rangle \langle N_\beta \rangle}{A} G_{\alpha\beta} \tag{2.18}$$

The KB integrals $G_{\alpha\beta}$ appear therefore linked to the second derivatives $\partial^2 \mathcal{J} / (\partial \mu_\alpha \partial \mu_\beta)$. As previously, the grand potential $\mathcal{J}$ and the Helmholtz free-energy are mutually Legendre transformed and, as a consequence, are related.

$$\mathcal{J} + \sum_{\alpha=1}^{p} N_\alpha \mu_\alpha = \mathcal{F} \tag{2.19}$$

where $\mathcal{F}(T, A, N_\alpha)$ depends naturally on the temperature T, area A (for the 2d case) and fixed number of particles N_α. The Legendre transform structure implies the following matrix-like relation:

$$\sum_{\gamma=1}^{p} \left(\frac{\partial^2 \mathcal{J}}{\partial \mu_\alpha \partial \mu_\gamma}\right)\Bigg|_{A,T} \left(\frac{\partial^2 \mathcal{F}}{\partial N_\gamma \partial N_\beta}\right)\Bigg|_{A,T} = -\delta_{\alpha\beta} \tag{2.20}$$

This relation formally connects the KB integrals ($\mathcal{J}$ terms) to the constant area derivatives of the chemical potential with respect to the number of particles

$$\left(\frac{\partial^2 \mathcal{F}}{\partial N_\gamma \partial N_\beta}\right)_{T,A,N_\gamma} = \left.\frac{\partial \mu_\gamma}{\partial N_\beta}\right|_{T,A,N_{\beta\neq\gamma}} \tag{2.21}$$

The usual KB theory aims at expressing the derivative of the chemical potentials with respect to the number of particles at constant pressure (or tension in the 2d case). In addition to inverting the matrix 2.20 this requires the use of a number of thermodynamic identities which involve the partial molar volumes (or molar areas in 2d case) and the systems' isothermal compressibility [5, 30]. In the scope of the current project, the above arguments can be summarized as follows:

$$\left.\frac{\partial \mu_\alpha}{\partial N_\beta}\right|_{T,\sigma,N_{\beta\neq\alpha}} = \left.\frac{\partial \mu_\alpha}{\partial N_\beta}\right|_{T,A,N_{\beta\neq\alpha}} - \frac{\mathcal{A}_\alpha \mathcal{A}_\beta}{A \chi_T} \tag{2.22}$$

with the partial molar areas defined by

$$\mathcal{A}_\alpha = -\left.\frac{\partial \mu_\alpha}{\partial \sigma}\right|_{T,\sigma,N_\beta} \tag{2.23}$$

Expressions in the multicomponent case are quite complicated because of the matrix inversion step. They are fortunately convenient in the binary case when $p = 2$. The central quantity of interest in this case would be $\mu_{12} = \partial\mu_2/\partial N_1|_{T,\sigma,N_2}$. The two other derivatives $\mu_{22} = \partial\mu_2/\partial N_2$ and $\mu_{11} = \partial\mu_1/\partial N_1$ can be expressed in terms of μ_{12} by means of the Gibbs-Duhem relation at constant temperature and tension

$$\begin{aligned} N_1 \mathrm{d}\mu_1 + N_2 \mathrm{d}\mu_2 &= (N_1\mu_{11} + N_2\mu_{12})\mathrm{d}N_1 + (N_1\mu_{12} + N_2\mu_{22})\mathrm{d}N_2 \\ &= 0 \end{aligned} \tag{2.24}$$

The result relations, using Ben-Naim notations [5] is as follows:

$$\begin{aligned} \Delta &= G_{11} + G_{22} - 2G_{12} \\ \eta &= \rho_1 + \rho_2 + \rho_1\rho_2\Delta \\ \zeta &= 1 + \rho_1 G_{11} + \rho_2 G_{22} + \rho_1\rho_2(G_{11}G_{22} - G_{12}^2) \\ \mu_{12} &= -\frac{k_B T}{A\eta} \\ \chi_T &= \frac{\zeta}{k_B T \eta} \end{aligned} \tag{2.25}$$

The above expressions give the fluid compressibility and the derivatives of the chemical potentials in terms of the average densities and the KB integrals. They are general and do not make assumptions on the nature of the intra and inter molecular interactions. These relations describe complex molecular fluids and do not depend in the thermodynamic limit on the arbitrary choice made for representing a given molecule.

2.5 Thermodynamics of regular solutions

The theory of regular solutions [4] is a simplified model for the Gibbs free-energy of a p-mixture of components, assuming that all pure components are found in the same "state" or "phase". It provides a simplified expression for the mixing part G^{mix} of the Gibbs free-energy defined as the difference between the mixed and separated states

$$G^{\mathrm{mix}}(T,\sigma,N_\alpha) = G(T,\sigma,\ldots N_\alpha \ldots) - \sum_\alpha G(T,\sigma,N_\alpha) \tag{2.26}$$

Extensivity allows to introduce molar fractions and write

$$\begin{aligned} N &= \sum_\alpha N_\alpha \\ x_\alpha &= \frac{N_\alpha}{N} \\ G^{\text{mix}}(T,\sigma,N_\alpha) &= NG^{\text{mix}}(T,\sigma,x_\alpha) \end{aligned} \tag{2.27}$$

The mixing free-energy is then approximated as

$$\frac{G^{\text{mix}}(T,\sigma,N_\alpha)}{k_BTN} = \sum_\alpha x_\alpha \ln(x_\alpha) + \sum_{\alpha,\beta} B_{\alpha\beta}x_\alpha x_\beta \tag{2.28}$$

Expression 2.28 is the sum of the ideal entropy of mixing term and a non ideal quadratic approximation term. In the binary case, one can write $x = N_1/(N_1+N_2)$ and

$$\frac{G^{\text{mix}}(T,\sigma,N_1,N_2)}{k_BTN} = x\ln(x) + (1-x)\ln(1-x) + Bx(1-x) \tag{2.29}$$

with $x = x_1$ and $1-x = x_2$ - concentrations of 1 and 2 components correspondingly.

This approximated expression provides explicit relations for the chemical potentials μ_1, μ_2 and the derivative μ_{12} [57], thus, establishing a connection between the theory of regular solutions and the Kirkwood-Buff integrals.

$$\begin{aligned} \mu_1(T,\sigma,x) &= \mu_1^{(0)} + k_BT(\ln(x) + B(1-x)^2) \\ \mu_2(T,\sigma,x) &= \mu_2^{(0)} + k_BT(\ln(1-x) + Bx^2) \\ \mu_{12} &= \frac{k_BT}{N}(-1+2Bx(1-x)) \end{aligned} \tag{2.30}$$

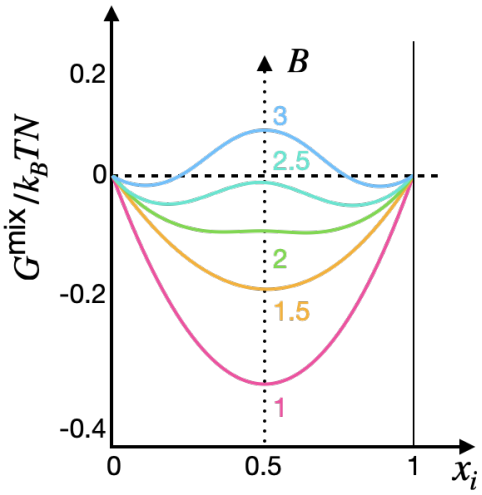


FIGURE 2.1: Illustration of the Gibbs energy of mixing for different values of the mixing parameter B.

Our objective is to estimate the coefficient B from the Molecular Dynamics simulations and eq 2.30 provides the desired connection. The final expression for the coefficient B using eq 2.25 reads as

$$B = \frac{\rho}{2}\left[\frac{\Delta}{1+\rho x(1-x)\Delta}\right] \tag{2.31}$$

The mixing parameter is a measure of the non ideality of the mixture. A positive B means that the interaction between similar molecules 1-1 or 2-2 are more favorable than the interaction between dissimilar molecules 1-2. A negative B corresponds to the opposite

situation. When the mutual interaction is too unfavorable, the system spontaneously separates into two phases, one enriched in species 1, the other in species 2. This occurs at a critical value $B_c = 2$ for equimolar mixtures $x = 0.5$. The critical value depends on x and becomes larger as x gets close to 0 or 1. For $x \neq 0.5$ there are in fact two consecutive critical values. The first one corresponds to the possibility of binary coexistence (binodal curve), the second one to the limit of stability of the single phase state (spinodal curve) (see Fig. 2.1).

2.6 Replicas and weakly coupled leaflets

The Kirkwood-Buff approach introduced above is valid for any 2d molecular fluid. This approximation should hold and also be applicable to a flat single leaflet provided one neglects all interaction with the opposite leaflet. It is also assumed that water is in sufficient amount so that the influence of the solvent can be effectively reduced to an implicit interaction between lipids (full hydration case).

In the absence of the interaction between leaflets, one can consider the pure bilayer as a mixture of two non-interacting fluids. If p lipid species are present, then the full bilayer becomes a $2p$ mixture. The current paragraph explains how to reconcile the single leaflet and the full bilayer description of a fluid. Let us remind to the reader, that it is assumed that the fluid is projected on a 2d flat space of coordinates xy. One introduces the monolayer density ρ_m and the bilayer density $\rho_b = 2\rho_m$. Similarly, $\rho_{\alpha,m}, \rho_{\alpha,b}$ are the mono and bilayer densities of species α, while N_α is the number of particles α in a monolayer.

The correlation functions $\tilde{h}_{\alpha\beta}(\mathbf{q})$ acquire a tensor-like structure, with additional indices u, l indicating which leaflet they pertain to. A general approach developed in the context of glassy and disordered fluids can be useful here [47, 49, 58]. In the scope of the current thesis, only a very particular case with two replicas $u, l = 1 \ldots 2$ is needed. Replica symmetry here implies that the leaflets share an identical composition. The generalized correlations $\tilde{h}_{\alpha\beta,ul}(\mathbf{q})$ can only take two values, depending on whether $u = l$ or $u \neq l$, namely

$$\tilde{h}_{\alpha\beta,ul}(\mathbf{q}) = \delta_{ul}\tilde{h}_{\alpha\beta}(\mathbf{q}) + (1 - \delta_{ul})\tilde{h}'_{\alpha\beta}(\mathbf{q}) \tag{2.32}$$

Strictly non-interacting leaflets then correspond to the situation $\tilde{h}'_{\alpha\beta}(\mathbf{q}) = 0$. The compressibility of the bilayer is then half of the compressibility of each monolayer.

If instead of each monolayer density mode $\hat{n}_{\alpha,u,\mathbf{q}}$ one aggregates both leaflets ($u = 1$ and $l = 2$) together, it can be shown that :

$$\langle(\hat{n}_{\alpha,1,\mathbf{q}} + \hat{n}_{\alpha,2,\mathbf{q}})(\hat{n}_{\beta,1,\mathbf{q}} + \hat{n}_{\beta,2,\mathbf{q}})\rangle = 2N_\alpha\delta_{\alpha\beta} + \frac{(2N_\alpha)(2N_\beta)}{A}\frac{\tilde{h}_{\alpha\beta}(\mathbf{q}) + \tilde{h}'_{\alpha\beta}(\mathbf{q})}{2} \tag{2.33}$$

with N_α,N_β the number of α, β lipids per leaflet.

Taking at once all the lipids in a bilayer amounts to formally replacing $\tilde{h}_{\alpha\beta}$ by $(\tilde{h}_{\alpha\beta} + \tilde{h}'_{\alpha\beta})/2$. If the leaflets are non-interacting, $\tilde{h}_{\alpha\beta}$ is replaced by $\tilde{h}_{\alpha\beta}/2$. Separate leaflets correlations and bilayers correlations then contain the same information and both allow for the determination of the central quantity $\tilde{h}_{\alpha\beta}(\mathbf{q})$.

The presence of the inter leaflet correlations can independently be assessed by a real-space cross-leaflet correlation function $g'_{\alpha\beta}(r)$. Any short-range deviation of $g'_{\alpha\beta}(r) - 1$ can be interpreted as the manifestation of some "cross-talk" between leaflets. Even in the absence of a direct cross-leaflet interactions, long wavelength undulation fluctuations are expected to induce some amount of $u \neq l$ correlation in $\tilde{h}'_{\alpha\beta}(\mathbf{q})$. It is therefore preferable to reduce as much as possible the undulation in such systems.

2.7 Linear response in density of a fluid

Linear response in density of a fluid is naturally formulated in the grand-canonical ensemble. Let us consider a pure fluid of N point-like particles that interact by means of a generic potential energy function $U_{\text{int}}(\{\mathbf{r}_j\})$. This potential depends on the configuration $\{\mathbf{r}_j\}$. By configuration here we mean a set of particles positions. Let us consider an external 1-body potential U_{ext} that is being added to the system and has the following form :

$$U_{\text{ext}} = \sum_{j=1}^{N} \phi(\mathbf{r}_j) \tag{2.34}$$

with $\phi(\mathbf{r}_j)$ being a potential that acts on jth particle only.

Assuming we work in the grand-canonical ensemble so the total number of particles N may vary while the chemical potential μ is held constant, we can define the density observable as follows :

$$n(\mathbf{r}) = \sum_{j=1}^{N} \delta(\mathbf{r} - \mathbf{r}_j) \tag{2.35}$$

where $\mathbf{r}$ points to an arbitrary space position. By definition of $n(\mathbf{r})$ the potential energy reads

$$U_{\text{ext}} = \int \mathrm{d}\mathbf{r}\, n(\mathbf{r})\phi(\mathbf{r}) \tag{2.36}$$

Then the average particle density becomes inhomogeneous and reads

$$\rho(\mathbf{r}) = \langle n(\mathbf{r}) \rangle \tag{2.37}$$

with $\langle \cdot \rangle$ denoting the grand-canonical average. It is interesting to note that in general case the average field $\rho(\mathbf{r})$ is non uniform and the precise relation between average density and external field is non trivial.

However, if the field is small enough, the dependence of ρ on the field is linear, and can be given by the following expression:

$$\rho(\mathbf{r}) = \int \mathrm{d}\mathbf{r}' \, R(\mathbf{r} - \mathbf{r}')\phi(\mathbf{r}') \tag{2.38}$$

for a pure system, and

$$\rho_\alpha(\mathbf{r}) = \sum_\beta \int \mathrm{d}\mathbf{r}' \, R_{\alpha\beta}(\mathbf{r} - \mathbf{r}')\phi_\beta(\mathbf{r}') \tag{2.39}$$

for a mixture, where the external potential ϕ_β is only exerted on the molecules of type β. The susceptibilities R, $R_{\alpha\beta}$ are related to the density mode correlations and therefore to the structure factors (see Appendix A for more detail).

This suggests an alternative possibility for obtaining structure factors from simulations. We specialize eq. 2.38 to the case of a harmonic external potential

$$U(x) = U_0 \cos(q_x x), \tag{2.40}$$

where $q_x = \frac{2\pi}{L}$ - "quantum number", U_0 - amplitude of the potential, x - coordinates of the lipid molecules. The potential is not applied to water molecules. This potential can alternatively be seen as an external applied force $F(x)$ onto the particles of the system

$$F(x) = -\frac{\mathrm{d}U(x)}{\mathrm{d}x} = U_0 q_x \sin(q_x x) \tag{2.41}$$

In this case, the linear modulation in the density of the system can be expressed as a response to the field

$$\rho(\mathbf{r}) = \rho_0 + \rho_1 \cos(q_x x), \tag{2.42}$$

where $\rho_0 = \frac{N}{LxL_y}$ is the equilibrium density of the system and ρ_1 is the magnitude of the Fourier coefficient of the inhomogeneous fluid. It turns out that this amplitude ρ_1 can be approximated directly in a numerical simulation by the average of the following discrete

sum :

$$\begin{aligned}\rho_1 &= \frac{2}{L_x L_y}\int_{\mathcal{S}} \mathrm{d}\mathbf{r}\cos(q_x x)\rho(\mathbf{r}) \\ &= \frac{2}{L_x L_y}\left\langle \int_{\mathcal{S}} \mathrm{d}\mathbf{r}\cos(q_x x)\sum_{j=1}^{N}\delta(\mathbf{r}-\mathbf{r}_j)\right\rangle \\ &= \frac{2}{L_x L_y}\langle\sum_{j=1}^{N}\cos[q_x x_j(t)]\rangle \end{aligned} \tag{2.43}$$

ρ_1 density amplitude dependent on the external potential, N - number of lipid molecules, L_x, L_y - characteristic size of the system in the x and y directions, $\mathcal{S}$ the integration domain of the repeated unit cell $[0, L_x] \times [0, L_y]$. Only wave vectors $\mathbf{q}$ commensurate with periodic boundary conditions are admissible. Comparing with expression 2.38 one finds

$$\begin{aligned}\rho_0 + \rho_1\cos(q_x x) &= \int \mathrm{d}\mathbf{r}'\, R(\mathbf{r}-\mathbf{r}')U_0\cos(q_x x') \\ &= U_0\int \mathrm{d}\mathbf{r}'\, R(\mathbf{r}'-\mathbf{r})\mathrm{Re}\left[e^{i\mathbf{q}\cdot(\mathbf{r}'-\mathbf{r})+iq_x x}\right] \\ &= U_0\mathrm{Re}[\tilde{R}(\mathbf{q})e^{iq_x x}] = U_0\tilde{R}(q)\cos(q_x x)\end{aligned} \tag{2.44}$$

where we have used that $R(\mathbf{r}'-\mathbf{r})$ is an even function and $\tilde{R}(q)$ a real quantity.

As is established in Appendix A, $\tilde{R}(q)$ is nothing but $-\frac{\rho_0}{k_B T}S(q)$, with ρ_0 the homogeneous equilibrium density and $S(q)$ the equilibrium structure factor of the pure fluid. Therefore, we found that in the case of an external harmonic (sine) potential, the structure factor is related to the amplitude of the cosine modulation of the density ρ_1, namely

$$S(\mathbf{q}) = -\frac{2k_B T}{U_0}\langle\frac{1}{N}\sum_{j=1}^{N}\cos[\mathbf{q}\cdot\mathbf{r}_j]\rangle = -\frac{k_B T}{U_0}\frac{\rho_1}{\rho_0}, \tag{2.45}$$

where it is assumed that $|\frac{U_0(x)}{k_B T}| \ll 1$. The structure factor appears as a wavelength dependent compressibility coefficient. It is not surprising that $S(q=0)$ is in the end related to the system compressibility.

Relation 2.45 provides an alternative route to the determination of the structure factor. If the usual approach is based on the equilibrium density fluctuations, the latter method consists in measuring the average value of such mode in the presence of a small external biasing potential. It is more direct as measures of average values are often easier than measures of correlations. The instantaneous amplitude $\hat{n}_{\mathbf{q}}(t)$ (right hand side of eq. 2.43) prior to averaging is a fluctuating quantity that evolves in a quadratic potential well whose stiffness is set by the inverse structure factor $S(\mathbf{q})^{-1}$. The presence of a biasing

potential U_0 displaces the equilibrium towards a new nonvanishing value ρ_1. The smaller U_0 is, the longer it takes for the sampled time average ρ_1 to emerge from the spontaneous fluctuating noise. A compromise must be found between keeping U_0 large to increase the measured ρ_1 while remaining in the linear response regime.

The fluctuating mode $\hat{n}_{\mathbf{q}}(t)$ has an intrinsic relaxation time resulting from the stiffness of the "trap" confining its fluctuations and the kinetic coefficient associated with the speed at which the mode relaxes towards equilibrium. The kinetic coefficient is here the collective diffusion coefficient D_c and in the low $\mathbf{q}$ hydrodynamic regime, the relaxation time $\tau(\mathbf{q})$ is expected to follow a behaviour $\tau(\mathbf{q}) = S(\mathbf{q})/(D_c\mathbf{q}^2)$ assuming that collective density fluctuations are dissipatively damped. This relaxation time quickly diverges as $\mathbf{q} \to 0$. Getting a well defined average value ρ_1 or a good correlation estimate $\langle \hat{n}_{\mathbf{q}} \hat{n}_{-\mathbf{q}} \rangle$ requires sampling on a time interval much longer than $\tau(\mathbf{q})$.

The linear response approach allows for some multiplexing $U_1 \cos\left(q_x^{(1)} x\right) + U_2 \cos\left(q_x^{(2)} x\right) + U_3 \cos\left(q_x^{(3)} x\right) + \ldots$ of the applied potential (or force). The corresponding reading of mode amplitudes $\rho_1^{(1)}$, $\rho_1^{(2)}$, $\rho_1^{(3)}$ can be done simultaneously.

2.8 Linear response and structure factors for mixtures

Density linear response in the case of a lipid mixture consists in applying a set of cosine potentials $U_\beta \cos(q_x x)$ to each species β and record the corresponding average harmonic density modulation $\rho_\alpha(\mathbf{r}) = \rho_{0,\alpha} + \rho_{1,\alpha} \cos(q_x x)$ from species α.

$$\rho_{1,\alpha} = -\frac{1}{k_B T} \sum_\beta \left[\delta_{\alpha\beta} \rho_{0,\alpha} + \rho_{0,\alpha} \rho_{0,\beta} \tilde{h}_{\alpha\beta}(\mathbf{q}) \right] U_\beta \tag{2.46}$$

Density linear response is entirely controlled by the continuous quantities $\tilde{h}_{\alpha\beta}(\mathbf{q})$ which can in turn be approximated numerically for commensurate wave vectors $\mathbf{q}$ by our simulation box density modes correlations $\langle \hat{n}_{\alpha,\mathbf{q}} \hat{n}_{\beta,-\mathbf{q}} \rangle$.

It is also possible to introduce generalized structure factors for mixtures. For instance

$$S_{\alpha\beta}(\mathbf{q}) = \delta_{\alpha\beta} + \sqrt{\rho_\alpha \rho_\beta} \tilde{h}_{\alpha\beta}(\mathbf{q}) \tag{2.47}$$

or explicitly in terms of numerical discrete density modes

$$S_{\alpha\beta}(\mathbf{q}) = \frac{1}{\sqrt{N_\alpha N_\beta}} \left\langle \hat{n}_{\alpha,\mathbf{q}} \hat{n}_{\beta,-\mathbf{q}} \right\rangle = \frac{1}{\sqrt{N_\alpha N_\beta}} \left\langle \sum_{j=1}^{N_\alpha} e^{i\mathbf{q}\mathbf{r}_{\alpha,j}} \sum_{j=1}^{N_\beta} e^{-i\mathbf{q}\mathbf{r}_{\beta,j}} \right\rangle$$
$$= \frac{1}{\sqrt{N_\alpha N_\beta}} \left\langle \sum_{j=1}^{N_\alpha} \cos(\mathbf{q}\cdot\mathbf{r}_{\alpha,j}) \sum_{j=1}^{N_\beta} \cos(\mathbf{q}\cdot\mathbf{r}_{\beta,j}) + \sum_{j=1}^{N_\alpha} \sin(\mathbf{q}\cdot\mathbf{r}_{\alpha,j}) \sum_{j=1}^{N_\beta} \sin(\mathbf{q}\cdot\mathbf{r}_{\beta,j}) \right.$$
$$\left. + i \sum_{j=1}^{N_\alpha} \sin(\mathbf{q}\cdot\mathbf{r}_{\alpha,j}) \sum_{j=1}^{N_\beta} \cos(\mathbf{q}\cdot\mathbf{r}_{\beta,j}) - i \sum_{j=1}^{N_\alpha} \cos(\mathbf{q}\cdot\mathbf{r}_{\alpha,j}) \sum_{j=1}^{N_2} \sin(\mathbf{q}\cdot\mathbf{r}_{\beta,j}) \right\rangle \tag{2.48}$$

where the complex term vanishes on average. Generalisation from the pure system to the mixtures seems quite natural. However, the kinetic relaxation time of the partial density modes $\hat{n}_{\alpha\mathbf{q}}$ is very different from their pure system analogues. On the one hand, for good mixing compounds, the thermodynamic unbalance giving rise to the equilibrium restoring forces is comparably small. The better lipid components mix, the less is the incentive for these molecules to demix. Second, the nature of the relaxation kinetics is different, and is expected to be controlled by a coefficient D whose value should be similar to the molecular self-diffusion coefficient D_s. The latter is certainly smaller than the collective coefficient D_c which is usually enhanced by the short range repulsive interaction between the lipids. Good mixing and slow interdiffusion combines to make the convergence of the crossed correlation $\langle \hat{n}_{\alpha,\mathbf{q}} \hat{n}_{\beta,-\mathbf{q}} \rangle$ significantly slower than for the pure case.

As a benchmark, we introduce the concept of "false mixtures" where a pure bilayer is treated as a binary mixture by simply relabelling randomly all the molecules and splitting them into two artificial classes. The mixture is by construction ideal and the thermodynamic restoring forces only originate from the Gibbs entropy of mixing of the two components.

Linear response theory provides in principle a shortcut for the determination of the mixing parameter B. The idea for a binary mixture of composition x_1, x_2 is to add two external potentials ϕ_1, ϕ_2 acting respectively on species 1 and 2.

$$\begin{aligned} \phi_2(\mathbf{r}) &= -U_0 x_1 \cos(\mathbf{q}\cdot\mathbf{r}) \\ \phi_1(\mathbf{r}) &= U_0 x_2 \cos(\mathbf{q}\cdot\mathbf{r}) \end{aligned} \tag{2.49}$$

This causes a linear density modulation response $\rho_{0,1} + \rho_{1,1}\cos(\mathbf{q}\cdot\mathbf{r})$, $\rho_{0,2} + \rho_{1,2}\cos(\mathbf{q}\cdot\mathbf{r})$. Then the combination

$$X = \frac{\rho_{1,1}}{\rho_{0,1}} - \frac{\rho_{1,2}}{\rho_{0,2}} = 1 + \rho x_1 x_2 (\tilde{h}_{11}(\mathbf{q}) + \tilde{h}_{22}(\mathbf{q}) - 2\tilde{h}_{12}(\mathbf{q})) \tag{2.50}$$

leads to the direct determination of the mixing parameter $B = (X-1)/(2Xx_1x_2)$. Similar relations enable to calculate the η and ζ parameters (see Appendix A).

2.9 Area compressibility

Fluid bulk compressibility is a measure of change in the volume resulting from the external pressure which is applied on the surface of the object. The isothermal compressibility can be expressed as follows [7]:

$$\chi_T = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T \tag{2.51}$$

where V is the volume (in 3d) and P is the external pressure.Alternatively one can introduce an isothermal bulk modulus (or a stretching modulus for the 2d case) K_A, defined as $K_A = 1/\chi_T$.

In the scope of the current project, we are interested in a lipid membrane which at first approximation can be treated as a 2d film. Let us assume that a membrane plane is perpendicular to the z axis. A fluid bilayer can withstand an anisotropic stress, *i.e.* a stress that have different Σ_{zz} and Σ_{xx}, Σ_{yy} components. The two latter components, however, must be equal and no stress in xy direction is allowed. The surrounding fluid, being a usual fluid, cannot contribute to the xz or yz stress anisotropy. The typical situation is therefore to assume that the anisotropic part of the stress arises directly from an internal membrane tension σ.

In practice, in simulations it means we can set an isotropic pressure P_x in the x, y direction and another constant pressure P_z in the z direction. Computing the resulting forces leads to a z component $P_z L_x L_y$ transmitted by the solvent on the top of the simulation box and a lateral x component $P_x L_z L_y$ along the membrane. The membrane tension accounts for the difference $\sigma = L_z \cdot (P_z - P_x)$ between the actual x force component, and the force that a usual fluid, which is subject to a pressure P_z, would be able to withstand. The state of a fluid membrane then requires 2 mechanical thermodynamic parameters: the isotropic fluid pressure P_z and the membrane tension σ [22]. Finally, it can be concluded that a pure fluid lipid membrane patch of symmetric composition should be described by the four independent thermodynamic parameters: $N_{\text{lipid}} = N_b$ number of lipids in a bilayer ($N_b = 2N_l$ number of lipids in a leaflet), T temperature, P_z isotropic fluid pressure and σ membrane tension. Those parameters are typically directly approachable in the Molecular Dynamics simulations. It is also assumed that enough water is available during the simulation, which correspond to a so-called "full hydration" situation. This is required so that the state of the membrane does not depend on the number of the solvent molecules.

As a result, the final system behaves like a membrane in contact with a reservoir of water molecules whose chemical potential μ is constant and restricted by P_z and T. In addition, the isotropic pressure is often set equal to the ambient pressure ($\sim$ 1 atm) and omitted from the discussion. The area of the box is then a fluctuating quantity depending on N_b, T and σ.

One can define a 2d compressibility χ_T as follows:

$$\chi_T = \frac{1}{A}\left(\frac{\partial A}{\partial \sigma}\right)_T \tag{2.52}$$

where $A = L_x L_y$ is the membrane area.

Lipid membranes have a natural reference state associated to $\sigma = 0$ and referred as "tensionless state". The associated equilibrium area per lipid (APL) a_0 is a central quantity, that has been measured experimentally [40], and is also one of the main outcome of numerical simulations. APL can be defined as follows:

$$a_0 = A_0/N_l \tag{2.53}$$

where A_0 is the equilibrium membrane area. Another important equilibrium characteristic of membranes is the bilayer thickness. It can be defined differently, based on the method of calculation and overall definition of the bilayer. One can strictly assign the membrane border to a lipid particle, head group, water-lipid surface, etc. In this work we use the simple definition:

$$z = \langle z_{\mathrm{PH}}^{\mathrm{up}} - z_{\mathrm{PH}}^{\mathrm{low}} \rangle \tag{2.54}$$

with $\langle . \rangle$ here being the average over time and lipids, $z_{\mathrm{PH}}^{\mathrm{up}}$ and $z_{\mathrm{PH}}^{\mathrm{low}}$ - the z coordinates of the phosphorus beads of the upper and lower leaflets correspondingly.

To calculate the compressibility of the membrane one can apply a set of different pressures and let the system equilibrate to a new configuration. Then, the equilibrium total lipid bilayer area A is expected to have a linear behavior with the tension around its reference ($\sigma = 0$) value $A_0(N_b, T)$, so that

$$\chi_T = \frac{1}{A}\frac{A - A_0}{\sigma} \simeq \frac{1}{A_0}\frac{A - A_0}{\sigma} \tag{2.55}$$

This approach can be naturally implemented in simulations by plotting the area *v.s.* the tension in a set of equilibrated simulations with anisotropic pressure conditions.

Another possibility to obtain the compressibility of the system from simulations is to use the area (in the 2d case) fluctuations of the box size of our system [7]:

$$\chi_T = \frac{\mathrm{var}^2(A)}{k_B T \langle A \rangle} \tag{2.56}$$

where $\langle A \rangle = A_0$ is the equilibrium area and $\mathrm{var}^2(A)$ is the variance of the area in a constant tension σ-temperature T ensemble.

Finally, the compressibility can also be given by the $q = 0$ limit of the structure factor of the system [91]:

$$S(\mathbf{q} \to 0) = \rho k_B T \chi_T \Rightarrow \chi_T = \frac{S(\mathbf{q} \to 0)}{k_B T \rho} \tag{2.57}$$

where $\rho = N_b / A_0$ is the number density of the system.

Generally speaking, such approaches can be applied to finite vesicle systems, as it was shown by Diamant [12] that the thermodynamics of a 3d membrane vesicle can be reduced to a generic 2d finite system enclosing the corresponding 3d volume. The above choice of (N, σ, P_z, T) is one possibility. The conjugated choice (N, A, V, T) is another possibility, with A the projected box area $L_x L_y$ and V the total volume. The first set of control variable must be used when one wishes to find the average area $\langle A \rangle$ and volume $\langle V \rangle$ (and consequently box size $\langle L_x \rangle, \langle L_y \rangle, \langle L_z \rangle$) associated to a given set of conditions $\sigma = 0$, $P_z = 1$ atm, T. Once these are known, we can simulate without barostat with constant box size conditions, which corresponds to the second set of conditions to obtain constant volume simulations of equilibrated system.

2.10 Features of the 2d structure factor of a molecular fluid

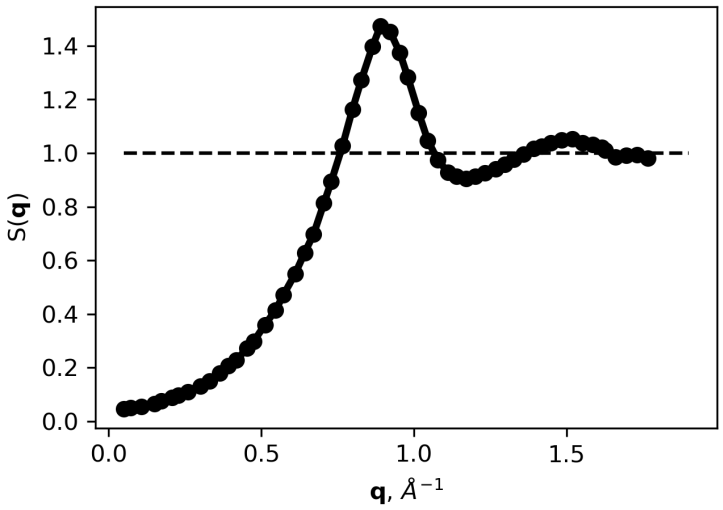


FIGURE 2.2: An example of the structure factor for fluid system.

Let us discuss in more detail a few special features of the structure factor $S(\mathbf{q}) = 1 + \rho\tilde{h}(\mathbf{q})$ in the case of a pure bilayer system treated as a two dimensional surface. We assume that each molecule is represented by a 2d point with coordinates (x_j, y_j), $j = 1 \dots N$, N is the total number of molecules. The central object is the equilibrium density modes $\sum_{j=1}^{N} \exp(i\mathbf{q} \cdot \mathbf{r}_j)$ correlation functions. The resulting structure factor cannot be compared directly with experiments, as in real samples all the atoms in the molecule participate in the scattering, while in simulation we are free to choose which particles to include into the consideration.

During the analysis of the obtained structure factors, it is important to distinguish between two following regimes. The first one is an intermediate $q \sim 1$ nm^{-1}. In this wave length one probes the local environment of the lipids, and so it is fair to expect a display of a correlation peak $S_{\text{peak}}(\mathbf{q})$. The height and shape of the first peak are seen to depend on the choice of the representative point of the molecule (center of mass, representative bead, etc.). the second regime is a low q regime, which probes the collective density fluctuations of the fluids. The structure factor is expected to be analytical in $\mathbf{q}$ in the absence of the critical fluctuations, and isotropic with a small q expansion $S(\mathbf{q}) = S(0) + C\mathbf{q}^2 + \dots$ As was mentioned previously, $S(0)$ is related to the fluid compressibility and should not depend on the choice of the representative point. The prefactor C however depends in general on this choice.

The dependence of $S_{\text{peak}}(\mathbf{q})$ on the representation choice can be rationalized by assuming a factorization $S_{\text{peak}}(\mathbf{q}) - 1 \simeq (S_{\text{best}}(\mathbf{q}) - 1)|F(\mathbf{q})|^2$ into a contribution $S_{\text{best}}(\mathbf{q})$ that best localizes the "center" of the lipid, and a "Debye-Waller" or form factor $F(\mathbf{q})$ which encodes the peculiarity of a each choice for representing a lipid molecule. This decomposition is appropriate for dilute fluids or in special situations, such as when the position and the orientation of molecules are statistically uncorrelated. Let us remind the reader, that the above arguments are only a working hypothesis because it is not possible to say a priori which method is the best to represent the lipid molecules. Now, assuming such "best" but unknown localization $\mathbf{r}_j^{(\text{best})}$ of lipid molecule j exists, and that the difference $\mathbf{r}_j - \mathbf{r}_j^{(\text{best})}$ with $\mathbf{r}_j$ being the actual position of the lipid, is described by a statistical distribution identical for all lipids and independent on their environment, one finds

$$\hat{n}_{\mathbf{q}} = \sum_{j=1}^{N} \exp(i\mathbf{q} \cdot \mathbf{r}_j) = \sum_j \exp\left(i\mathbf{q} \cdot \mathbf{r}_j^{(\text{best})}\right) \exp\left(i\mathbf{q} \cdot (\mathbf{r}_j - \mathbf{r}_j^{(\text{best})})\right) \tag{2.58}$$

$$\hat{n}_{\mathbf{q}}\hat{n}_{-\mathbf{q}} = N + \sum_j \sum_{k \neq j} \exp\left(i\mathbf{q} \cdot (\mathbf{r}_j^{(\text{best})} - \mathbf{r}_k^{(\text{best})})\right) \exp\left(i\mathbf{q} \cdot (\mathbf{r}_j - \mathbf{r}_j^{(\text{best})} - \mathbf{r}_k + \mathbf{r}_k^{(\text{best})})\right) \tag{2.59}$$

Averaging $\exp\Big(i\mathbf{q}\cdot(\mathbf{r}_j-\mathbf{r}_j^{(\text{best})}-\mathbf{r}_k+\mathbf{r}_k^{(\text{best})})\Big)$ as

$$\begin{aligned}\Big\langle\exp\Big(i\mathbf{q}\cdot(\mathbf{r}_j-\mathbf{r}_j^{(\text{best})}-\mathbf{r}_k+\mathbf{r}_k^{(\text{best})})\Big)\Big\rangle &= \Big\langle\exp\Big(i\mathbf{q}\cdot(\mathbf{r}_j-\mathbf{r}_j^{(\text{best})})\Big)\Big\rangle\Big\langle\exp\Big(i\mathbf{q}\cdot(\mathbf{r}_k-\mathbf{r}_k^{(\text{best})})\Big)\Big\rangle^* \\ &= F(\mathbf{q})F(\mathbf{q})^* \end{aligned} \tag{2.60}$$

$$\begin{aligned}&\Big\langle\sum_j\sum_{k\neq j}\exp\Big(i\mathbf{q}\cdot(\mathbf{r}_j^{(\text{best})}-\mathbf{r}_k^{(\text{best})})\Big)\exp\Big(i\mathbf{q}\cdot(\mathbf{r}_j-\mathbf{r}_j^{(\text{best})}-\mathbf{r}_k+\mathbf{r}_k^{(\text{best})})\Big)\Big\rangle \\ &\qquad= N\Big\langle\sum_{k\neq j}\exp\Big(i\mathbf{q}\cdot(\mathbf{r}_j^{(\text{best})}-\mathbf{r}_k^{(\text{best})})\Big)\Big\rangle|F(\mathbf{q})|^2 \\ &\qquad= N(S_{\text{best}}(\mathbf{q})-1)|F(\mathbf{q})|^2\end{aligned} \tag{2.61}$$

we obtain $S_{\text{peak}}(\mathbf{q}) = 1+|F(\mathbf{q})|^2(S_{\text{best}}(\mathbf{q})-1)$, with $S_{\text{best}}(\mathbf{q})$ being the structure factor which would be obtained if the best localization $\mathbf{r}_j^{(\text{best})}$ was considered. Here, * is marking the conjugate value. The form factor $|F(\mathbf{q})|^2$ is usually a positive decreasing function of $\mathbf{q}$, such that $|F(0)|^2 = 1$. Let us assume that fluctuations of $\mathbf{r}_j$ around $\mathbf{r}_j^{(\text{best})}$ are small. In other words, one can imagine the two positions $\mathbf{r}_j$ and $\mathbf{r}_j^{(\text{best})}$ to be connected with a spring and, thus, $\mathbf{r}_j$ is experiencing Gaussian fluctuations around $\mathbf{r}_j^{(\text{best})}$. In such situation the form factor should be bell-shaped with a width of the order of the inverse of the lateral size of a lipid. Therefore, if such a factorization holds in our case, the shape of $S_{\text{peak}}(\mathbf{q})-1$ would be given by the product of a sharply peaked reference structure factor $S_{\text{best}}(\mathbf{q})-1$ multiplied by a decreasing positive prefactor that is itself decreasing on the same scale. This is a likely an explanation of the dependence of $S_{\text{peak}}(\mathbf{q})$ on the choice of the reference lipid positions, such as described in Chapter 4.

On the other hand, the behavior of $S(\mathbf{q})$ in the low $\mathbf{q}$ region seems somewhat unrelated to what rules the neighborhood of the first peak. The fact that $|F(0)|^2 = 1$ underlines that $S(q=0)$ should become independent of the choice of the reference lipid representation. This in turns means that in the very coarse-grained hydrodynamic limit, lipid molecules loose their identity. This does not mean however that the next term in the analytical $\mathbf{q}$ expansion near 0 is independent of this choice, as confirmed in simulations (see Chapter 4).

An important difference between the first peak region and the low $\mathbf{q}$ region originates from the way the structure factor is built. In the first peak region, $S_{\text{peak}}(\mathbf{q})$ is mostly dependent on the relative positions of a few variables $\mathbf{r}_j$ associated to a lipid j and its close environment. By contrast, the structure factor $S(\mathbf{q})$ in the low $\mathbf{q}$ region is a very collective object, that combines hundreds of positions $\mathbf{r}_j$ in our typical simulations. At these scales, one can introduce a collective density field $\rho(\mathbf{r},t)$ corresponding to a coarse-grained average of the density field observable $n(\mathbf{r}) = \sum_j \delta(\mathbf{r}-\mathbf{r}_j(t))$. This field is non homogeneous and fluctuates slowly. Such kind of an approach is the description commonly referred as

"fluctuating hydrodynamics". The density modes now correspond to

$$\hat{n}_{\mathbf{q}}(t) = \sum_{j=1}^{N} \exp\big(i\mathbf{q}\cdot\mathbf{r}_j(t)\big) \simeq \int_{\mathcal{S}} \mathrm{d}\mathbf{r}\, \rho(\mathbf{r},t) e^{i\mathbf{q}\cdot\mathbf{r}} \tag{2.62}$$

with unit cell $\mathcal{S}$. The structure factor in the low $\mathbf{q}$ region is therefore associated to the statistical equilibrium correlations of the coarse-grained hydrodynamic density field $\rho(\mathbf{r},t)$.

A possible reason for the $\mathbf{q}^2$ dependence of $S(\mathbf{q})$ near $\mathbf{q}=0$ might be explained at phenomenological level by the presence of an effective 2-bodies pair interaction between lipids. For instance, if one assumes that the free-energy of a weakly modulated density profile $\rho(\mathbf{r}) = \rho_0 + \rho_1 \cos(q_x x)$ is given approximately by a sum over all pairs of molecules of an effective pair-interaction $v_{\mathrm{int}}(\mathbf{r})$, one finds a contribution to the free-energy in the following form:

$$\int_{\mathcal{S}} \mathrm{d}\mathbf{r}_1 \int_{\mathcal{S}} \mathrm{d}\mathbf{r}_2\, \rho(\mathbf{r}_1)\rho(\mathbf{r}_2) v_{\mathrm{int}}(\mathbf{r}_1 - \mathbf{r}_2) \tag{2.63}$$

which by integrating over a repeated unit cell $\mathcal{S}$ with periodic boundary conditions as in a simulation box leads to an expansion in powers of the amplitude ρ_1.

$$N\rho_0 \left[\tilde{v}_{\mathrm{int}}(0) + + \frac{\rho_1^2}{2\rho_0^2} \tilde{v}_{\mathrm{int}}(q_x) \right] \tag{2.64}$$

As the free-energy now explicitly depends on $\mathbf{q}$, the amplitude of the fluctuating density modes $\langle \rho_1^2 \rangle$ is likely to be also $\mathbf{q}$ dependent, and so does $S(\mathbf{q})$. If two different choices of lipids representation leads to two different effective interactions potentials $\tilde{v}_{\mathrm{int}}(q_x)$, then the $C\mathbf{q}^2$ terms naturally differ. In addition if $v_{\mathrm{int}}(\mathbf{r})$ has a natural length scale of the order of the horizontal separation between lipids, then $\tilde{v}_{\mathrm{int}}(q_x)$ can account for the smooth curvature of $S(\mathbf{q})$ from $\mathbf{q}=0$ to the first correlation peak.

2.11 Fluctuating bilayers

So far we have assumed that the membrane can be treated as a planar surface. However, it is a well-known fact that biological membranes fluctuate not only in planar but also in vertical direction. Thus, and it is important to examine how much the latter changes the results exposed in the previous sections.

As was discussed before, the simplest possible description of a fluctuating bilayer consists in assimilating it to a fluctuating surface. One can imagine this as an extreme coarse-graining. This fluctuating surface is assumed to be infinitely thin and continuous. Membrane fluctuations are then controlled by a bending modulus κ, usually of the order of 10-20 $k_B T$ in the fluid phase. Such large value means that small bilayer portions are nearly

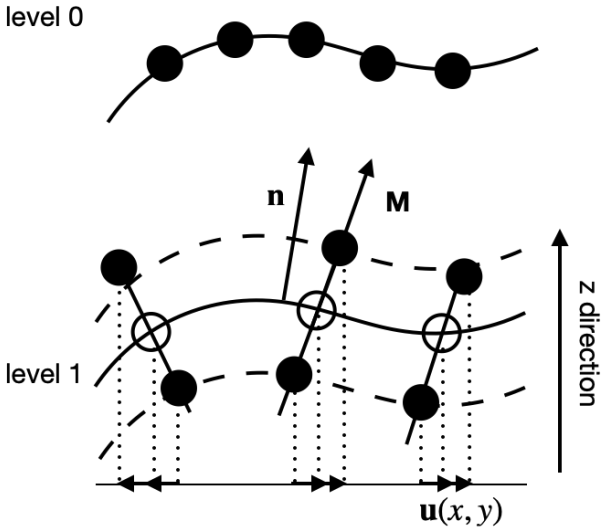


FIGURE 2.3: Bilayer leaflet representation as an infinitely thin surface (level 0) and as a surface with a director **M** and a thickness (level 1). Normal vector to the surface **n** and director **M** are shown. Representative beads in each case are shown in black, and ideal infinitely thin reference points are shown in white in the case of the level 1 description. The displacement field **u** corresponds to the horizontal projection of the vector linking the ideal position (white) to the real positions (black) of the lipids. $\mathcal{D}$ (see text) is the typical distance between white and black symbols. The white symbols are associated to the neutral surface (thick black line).

flat. However, even weak undulations may affect substantially the nature of the effective projected density modes correlations.

As a first step a refinement of the fluctuating surface approximation, it is possible to distinguish between the two leaflets and assign to them a vector order parameter $\mathbf{M}_\ell(\mathbf{r})$ ($\ell = 1, 2$ standing for the upper and lower leaflet correspondingly). The vector order parameter is a quantity, coarse-grained on the scale of one or a few lipid molecules, that expresses the average spatial orientation of a lipid from the tail to the head group, as discussed for instance in [85, 86]. In the following text, this unit vector **M** will be referred as the "director" (see Fig. 2.3). A lipid is said to be tilted when the director does not coincide with the local membrane normal orientation (which is another "director" vector, but coarse-grained at a much larger lateral scale). In the current project the latter in the ideal case would coincide with the z axis. By ideal case here we mean completely flat membrane.

Assuming that the bilayer leaflets are only weakly coupled mechanically, it is natural to assume that in each leaflet a so-called "neutral surface" should exist. By neutral surface we mean a surface with respect to which the lipid directors can pivot and tilt ($\mathbf{M} \neq \mathbf{n}$). Tilting only a single lipid molecule with respect to its neighbor is not energetically favorable,

however, tilting a bundle of neighboring lipids is more favorable. The latter can lead to the cancel out of the effects in a long-range perspective, meaning that expanding in one leaflet leads to contracting in the opposite one (see Fig. 2.4). One can associate an elastic free-energy to the tilt degrees of freedom, which is inspired from the description of the nematic liquid crystals [22].

The average tilt of a lipid is non vanishing in a number of low temperature states of lipid membranes, known as $L_{\beta'}$ or gel phases. In the high temperature fluid phase that we focus on most of the time, the average tilt is zero and one can consider that there is a thermodynamic restoring torque acting on the tilted lipids. To express mathematically the nature of such restoring potential energy in the case of a flat leaflet ($\mathbf{n} \parallel z$), one decomposes $\mathbf{M}$ as $\mathbf{M}_{\parallel} + \mathbf{M}_{\perp}$, $\parallel$ being the x, y component and $\perp$ the z component. The tilt elasticity density contains a quadratic term $K_0 \mathbf{M}_{\parallel}^2/2$ and one or more quadratic gradient terms $K_1 \nabla \mathbf{M}_{\parallel} : \nabla \mathbf{M}_{\parallel}$. The energy equipartition theorem predicts that the Fourier modes $\tilde{\mathbf{M}}_{\parallel,\mathbf{q}}$ associated to $\mathbf{M}_{\parallel}$ fluctuate with amplitude $k_B T L_x L_y/(K_0 + K_1 \mathbf{q}^2)$ (the three nematic elastic constants are here taken all equal to simplify). K_0 is responsible for efficiently restoring the lipid director field towards its vertical orientation. K_1 is the term preventing two neighboring lipids to orient in different directions.

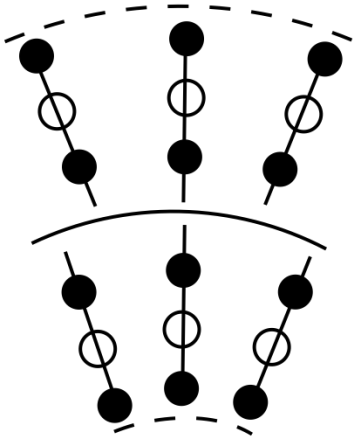


FIGURE 2.4: Illustration of the mean curvature effect of the bilayer: expansion of the black symbols in the upper leaflet and contraction in the lower leaflet. Representative beads are shown in black, and ideal infinitely thin reference points are shown in white.

Let us assume that the so-called neutral surface does exist for each leaflet and lipids oscillate around it inside a membrane. It is natural to assume that the best reference point of the molecule $\mathbf{r}^{(\text{best})}$ should be located on the intersect of the neutral surface and the lipid. In this way, the lipid tilt does not change $\mathbf{r}^{(\text{best})}$, and the associated $S_{\text{best}}(\mathbf{q})$ structure factor remains immune to any tilt fluctuations. However, if the representative point of a lipid lies at a distance $\mathcal{D}$ above or below the neutral surface, a non vanishing tilt induces a horizontal displacement $\mathbf{u}(\mathbf{r}_j) = \mathbf{r}_j - \mathbf{r}_j^{(\text{best})} \simeq \mathcal{D}\mathbf{M}_{\parallel}(\mathbf{r})$ on each particle j in the system. The coarse-grained local displacement $\mathbf{u}(\mathbf{r}_j)$ is then responsible for a relative excess in the density $\delta\rho(\mathbf{r}) = -\rho_0 \text{div}(\mathbf{u}(\mathbf{r})) = -\rho_0 \mathcal{D} \text{div}(\mathbf{M}_{\parallel})$. The coarse-grained displacement field $\mathbf{u}(\mathbf{r})$ is associated to the fluctuating coarse-grained density $\rho(\mathbf{r})$. The displacements $\mathbf{r}_j - \mathbf{r}_j^{(\text{best})}$ are supposed to be smooth and given by the coarse-grained field $\mathbf{u}(\mathbf{r}_j)$.

Thus, the tilt degrees of freedom are about to create new density fluctuation modes $\rho_0 \mathcal{D}\, i\mathbf{q}\tilde{\mathbf{M}}_{\parallel,\mathbf{q}}$, uncoupled at first order to the spontaneous density fluctuations responsible for the building of $S_{\text{best}}(\mathbf{q})$. The tilt modes come as an extra term in the structure

factor and contribute to the product $NS(\mathbf{q})$ as follows (see details in Appendix A):

$$\begin{aligned} NS(\mathbf{q}) &= NS_{\text{best}}(\mathbf{q}) + \rho_0^2 \mathcal{D}^2 \mathbf{q}^2 \langle \tilde{M}_{\parallel}^2 \rangle \\ S(\mathbf{q}) &= S_{\text{best}}(\mathbf{q}) + \rho_0 \mathcal{D}^2 k_B T \frac{\mathbf{q}^2}{K_0 + K_1 \mathbf{q}^2} \end{aligned} \tag{2.65}$$

Although, it can be seen now that the tilt modes around a flat leaflet cannot change the value of the structure factor near the origin $\mathbf{q} = 0$, they modify the quadratic deviation coefficient C from the expansion of $S(\mathbf{q})$ with a positive contribution, varying quadratically with the distance $\mathcal{D}$ to the neutral surface.

We now assume that the surface describing the leaflet as a first approximation is not flat any more, but described by an elevation function $z(\mathbf{r})$ (Monge representation). The normal vector to the surface has components $\mathbf{n} = (-z_{,x}, -z_{,y}, 1)/\sqrt{1 + (\nabla z)^2}$, $z_{,x}, z_{,y}$ being the spatial derivatives of $z(\mathbf{r})$.

The next level of approximation associates the surface $z(\mathbf{r})$ to the neutral surface of the membrane. The main difficulty when dealing with this general situation is that there is no obvious way one can put a non flat membrane in relation with a flat plane. A rough arbitrary surface has a non vanishing total (Gauss) curvature that prevents from finding a mapping to a flat surface while preserving distance. Rough and flat surfaces are generally non isometric.

In the most general situation, one can conceive that membrane geometry and membrane local density fluctuations are coupled. This is the case for instance if the lipid components have a "conical" or "inverse conical shape", a property known as "intrinsic lipid curvature" that influences the curvature of the leaflets. However, we restrict ourselves to a situation where membrane is only marginally curved, and neglect all curvature-composition couplings. In addition we do not attempt to determine the neutral surface shape and only deal with the (x, y) projected components of the lipid representative coordinates. The question that needs to be solved is then to relate the apparent projected density modes to the real local density changes that could be observed in the local frame tangent to the neutral surface.

A first apparent effect of the above arguments arises from the geometrical projection, and the expression of the local surface element in terms of the projected surface. It is well known that the local surface above a small integration element $\mathrm{d}x\mathrm{d}y$ is $\sqrt{1 + (\nabla z)^2}\,\mathrm{d}x\mathrm{d}y$ (see *e.g.* [65]). This means that a homogeneous density on the fluctuating surface creates an apparent modulation $\delta\rho$ of the projected density equal to

$$\delta\rho(\mathbf{r}) = \rho_0 \frac{(\nabla z)^2}{2} \tag{2.66}$$

This term depends on the ratio $(k_BT/\kappa)^2$ and its variance has a quartic dependence $|\nabla z|^4$ in the slope of the bilayer. We therefore consider that this term does not influence significantly the projected structure factor because due to the simulation conditions (PBC enforce the horizontality of the membrane) this effect should be comparably small in our case.

Similar to the tilt example, deviation from the neutral surface $z(\mathbf{r})$ should also be responsible for a horizontal relative displacement $\mathbf{u}(\mathbf{r}) = \mathcal{D}\mathbf{n}_{\|}$ in the x, y plane, $\mathbf{n}$ being the normal vector to the neutral surface. Such displacement leads in turn to a local excess density $-\rho_0 \mathrm{div}(\mathbf{u})$. Neglecting the normalization term $\sqrt{1+(\nabla z)^2} \simeq 1$ leads to the following expression:

$$\mathrm{div}(\mathcal{D}\mathbf{n}_{\|}) = -\mathcal{D}\Delta z \tag{2.67}$$

involving the Laplacian of the surface (local mean curvature) which Fourier transforms as $-q^2\tilde{z}_{\mathbf{q}}$. Similar to the above derivation (see Appendix A) one writes

$$NS(\mathbf{q}) = NS_{\mathrm{best}}(\mathbf{q}) + \rho_0^2\mathcal{D}^2\mathbf{q}^4\langle|\tilde{z}_{\mathbf{q}}|^2\rangle \tag{2.68}$$

The average $\langle|\tilde{z}_{\mathbf{q}}|^2\rangle$ is known as the Helfrich spectrum and in the absence of the surface tension corresponds to a $\frac{k_BTL_xL_y}{\kappa\mathbf{q}^4}$ term. The two factors $\mathbf{q}^4$ compensate so that $\mathcal{D}^2\mathbf{q}^4\langle|\tilde{z}_{\mathbf{q}}|^2\rangle$ does not vanish in the $\mathbf{q} \to 0$ limit. Instead, it gives rise to a quadratic constant term proportional to the deviation $\mathcal{D}$ of the representative point $\mathbf{r}$ of the lipid molecule relative to the neutral surface. Finally, one gets

$$S(\mathbf{q}) = S_{\mathrm{best}}(\mathbf{q}) + \rho_0\mathcal{D}^2\frac{k_BT}{\kappa} \tag{2.69}$$

It can be concluded, that the membrane undulations are likely to cause the strongest corrections to the low $\mathbf{q}$ structure factor. They seem to be the factor able to change the value of $S(0)$ while other contributions (effective pair-wise interactions, local tilt) only act by changing the $C\mathbf{q}^2$ expansion correction. The effect of the membrane undulations must therefore be minimized by choosing the representative positions for lipid molecules as close as possible to the neutral surface.

An interesting and important corollary of the undulation correction mechanism is that it should act simultaneously on both leaflets in an opposite manner, in other words the undulation corrections of both leaflets are anti-correlated. Taking both leaflets together to define a single global density term (as in eq. 2.33) should neutralize and minimize the influence of such undulation corrections. It also provides an explanation for possible numerical discrepancies between the single leaflet and the full bilayer approach to the determination of the structure factor in pure system and in consequence, the compressibility

coefficient.

In the case of the multicomponent lipid mixtures, we also observe that undulation corrections act in a correlated manner on all lipid types in a given leaflet, provided the lipids of different species are sufficiently similar in shape and size. For instance, if a binary mixture $\alpha = 1, 2$ of 2 phospholipid types is considered, and if the representative point of these lipid is defined in a similar way (*e.g.* the center of mass of each molecule), then undulation corrections should mostly cancel out in the cross-terms involved in the calculations of some thermodynamic properties. For example, in combinations such as the one appearing in the $\Delta = \tilde{h}_{11} + \tilde{h}_{22} - 2\tilde{h}_{12}$ term (eq. 2.25) introduced by Ben-Naim in his Kirkwood-Buff treatment of the regular solutions and used for the calculation of the mixing parameter in this work. Therefore, we expect that the undulation corrections, even though indeed present in the system, should affect less the value of the non ideality parameter B than the full compressibility of the membrane χ_T.

Watson et al. in their work [85] demonstrated the distinction between the local membrane inclination due to undulations and the tilt with respect to the normal membrane direction is artificial. They propose to deal directly with the fluctuations of the local director vector field $\mathbf{M}(\mathbf{r})$, and show that the lipid tilt elasticity coefficients and the membrane bending modulus are indeed strongly related. In simulations with periodic boundary conditions, the global fluctuations of $\mathbf{M}$ are constrained. In particular the integrated change in density $\int_S \mathrm{d}\mathbf{r}\, \mathrm{div}(\mathbf{M}_\parallel)$ vanishes due to the fact that the total number of lipids in the system is constant. However, finite wavelength modes $\tilde{\mathbf{M}}_{\parallel,\mathbf{q}}$ are not bound to such a constraint and are expected to fluctuate freely. Their fluctuations, computed by Watson et al., are inversely proportional to the membrane bending modulus. It can be concluded, that their approach should be consistent with our argument which is based on the Helfrich undulation spectrum. Fluctuations of $\tilde{\mathbf{M}}_{\parallel,\mathbf{q}}$ is similar to the density fluctuations in a flat fluid in the canonical ensemble, which are constrained in the $q = 0$ case, but free otherwise. In other words, $\tilde{\mathbf{M}}_{\parallel,\mathbf{q}}$ is expected to fluctuate as if there was no constraint imposed by the periodic boundary conditions. Thus, the apparent structure factor of a pure system should obey a generalized relation:

$$S(0) = \rho_0 \frac{k_B T}{K_A} + \rho_0 \mathcal{D}^2 \frac{k_B T}{\kappa} \tag{2.70}$$

where the first term accounts for the in plane density fluctuations of the fluid controlled by a membrane compressibility coefficient $K_A = \chi_T^{-1}$ and the director fluctuations (out of plane undulation corrections) controlled by the membrane bending modulus. A coupling term with the dimension of the square of the length $\mathcal{D}^2$ expresses how much tilt and undulations act on the apparent projected density modes. It can be concluded that the true

value of $S(0)$ are expected to be available at the minimal $\mathcal{D}$ distance from the monolayer neutral plane.

CHAPTER 3

Methodology and simulation details

To study physical, chemical, or other phenomena one relies a lot upon computer simulations. In the field of molecular physics and chemistry, it is possible to adopt two main strategies: Monte Carlo (MC) and Molecular Dynamics (MD) simulation. MC is a probabilistic method that enable fast and efficient sampling of complex configuration spaces [48]. A variety of fields utilize MC simulations, including finance, engineering, supply chain, and science. **Molecular Dynamics** is based on solving Newton's equations of motion, giving a time evolution of the desired system [1, 2]. It is well-suitable for scientific research of different physical problems, such as phase transition, diffusion, phase separation, etc. It can be applied to both biological and purely physical systems, i.e. DNA, proteins, glass-forming systems, etc. A particularly important application of MD simulation is to determine how a biomolecular system responds to some external perturbation [16]. While MC enables in principle a faster sampling of the configurations and requires only the configurational energy, MD on the other hand requires the forces, and describes more accurately the dynamical properties of the system. Another benefit of using MD algorithms is that it can be efficiently parallelized through spacial decomposition.

There are many different softwares available and capable of doing great efficient computer experiments: GROMACS, DL POLY, CHARMM, AMBER, NAMD ...to name a few. In this work the **LAMMPS** (Large-scale Atomic/Molecular Massively Parallel Simulator) package was used. It is an open-source code that is intensively developed and updated to fit even more scientific purposes. It allows flexible possibilities of external

force application and choice of the thermostat/barostat options. LAMMPS has a strong and active user community which allows discuss implementations of new methods and physical accuracy of the software.

Moreover, this software can be run in parallel to provide accelerated performance which is why it was possible to use supercomputer facilities during this work. LAMMPS gives users an opportunity to work with different units, and for most bio inspired simulations the `real` units are appropriate. The most important of these units are given in Tab. 3.1.

TABLE 3.1: LAMMPS real units.

Value	Unit
mass	grams/mole
distance	Angstrom
time	femtoseconds
energy	Kcal/mole
temperature	Kelvin
pressure	atmospheres
force	kcal/(mol·Å)

For visualization and creation of snapshots along this thesis Visual Molecular Dynamics (VMD) software was used. Movies were generated using the Ovito package as it provided faster rendering engine compared to VMD, while somewhat lower resolution.

3.1 Classical Molecular Dynamics

In this work, the MD approach was used to study the evolution of a lipid membrane in water solution. Let us discuss this method in more detail.

Let us consider a system of N particles with a fixed volume V and total energy E in the microcanonical ensemble. The evolution of the microscopic configuration of the system can be determined based on the positions and velocities of all particles of the system at a certain time t. MD implies numerically solving classical Newtons equations of motion:

$$\mathbf{F}_i(t) = m_i \frac{d^2\mathbf{r}_i(t)}{dt^2} = m_i \mathbf{a}_i(t) \tag{3.1}$$

with $\mathbf{F}_i(t)$ being the force applied to particle i, m_i, $\mathbf{r}_i(t)$ and $\mathbf{a}_i(t)$ - mass, position and acceleration of ith particle, $i = 1...N$ - index of the particle, and t - time. The total force that acts on particle i can be defined as follows:

$$\mathbf{F}_i(t) = -\frac{\partial u_N(\mathbf{r}_1(t), ..., \mathbf{r}_N(t))}{\partial \mathbf{r}_i(t)} \tag{3.2}$$

where $u_N(\mathbf{r}_1(t), ..., \mathbf{r}_N(t)) = u(\mathbf{r}^N)$ is the potential energy of the system. It is important to mention here that in real MD calculations usually only pair-wise interactions are computed.

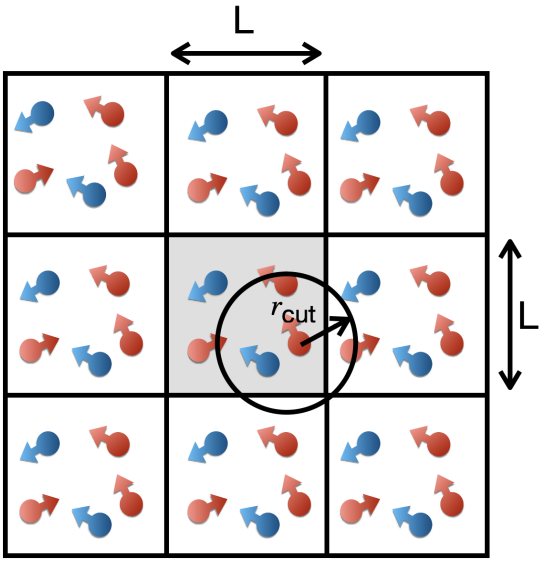


FIGURE 3.1: Schematic illustration of the periodic boundary conditions. The unit cell is marked gray and has size L, it consists of red and blue particles, r_{cut} is the maximum interaction distance. Images of the simulation cell are shown around the initial unit cell.

To minimize boundary effects during simulations it is common to use Periodic boundary conditions (PBC). In computer simulations PBC replicate the system in order to simulate an infinite space. A representation of the PBC is shown in Fig. 3.1. One can see that it means creating images of all particles confined in the box and periodically replicating them in all spatial directions. To distinguish which particles are placed within the interaction radius r_{cut} and to calculate the actual distance between the given particles, the minimum image convention is used [48]. Its idea lies in finding the minimum distance between two particles or the images of the neighboring replicas. Then a simple rule is applied to calculate the final distance between particles i and j:

$$r_{ij} = \begin{cases} r_{ij}, & \text{if } |\mathbf{r}_i - \mathbf{r}_j| < r_{\text{cut}} \\ r_{ij'}, & \text{if } |\mathbf{r}_i - \mathbf{r}_j| > r_{\text{cut}} \end{cases} \tag{3.3}$$

where $\mathbf{r}_i, \mathbf{r}_j$ are the positions of the i, jth particles in the original simulation box and $\mathbf{r}_{j'}$ is the position of the nearest image of jth particle. It is important to note here that the cut-off distance has to be less than half the size of the box $L/2$. Afterward, this distance is used in the interaction calculations between pairs of particles.

Let us consider a time increment δt that corresponds to a "time-step" in MD simulation. To perform a MD run basically means to **integrate the equations of motion** 3.1 for each particle in the system. With the information about the position $\mathbf{r}_i(t)$ of each particle at a given time t one can calculate the corresponding force $\mathbf{F}_i(t)$ exerted on ith particle. Next

step would be to get an estimation of the new position $\mathbf{r}_i(t+\delta t)$ and velocity $\mathbf{v}_i(t+\delta t)$ at time $t+\delta t$. For this one can use Taylor expansion:

$$\begin{aligned}\mathbf{r}_i(t+\delta t) &\approx \mathbf{r}_i(t) + \mathbf{v}_i(t)\delta t + \frac{\mathbf{F}_i(t)}{2m_i}\delta t^2 \\ \mathbf{v}_i(t+\delta t) &\approx \mathbf{v}_i(t) + \frac{\mathbf{F}_i(t)}{m_i}\delta t + \frac{\dot{\mathbf{F}}_i(t)}{2m_i}\delta t^2\end{aligned} \tag{3.4}$$

where $\dot{\mathbf{F}} = \frac{d}{dt}\mathbf{F}_i(t)$ and $i = 1..N$.

The above procedure is repeated n_{run} times during a run and corresponds to the number of simulation steps. This leads to a discrete trajectory of the system, $\mathbf{r}(t_{\text{step}} = n_{\text{step}}\delta t)$ with $n_{\text{step}} = 0..n_{\text{run}}$.

On the other hand, one can see that eqs. 3.4 are not completely consistent with the original eq. 3.1 because they lose the time-reversibility. That means if we try to "go back in time", $t \to -t$, we will not be able to get the same solution as before. To fix this issue a more modern and widely used integration algorithm is introduced - the **Velocity Verlet algorithm** [2, 82]:

$$\begin{aligned}\mathbf{r}_i(t+\delta t) &\approx \mathbf{r}_i(t) + \mathbf{v}_i(t)\delta t + \frac{\mathbf{F}_i(t)}{2m_i}\delta t^2 \\ \mathbf{v}_i(t+\delta t) &\approx \mathbf{v}_i(t) + \frac{\mathbf{F}_i(t) + \mathbf{F}_i(t+\delta t)}{2m_i}\delta t\end{aligned} \tag{3.5}$$

with $\mathbf{F}_i(t+\delta t)$ calculated for $\mathbf{r}_i = \mathbf{r}_i(t+\delta t)$, $i = 1..N$.

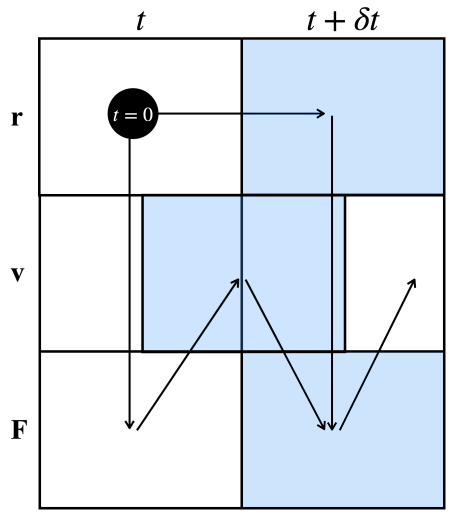


FIGURE 3.2: Schematic illustration of the Velocity Verlet method showing the sequence of the integration algorithm.

In Fig. 3.2 one can see a schematic representation of the Velocity Verlet algorithm. Let us give here the most important steps of this method. After the initial assignment of positions and velocities at time t, the total force on the current step $\mathbf{F}(t)$ is calculated. Next, new positions $\mathbf{r}(t+\delta t)$ are evaluated and current force $\mathbf{F}(t)$ is used to get a half-time velocity $\mathbf{v}(t+\frac{\delta t}{2})$. After this a new force $\mathbf{F}(t+\delta t)$ is computed for all particles and finally the velocity of the next time-step $\mathbf{v}(t+\delta t)$ is obtained. Iterating over this scheme during the simulation time gives the evolution of the positions, velocities, and forces of the system.

It is also important to note here that the value of a time-step δt should not exceed some characteristic

time for the given system so that interactions are computed numerically accurate and no particle is "lost" on the next computation step. In the case of biological coarse-grained simulations, a step of 10 fs is usually a satisfactory choice.

3.1.1 Keeping temperature and pressure constant

The microcanonical ensemble has a fundamental importance but cannot be efficiently implemented on computers. The reason is that the discretization of the Newton dynamics does not conserve the total energy exactly. This needs to be artificially maintained. In addition it is often more relevant in soft matter science and biophysics to study systems in the canonical ensemble, that is at constant temperature. Therefore, the Newton equations need to be modified so the time averaged sampling of the phase and configuration space coincide with a canonical NVT distribution. This must be done in a way that preserves as much as possible the characteristics of the original Newton equations. While in experiments the system exchange energy through its boundaries or radiation, in simulations one introduces a weak coupling of all the degrees of freedom of the system to a heat reservoir called thermostat.

Thermostated simulation do not suffer from drifts in energy, while it is the case for the microcanonical simulations. Thus, such simulations are intensively used for scientific research. However, in some cases it is also desirable to let the volume V fluctuate freely and keep the pressure P constant. This is achieved naturally by allowing the simulation box size L_x, L_y, L_z to change. In other words, the simulation box dimensions become dynamical variables that obey mechanical equations alongside the particle coordinates. As for temperature, the change in box size dimensions must not interfere excessively with the particles degrees of freedom in order to preserve the nature of the Newtonian trajectories as much as possible. Algorithms that performs such control of the box size and coupling to the particles positions are called barostats.

A number of different thermostating algorithms are available [2, 78]: Nose-Hoover, Berendsen, Anderson, Langevin, etc. We use in this work an algorithm provided by LAMMPS (`fix nvt`) that is essentially equivalent to Nose-Hoover (NH) [42]. Let us present here the principle of the NH algorithm. The NH thermostat provides a way to simulate a system that samples the NVT ensemble at large times. Firstly, a friction coefficient $\xi(t)$ is introduced that slows or accelerates particles until the target temperature is achieved. It is important to note here that the temperature is not exactly fixed but is rather enforced

on average. The equations of motion then are as follows:

$$\begin{aligned}
\frac{\mathrm{d}\mathbf{r}_i}{\mathrm{d}t} &= \mathbf{v}_i \\
m_i\frac{\mathrm{d}\mathbf{v}_i}{\mathrm{d}t} &= \mathbf{F}_i - \xi m_i \mathbf{v}_i \\
\frac{\mathrm{d}\xi(t)}{\mathrm{d}t} &= \frac{1}{Q}\left[\sum_{i=1}^{N} m_i \mathbf{v}_i^2 - 3Nk_BT\right]
\end{aligned} \tag{3.6}$$

with Q - thermal inertia coefficient which behaves as an effective mass and k_B the Boltzmann constant. The $\xi(t)$ is a new dynamical variable that extends the Newton equations and plays the role of pumping energy in and out of the system. It is sometimes necessary to couple separately different subgroups of particles to different thermostats, to avoid the so called "hot solvent"-"cold solute" problem.

To realise a simulation of a system in the isobaric-isothermal NPT ensemble one can use an extension of eq. 3.6 proposed by Martyna, Tobias and Klein [43]. The isotropic barostat implemented in LAMMPS essentially performs the same task (`fix npt`). The barostat extends the isothermal NH scheme by introducing the expansion rate $\zeta = \frac{\dot{V}}{3V}$ to describe changes in volume V with time as a new dynamical variable. The equations of motion change as follows:

$$\begin{aligned}
\frac{\mathrm{d}\mathbf{r}_i}{\mathrm{d}t} &= \frac{\mathbf{p}_i}{m_i} + \zeta(t)\mathbf{r}_i \\
\frac{\mathrm{d}\mathbf{p}_i}{\mathrm{d}t} &= \mathbf{F}_i - \left(1 + \frac{3}{N}\right)\zeta(t)\mathbf{p}_i - \xi(t)\mathbf{p}_i \\
\frac{\mathrm{d}V}{\mathrm{d}t} &= 3V\zeta(t) \\
\frac{\mathrm{d}\zeta}{\mathrm{d}t} &= \frac{3V(t)}{W}(P(t) - P_e) + \frac{3}{WN}\left(\sum_{i=1}^{N}\frac{\mathbf{p}_i^2}{m_i}\right) - \xi(t)\zeta(t) \\
\frac{\mathrm{d}\xi}{\mathrm{d}t} &= \frac{1}{Q}\left(\sum_{i=1}^{N}\frac{\mathbf{p}_i^2}{m_i} + W\zeta^2(t) - (3N+1)k_BT\right)
\end{aligned} \tag{3.7}$$

where P_e is a fixed external pressure imposed on the system, W is the relaxation rate of the barostat that determines the frequency of volume fluctuations, and $P(t)$ is an instantaneous pressure of the system defined as a virial for pairwise additive interactions:

$$P(t) = \rho k_B T + \frac{1}{3V}\langle\sum_{i<j}\mathbf{F}_{ij}\cdot\mathbf{r}_{ij}\rangle \tag{3.8}$$

and $\rho = \frac{N}{V}$ is the total density of the system. It has to be noted here that in the above MD equations the dimension d is 3 and V is the volume of the simulation box. Thus, there is

an additional degree of freedom associated with the dynamical variable ζ.

Simulations of lipid bilayers require additional precautions which are linked to the anisotropy of the molecular systems. Restricting the volume dynamics to only isotropic changes makes it impossible to enforce simultaneously the solvent pressure P_z and the membrane tension σ. It is important to use a semi-isotropic barostat. It means that the box size can fluctuate independently in the x, y and z directions (if the membrane is oriented perpendicular to the xy plane) and that an initially cubic cell does not stay cubic but becomes tetragonal.

In LAMMPS this can be simply achieved by modification of `fix npt` command in the input script as shown below:

```
fix 1 all npt temp Tstart Tstop Tdamp
      x Pstart Pstop s
      y Pstart Pstop Pdamp
      z Pstart Pstop Pdamp
      couple xy
```

where `all` specifies that the command is applied to the whole system, T_{start}, T_{stop} are the external temperatures at start/end of run, T_{damp} is the temperature damping parameter (time units), P_{start}, P_{stop} are the external stress tensor components at start/end of run (pressure units), P_{damp} is the stress damping parameter (time units), `couple xy` specifies coupling of xy stresses. By damping parameters here we mean relaxation time, that determines the time scale on which the corresponding T or P is relaxed. In the current work T_{damp}=500.0 fs and P_{damp}=5000.0 fs.

3.2 SPICA force field

Let us define a force field (FF) before moving on to the detailed description of SPICA. To be precise, a force field is an empirical description of the forces that originate from the quantum dynamics of atoms and electrostatic interactions of static charges on them. There are different types of FFs available currently: from fully atomistic models that take into account all details of the given molecule [32, 37] to extreme coarse-grained models that simulate a whole molecule as one point [13, 15]. They all have their pros and cons. For example, all-atom models give much more realistic and detailed results concerning molecular structure and lipid-protein interactions, but they require enormous computational power. On the other hand, coarse-grained simulations are much faster and can

achieve longer physical times while giving reliable results on thermodynamics and the macroscopic structure of the system.

A FF is "shaping" the molecule, determining all interactions between different atoms connected by bonds (bonded interactions) and between atoms that are not connected by bonds (nonbonded interactions). These interactions are also called intra molecular and inter molecular respectively.

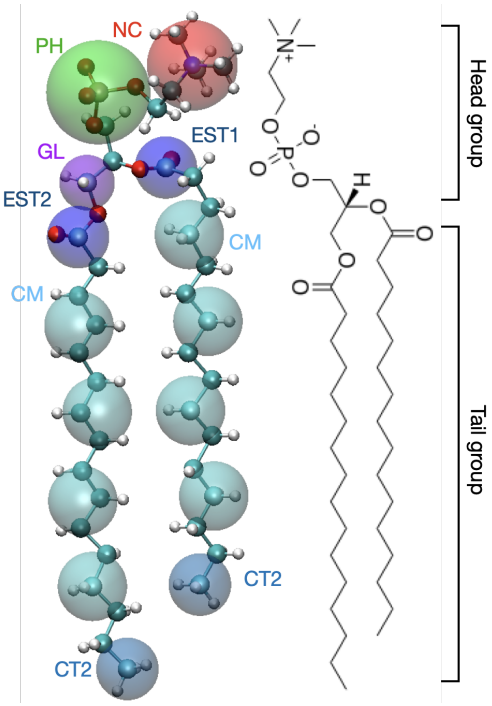


FIGURE 3.3: Schematic mapping of atomistic representation of DPPC molecule to SPICA force field beads.

As was mentioned before, in this project the SPICA coarse-grained force field [69] was used. An example of mapping of an atomistic DPPC molecule to SPICA DPPC is shown in Fig. 3.3. In SPICA FF, on average 3 real atoms are mapped into 1 SPICA bead and 3 water molecules are mapped into 1 SPICA water bead. Beads in this CG model have different masses which correspond to the atomistic masses of atoms they consist of, making this FF closer to the AA representation. Consecutive beads are connected by bonds, with an angle between the bonds for further beads. A description of SPICA parameters can be found in the Appendix B.

Let us look at the interactions included in the SPICA model in more detail. A harmonic bond stretching and angle bending potentials for 1-2 and 1-2-3 (see Fig. 3.5) bonded pairs are considered for **intra molecular interaction**:

$$\begin{aligned} U_{1-2} &= \sum_{\text{bond}} k_b (r_{ij} - r_0)^2 \\ U_{1-2-3} &= \sum_{\text{angle}} k_\theta (\theta_{ijk} - \theta_0)^2 \end{aligned} \tag{3.9}$$

where k_b and k_θ are the force constants, r_{ij} and θ_{ijk} - the distance and angle between i,j-th and i,j,k-th consecutive beads respectively, and r_0 and θ_0 are the distance and angle at the minimum energy, respectively. By 1-2 and 1-2-3 here we mean interactions between two (bonds) and three (angles) consecutive bonds in one molecule.

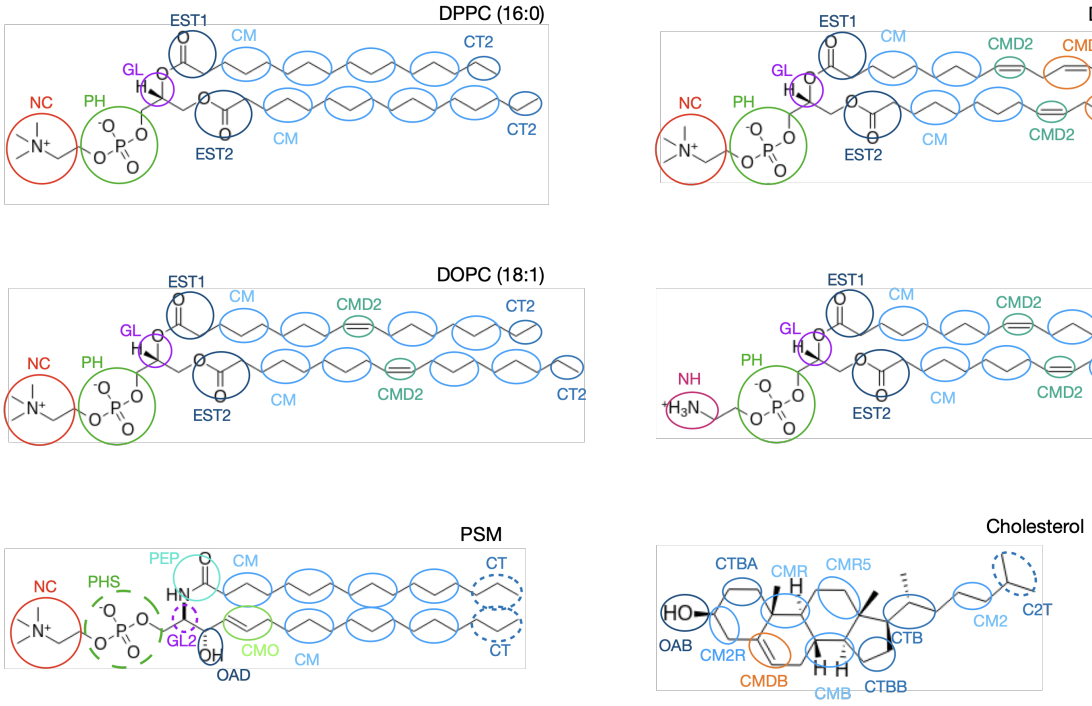


FIGURE 3.4: Schematic mapping of chemical formulas of DPPC, DOPC, DOPE, DLiPC, PSM, and Cholesterol to SPICA force field beads. Sizes of beads represent mass distribution in the lipid.

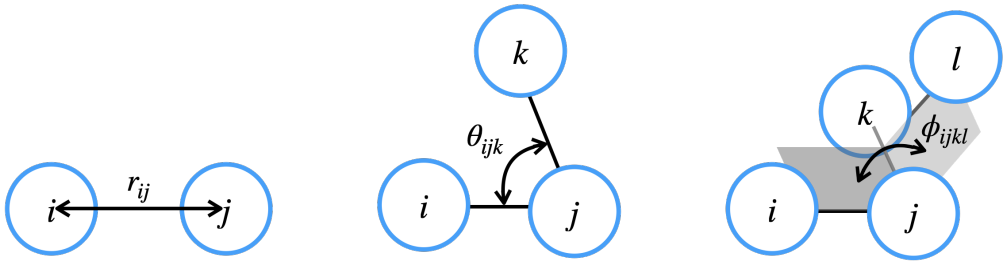


FIGURE 3.5: Illustration of different types of interactions included in the model (from left to right): bond stretching, angle bending, bond rotation or torsion (dihedrals).

Torsion potential is not used a lot in SPICA but is nevertheless applied to selected pairs in the sterol tail (for example in Cholesterol):

$$U_{1-2-3-4} = \sum_{\text{dihedral}} k_\phi \left[1 + \cos\left(n(\phi_{ijkl} - \phi)\right)\right] \tag{3.10}$$

where k_ϕ is the multiplicative constant, ϕ_{ijkl} is the torsion angle, ϕ is the phase shift, n is the integer constant that indicates the periodicity, and i, j, k, l corresponds to 1, 2, 3, and 4th consecutive beads.

For the pairs separated by more than two bonds **nonbonded interactions** are used. Depending on the pair of particles nonbonded interaction is described by the Lennard-Jones

(LJ) function and/or Coulomb interaction:

$$U_{\mathrm{LJ}}(r_{ij}) = \begin{cases} \frac{3\sqrt{3}}{2}\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{4}\right], & \text{pairs with water} \\ \frac{27}{4}\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{9} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right], & \text{all other pairs} \end{cases} \tag{3.11}$$

where r_{ij} is a distance between particles i and j, ϵ_{ij} is the depth of the potential well and σ_{ij} is a relative particle size. The cut-off for SPICA is $r = r_{\mathrm{cut}} = 1.5$ nm.

The Coulomb interaction is:

$$U_{\mathrm{C}}(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r r_{ij}} \tag{3.12}$$

where q_i, q_j are the charges of the particles i, j, ϵ_0 is the dielectric constant, and $\epsilon_r = 80$ is set to incorporate the screening effect of water in this model.

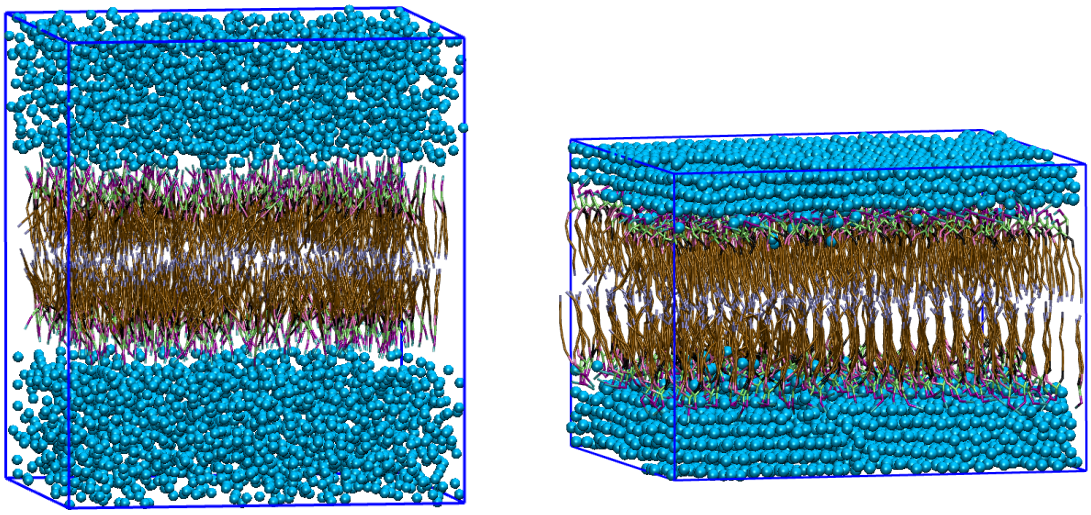
3.3 Simulation details

Let us discuss in more detail simulation that was performed during this work. We introduce here a protocol of the initial configuration creation and the typical molecular dynamics run. Firstly, we discuss the importance of the primary configuration of the bilayer membrane and the difficulties one might encounter during this step. Next, we lead the reader through our typical MD run and introduce specific protocols applied for compressibility and external potential cases.

3.3.1 Getting the initial configurations

As we are interested in the thermodynamic of mixing, we need to create bilayer systems of controlled composition. Firstly, we had to build a bilayer sheet in water solution, while maintaining the magnitude of Area per lipid (APL) (see Chapter 2) and density of such systems. The first criteria is needed to obtain x, y dimensions and the second one - to find a z size of the system. A good guess while creating initial configurations can help one build an efficient simulation with faster equilibration.

Two methods for creating the initial configurations in this project were tested: using Packmol and MolTemplate softwares. In both methods, the same number of particles and initial box sizes (densities) were used. We found that the Packmol package produces randomly organized configurations while MolTemplate creates highly ordered structures. For both cases, a predefined fixed bilayer was put between two water slices. By predefined bilayer here is meant that the head group and last beads in the tail were constrained in a xy plane in such a way that every lipid is oriented along the z direction. Water was



(A) Packmol. Bilayer in a fluid state. (B) MolTemplate. Bilayer in gel/fluid state.

FIGURE 3.6: Snapshot of the initial configuration of 512 DPPC lipids and 4096 solvent beads after a 100ns equilibration run at $T = 298.15$K. Water is marked cyan, the tail group is brown and head group beads have blue (NC), violet (PH), and mint (GL) marking. Ordering can be seen in panel (B) which signals that the bilayer is in the gel phase and water is crystallized. Water beads are shown to highlight the ordering in the (B) case.

placed above and below these planes (see Fig. 3.6). As a consequence lipid bilayers obtained from MolTemplate were in the gel phase and required a melting procedure with a follow-up cooling to reach the fluid phase. On the contrary, Packmol generated fluid-like disordered initial configurations that did not require additional heating-cooling protocol. As in this work we are more interested in lipids in the fluid phase, we have chosen to work with the Packmol software for further configurations creation.

For our purpose, a Python script to create input files for Packmol was developed. This utility takes Protein Data Bank (PDB) structures adapted for SPICA FF with desired number of lipid and water molecules, and creates a file that can be directly used with Packmol. The parameter for APL is approximated to be 70 Å^2/lipid, which is in typical magnitude in phospholipid systems. Density is taken as $\rho = 3.6$ particle/nm^3 so that the overall packing of the initial system is close to a thermalized one. This parameter was chosen after preliminary tests on pure DPPC bilayers.

To create mixtures of different lipid types we used the same software but with different amount of lipids. Packmol makes the creation of standard mixture systems exceedingly simple, while maintaining the bilayer structure. In this project we considered the following family of lipids: phospholipids, sphingomyelins, phosphoethanolamine, and

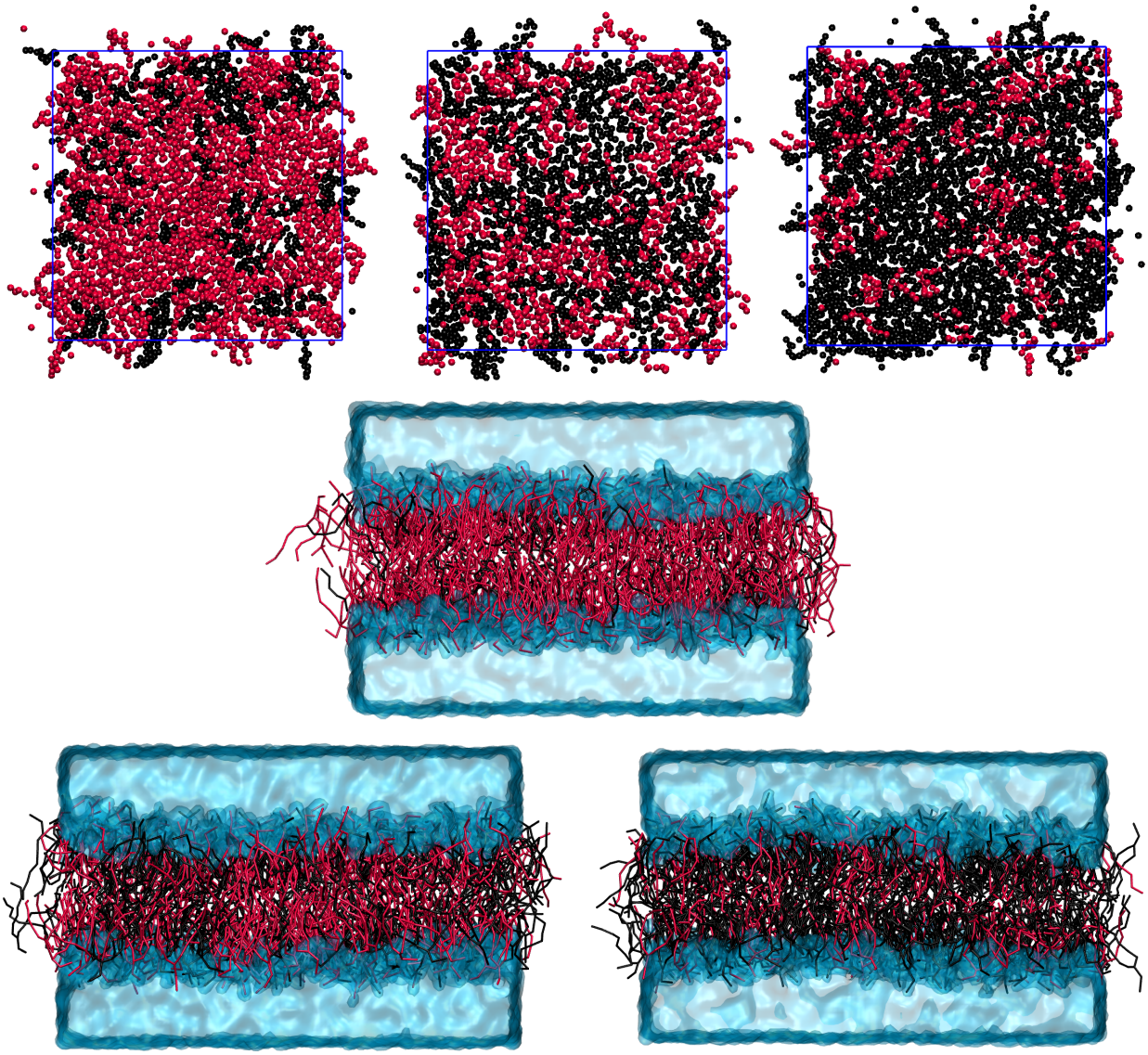


FIGURE 3.7: Configurations of DLiPC/DPPC mixtures after equilibration of 10μs with different ratios: 1:3, 1:1, 3:1 from left to right. Top and side views are shown. Top view shows particles in the upper leaflet only, so domain formation in a monolayer can be observed. DLiPC lipids are marked black, DPPC - red, and water molecules - cyan (on the top view are not shown). The total number of lipids is 512 and water beads is 8192.

sterol. In Fig. 3.4 one can see a mapping of all the lipids used to the SPICA FF: 1,2-dipalmitoyl-sn-glycero-3-phosphatidylcholine (DPPC) ((16:0), saturated), 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) ((18:1), unsaturated), 1,2-dilinoleoyl-sn-glycero-3-phosphocholine (DLiPC) ((18:2), unsaturated), 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE) ((18:1), unsaturated), Palmitoyl sphingomyelin (PSM) and Cholesterol. In Fig. 3.7 one can see an example of DLiPC/DPPC system at different concentrations after 10 μs NVT run. For pairs with Cholesterol we have investigated concentrations of 1:3 and 1:1 (see Fig. 3.8). In the following text when talking about concentration we make

use of the X:Y nomenclature, meaning that X molecules of one lipid type are found for Y molecules of the second lipid type. As an example, a 1:3 A/B mixture corresponds to a mixture with 25% lipid A and 75% lipid B.

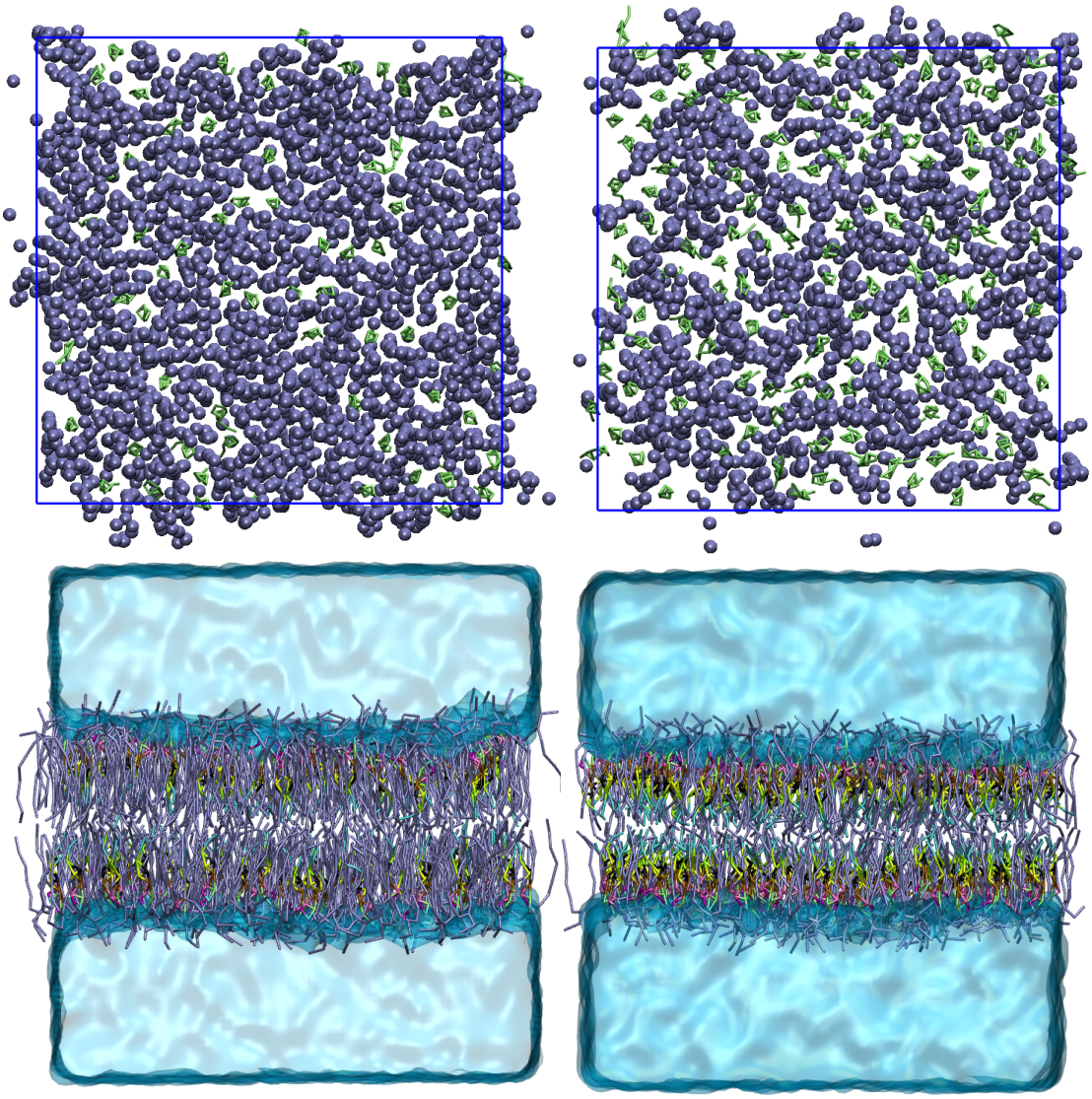


FIGURE 3.8: Configurations of SSM/Cholesterol mixtures after equilibration of 10μs with different ratios: 25% Cholesterol, 50% Cholesterol from left to right. Top and side views are shown. SSM lipids are shown in iceblue, Cholesterol in lime, and water molecules - in cyan (on the top view are not shown). The total number of lipids is 512 and water beads is 8192.

One of the first challenges encountered during this work was to find the suitable size of the membrane to study. Our systems had to be large enough to gather sufficient statistics for the calculations and probe collective fluctuation modes. On the other hand, the bilayer should not have too many undulations in the z direction (see Chapter 2). Firstly, pure systems of 64, 128 and 256 lipids were created and equilibrated. It was found that the statistics of 64 and 128 bilayers was not satisfying and consistent values for zero-limit

TABLE 3.2: System configuration for the studied lipids.

Lipid		N_{lipid}	$N_{\text{water}}^{\text{beads}}$	NVT time
DPPC		1024, 512, 256	8192, 4096	10 μs
DLiPC		512	8192	10 μs
DOPC		512	8192	100 ns
DOPE		512	8192	100 ns
PSM		512	8192	100 ns
SSM		512	8192	100 ns
Lipid pair	Ratio	N_{lipid}	$N_{\text{water}}^{\text{beads}}$	NVT time
DLiPC/DPPC	1:3, 1:1, 3:1	1024, 512	16384, 8192, 4096	10 μs
DPPC/Chol	3:1, 1:1	512	8192	10 μs
DOPC/Chol	3:1, 1:1	512	8192	10 μs
DLiPC/Chol	3:1, 1:1	512	8192	10 μs
DOPE/Chol	3:1, 1:1	512	8192	10 μs
SSM/Chol	3:1, 1:1	512	8192	10 μs
PSM/Chol	3:1, 1:1	512	8192	10 μs

structure factor could not be obtained (see Chapter 2 for the definition of the structure factor). Considering the above arguments, we tried the same protocol with 512 and 1024 pure lipid membranes. It was found that both were sufficient for our studies, while the 1024 system does not significantly improve the statistics but takes a longer time to equilibrate. The gain in $\mathbf{q}$ range is somewhat compensated by the loss in statistical accuracy. Taking into account all of the above statements, we concluded that for pure bilayer systems it was sufficient to study systems of 256 and 512 lipids in total. Although the system size plays an important role as far as pure membranes results are concerned, the total simulation time was found to be less influential. We obtained small to no differences in the calculations for pure bilayers of DPPC when increasing the run length from 100 ns to 10 μs. In the end it was decided that for the pure lipid bilayers, a 100 ns NVT run would be sufficient.

In the case of the mixtures of two phospholipids, we found that even systems of 512 lipids in total are still "small" and did not give the desired precision. We analyzed systems of 1024 lipids in total and checked the system size effects for the pair of DLiPC and DPPC during NVT 10 μs runs. For the pairs with Cholesterol, it was found that systems of 512 lipids in total were enough due to the strong interactions between Cholesterol and lipid molecules. A detailed list of all studied compositions is given in Tab. 3.2.

3.3.2 Simulation protocol

We performed the MD simulations using the Velocity Verlet algorithm in an orthogonal box with periodic boundary conditions at $T = 298.15$ K and $P = 1$ atm. The time-step for the phospholipid and sphingolipid simulations was $\delta t = 10$ fs, while for the runs with Cholesterol it was $\delta t = 5$ fs. A schematic representation of the simulation protocol can be seen in Fig. 3.9.

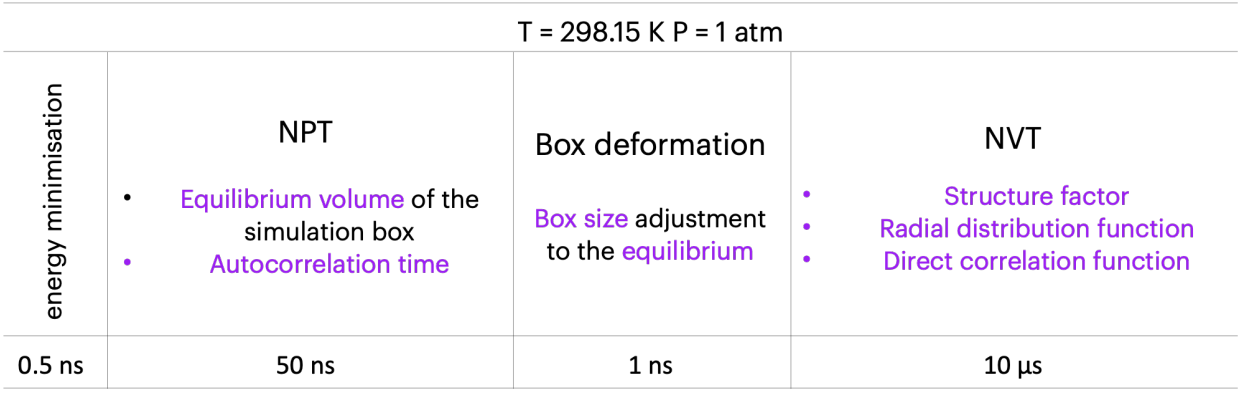


FIGURE 3.9: A schematic representation of the simulation protocol used in this work.

A general simulation begins with a very short energy minimization run of 0.5 ns followed by a 50 ns NPT run. The last 25 ns of the latter are used to calculate the average volume of the simulation box for a given pressure value and to estimate the autocorrelation time of the given lipid system. To confirm that a system is in an equilibrium state, the total area was checked as a function of time until the plateau is observed. Next, a short box deformation run is performed to softly adjust the system size to its equilibrium NPT volume. This part is needed to smoothly change the system volume and avoid volume jumps on the next step. Finally, a comparably long equilibrium NVT run is done. As was mentioned before, for most of the systems we used runs of 10 μs total. During this last step, we calculate the structure and thermodynamic parameters of interest (see Chapter 2).

It is worth mentioning here that all simulations were done on the Jean Zay IDRIS supercomputer which is located in Orsay, France (grant A10 A0120712495). This allowed us to use parallelization to speed up calculations and obtain results faster and more efficiently. Unfortunately, even with all the available resources, our system had its own constraints. Due to the high demand of electrostatic interactions for the given system sizes, it was not efficient to parallelize on more than 40 cores. That lead to a limitation of approximately 200 ns calculation per one simulation package. We have established the procedure of sending successive runs of 50 packages to achieve the desired 10 μs trajectory, which took

approximately two weeks to finish. Further analysis was performed after merging these trajectories into a single one.

3.3.3 Stretching of the membrane

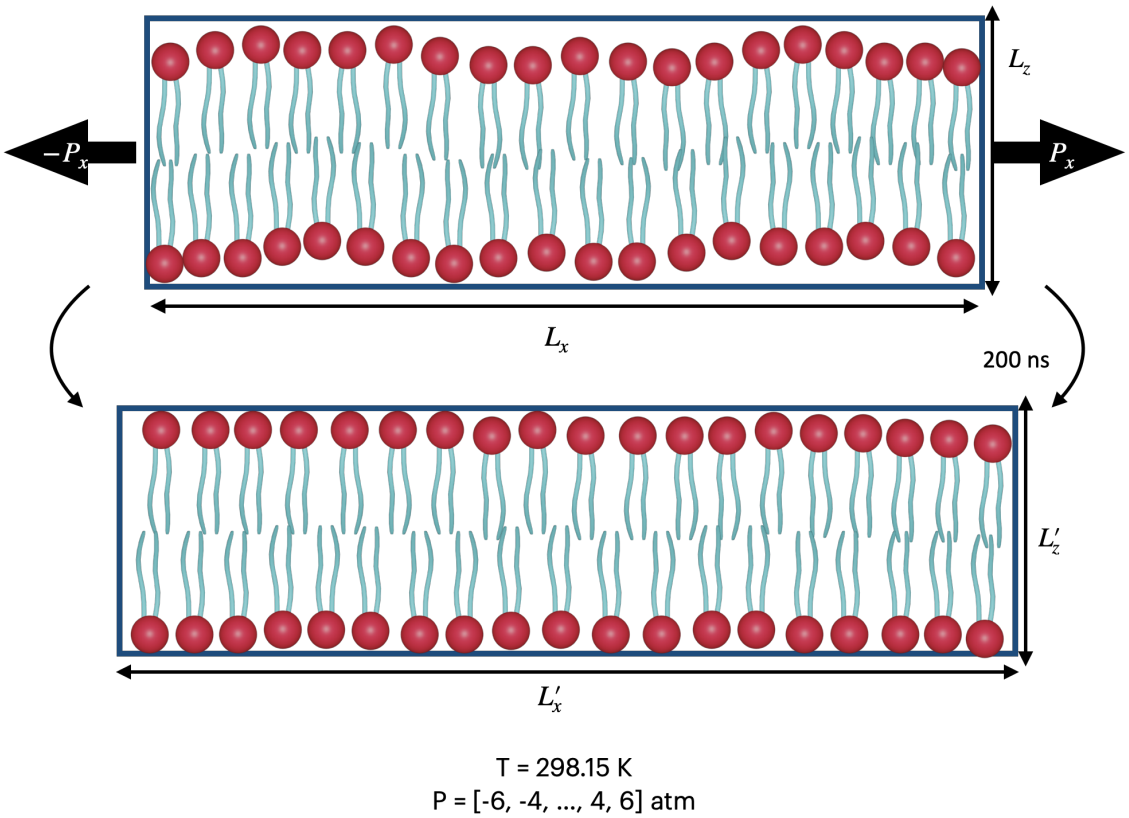


FIGURE 3.10: A schematic representation of the simulation protocol used for the membrane stretching.

To determine the compressibility of the membrane a stretching protocol was performed (see Fig. 3.10). We used pressures $P = -6, -4, ..., 4, 6$ atm applied in the xy plane as a form of direct membrane stretching. This was done using separate keywords in the LAMMPS barostat command. For example,

```
variable P equal ${Px} # Px is given by the submitting script
     fix 1 all npt temp 298.15 298.15 500.
           x ${P} ${P} 5000.  y ${P} ${P} 5000.  z 1.  1.  5000.  couple xy
```

means that temperature is fixed at $T = 298.15$ K, pressure along x and y direction is $P_x = P_y = P$, along z is 1.0 atm, damping parameters are $T_{\text{damp}} = 500, P_{\text{damp}} = 5000$ for all particles in the system. Here, as was mentioned above, an anisotropic barostat is used, i.e. xy are coupled together. These were 5 short 200 ns runs for each pressure value

that were averaged afterward. For a detailed explanation of the analysis please refer to Chapter 2.

3.3.4 External potential

As was proposed in Chapter 2 we can also apply an external force to bias the system and calculate the structure factor from the linear response of the density modes. By an external force here we mean a force that is applied to specific beads, to the Center of mass (COM) or Pseudo center of mass (pCOM) (this is explained in Chapter 4) or to all the atoms in the bilayer. To apply an external force in the simulation one uses the `fix addforce` command of LAMMPS. This command allows us to give an additional "push" to atoms in a given direction by specifying each of the f_x, f_y, and f_z components of the total force. It is also possible to specify atom groups to which the force is applied and thanks to that one could implement the external force in two ways: applied to a selected bead in each lipid molecule or to the whole lipid.

There are different possibilities to treat lipids during the MD simulation. LAMMPS gives us the possibility to explore all kinds of cases as the `addforce` command is applied to each particle of the simulation system, it is the so-called "per-atom" quantity. Considering that each bead in a lipid (Fig. 3.3) is treated as a separate atom in simulation, we can apply a force to any of them, or to some of them at the same time. So far we have explored the case of a force $\mathbf{F}(x)$ applied to the GL bead and all the beads of each lipid. Two cases are technically the same and only differ in the prefactor of the $\mathbf{F}(x)$ expression, meaning that in the latter situation we have to normalize by the total number of beads in the lipid to have the same total force applied as in the case of one bead. In addition, the force has to be normalized by the total number of lipids in the system in both cases.

One way to apply a bias to the system during a LAMMPS simulation is to use the Python interface. That means that every N timesteps LAMMPS calls a specified Python script and perform commands inside it. In the scope of the current project, this feature could be used to apply the external force directly to the COM or pCOM of the lipids. As per the knowledge of the writer, native LAMMPS functionality does not provide one with such ability. It is indeed possible to calculate COMs of separate molecules, whole leaflets, etc. but it is not possible to apply per-atom transformations to such objects. On the other hand, constant discussion between LAMMPS and Python is computationally costly. For example, a simulation with a periodic external potential applied to COM took approximately 10 hours for 10 ns, while the same field applied to the bead of the lipids took approximately 8 hours for 100 ns. We did obtain consistent results for COM and pCOM simulations with Python, but in the end, decided to use the native LAMMPS tools and

further external force simulations will be discussed from the context of the `addforce` command. The best option is to apply the force to all the lipid beads proportionally to their mass (in case of COM) or equally distributed among beads (in case of pseudo COM).

It is possible to apply a force corresponding to a few different modes at the same time. We call this "multiplexing". This is done simply by including all the modes explicitly into the calculation of $F(x) = U_0 q_1 \sin(q_1 x) + U_0 q_2 \sin(q_2 x) + \ldots$. We went up to 7 modes applied simultaneously in the pure case and up to 3 modes in the case of mixtures. Even though the algorithm in the case of phospholipid mixtures is a bit more complex than for the pure system, the idea behind the simulation is the same (see Chapter 2). The external force was applied with the use of the `addforce` command in both situations (see Appendix D for an example of the LAMMPS script).

3.4 Error treatment

Let us introduce here the idea behind the estimation of the confidence of the obtained results. Bilayer membranes are complex objects with multiple fluctuation degrees of freedom that might lead to high errors during observable calculations. In this section we show how one can estimate such errors based on the auto correlation time or resampling of the trajectories. First method allows us to predict the necessary time interval to perform resampling and can also provide some confidence interval information. Calculations of auto correlation were done for box size fluctuations during the fixed pressure run, while resampling was performed for fixed volume long runs.

3.4.1 On the integrated auto correlation time

For proper handling of such complex systems as a membrane, we have to be sure that all calculations are done after the system has reached the equilibrium state. This can be checked knowing the relaxation time for a given system and a given observable. In our case an **integrated auto correlation time** (τ, IATF) of the box size was used [75] to estimate if the system reached its equilibrium state. The IATF provides a useful estimate of the time required for considering that two consecutive intervals are effectively uncorrelated. Such long time intervals can then be used in a bootstrap resampling approach to find out the confidence interval of a sampled average.

Let us define an unnormalized autocorrelation function:

$$C(t) = \langle f_s f_{s+t} \rangle - \mu^2 \tag{3.13}$$

with $\{f_t\}$ being real-valued stationary stochastic process and μ - its mean.

Then integrated autocorrelation time can be defined as

$$\tau_{int} = \frac{1}{2}\sum_{t=-\infty}^{\infty}\frac{C(t)}{C(0)} = \frac{1}{2}\sum_{t=-\infty}^{\infty}\rho(t) \tag{3.14}$$

where $\rho(t)$ is the normalized autocorrelation function.

According to ref. [75] the "natural" estimator of μ would be

$$\bar{f} = \frac{1}{n}\sum_{i=1}^{n} f_i \tag{3.15}$$

and thus its variance

$$\mathrm{var}(\bar{f}) = \frac{1}{n}\sum_{t=-(n-1)}^{n-1}(1-\frac{|t|}{n})C(t) \approx \frac{1}{n}2\tau_{int}C(0) \tag{3.16}$$

where infinite summation is replaced by a sum over $n \gg \tau_{int}$ - simulation time.

Before calculating any observable and its variance we have to establish how to find τ_{int}. In analogy with previous statement "natural" estimator for τ_{int} is

$$\tau_{int} = \frac{1}{2}\sum_{t=-(n-1)}^{n-1}\rho(t) \tag{3.17}$$

But as is stated in ref. [75] this is wrong. The estimator defined in eq. 3.17 is not accurate as, roughly speaking, the sample autocorrelations for $|t| \gg \tau_{int}$ are mostly noise rather than real signal data. The solution is to cut off the sum using "window" $\lambda(t)$:

$$\tau_{int} = \frac{1}{2}\sum_{t=-(n-1)}^{n-1}\rho(t)\lambda(t) = \begin{cases} \frac{1}{2}\sum_{t=-(M-1)}^{M-1}\rho(t), & |t| \leq M \\ 0, & |t| > M \end{cases} \tag{3.18}$$

where M is some arbitrary lag time limit. In the current project this parameter was chosen by the following algorithm. The lag time for the calculation of $C(T)$ was gradually increased until $C(T)$ reaches ≈ 0.5. Then the corresponding lag value was multiplied by a factor of 10 and τ_{int} was calculated on the obtained interval.

Thus, the variance of the sample averaged observable can be expressed as

$$\mathrm{var}(\bar{f}) = \frac{1}{n}2\tau_{int}C(0) = \frac{1}{n}2\tau_{int}\sigma^2 = \frac{1}{n}4\tau_{exp}\sigma^2 \tag{3.19}$$

where σ is the original data variance and τ_{exp} is the exponential autocorrelation time we measure.

Finally, the standard deviation for the sample average of this observable should be

$$\mathrm{std}(\bar{f}) = \frac{2\sigma}{\sqrt{n/\tau_{exp}}} \tag{3.20}$$

we can estimate the relaxation time of our system with τ_{exp} and define a proper time sampling.

TABLE 3.3: Typical relaxation times of studied systems from the area fluctuations.

	τ, ns
DPPC	0.65
DLiPC	0.4
DLiPC/DPPC 1:3	0.55
DLiPC/DPPC 1:1	0.50
DLiPC/DPPC 3:1	0.45
DLiPC/DOPC 1:3	0.75
DLiPC/DOPC 1:1	0.42
DLiPC/DOPC 3:1	0.44

3.4.2 Bootstrap analysis

Even though nowadays technology is developing very fast, computational power that is available for research is still limited. It matters to get the best of the available simulation time and to quantify the confidence in an averaged observable sampled during a finite time. In order to achieve this we decided to resort to one particular resampling method: the bootstrap algorithm [14]. The central idea is to use the only available information (the available sample) to generate plausible synthetic samples and estimate the variability of the resampled averages. In practice this means slicing a long sample into a few shorter samples and creating artificial data sets using parts of the original data set. Synthetic samples are obtained by drawing randomly with replacement the small samples and building new long ones, which have the same length as the original data. Such kind of resampling is usually called sampling with a replacement. If the short samples of the data set are longer than the observable autocorrelation time, they can be considered as statistically independent, and the bootstrap procedure makes it possible to estimate the intrinsic variability of the sampled average. This variability is converted into a confidence interval by multiplying by 2 the corresponding mean square deviation (MSD).

In this work, trajectory of particles in the systems was used as a primary data set. Fig. 3.11 shows an example of the resampled trajectory for the upper and lower leaflets of one configuration. The bootstrap procedure does not require the integrated auto correlation time (τ, IACT) but simply assumes it is shorter than the short trajectory slices.

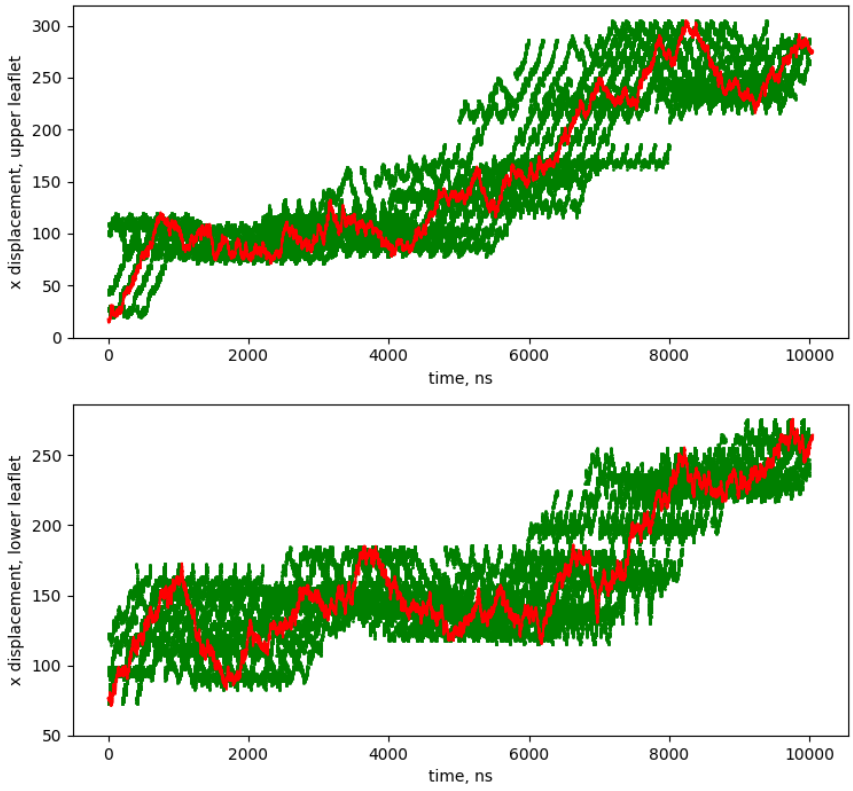


FIGURE 3.11: An example of a bootstrapped trajectory for the upper and lower leaflet.

Let us estimate the confidence interval for such measurements. We consider a set of N samples $A = \{x_i\}$, $i = 1...N$ as original data obtained during simulation. Using the bootstrap analysis one can generate M artificial samples $A^\alpha = \{x_i^\alpha\}$, $i = 1...N$, $\alpha = 1...M$. Then the variability of the average reads

$$\langle f \rangle^\alpha = \frac{\sum_{i=1}^{N} f(x_i^\alpha)}{N} \tag{3.21}$$

Finally, we obtain a confidence interval 2σ for the sampled average with the following expression:

$$\sigma^2 = \frac{1}{M-1} \sum_{\alpha=1}^{M} \left(\langle f \rangle^\alpha - \frac{1}{M} \sum_{\beta} \langle f \rangle^\beta \right)^2 \tag{3.22}$$

with large enough M. In the scope of the current research, $M = 10$ gives satisfactory results. As can be seen from Tab. 3.3, studied systems show low τ compared to the whole trajectory length. For example, even with a run of 100 ns, diving it by $M = 10$ means that we will average over slices of 10 ns each, which is magnitude higher than the obtained τ.

Let us also introduce here a somewhat intermediate approach between bootstrap and IACT estimates. It consists of considering a long sample, sliced into 5 parts, thus, one can estimate the variance of the averages associated with the 5 shorter samples. If the sampled average obeys the same convergence law as in eq 3.20 then the mean square deviation calculated from the 5 short samples is expected to be of the same order of magnitude as the 2σ confidence interval associated with eq 3.20. It can help us avoid IACT estimation and complex bootstrap procedure, while still giving proper error estimate.

CHAPTER 4

Pure lipid bilayers

In this chapter, we apply our density mode fluctuation analysis to the pure bilayer systems in the fluid phase. Main focus is given to the pure DPPC membrane as it was used as a model system to establish the framework of the current project. We tried and compared various approaches that were discussed before. The number of lipid molecules was varied from 256 to 1024 in both leaflets, each leaflet has equal number of lipids. We systematically compared a number of different ways of representation and projection of the DPPC molecules onto the horizontal plane (here, xy plane). This pure bilayer system also served as a "laboratory" for estimating uncertainties, either by dealing with the time autocorrelation of the sampled observables, or a bootstrap statistical approach (see Chapter 3). Most of the results presented here will be discussed for the DPPC pure bilayer, although the simplified version of the analysis was also performed for other model lipids: DLiPC, DOPC, DOPE, PSM and SSM.

As was discussed previously, it is interesting to consider application of the theory of replicas to the lipid bilayers. Thus, we studied here the influence of the coupling between the leaflets by comparing our results if both leaflets or only one is considered for calculations. We tried to assess the degree of coupling between the two "replicas-leaflets" based on the deviations of the obtained data. One of the interesting characteristics to be compared is the bilayer area compressibility. This quantity was computed using the method of area fluctuations, the direct application of an anisotropic stress and the extrapolation to the origin of the structure factor. Comparison of the different methods is instructive as it can indirectly show the influence of the coupling between the monolayers.

In the case of the pure DPPC bilayer, we also tested directly the potential energy biasing method, or density linear response, to compare with the usual equilibrium fluctuation method for computing structure factors. We tested the possibility to apply the potential to one of the lipid beads, its center of mass or all beads at the same time.

4.1 Position of the problem

Our simulations need to reconcile two opposite requirements. The system must be as big as possible to approach the thermodynamic limit and reach the smallest possible wave-vectors $\mathbf{q}$. On the other hand, the relaxation times $\tau(\mathbf{q})$ of such modes diverge as $\mathbf{q}$ tends to 0. The quality of the available statistics is proportional to $\tau(\mathbf{q})^{-1}$. In other words, the simulation time required to collect an equivalent number of "statistically independent" sampling values of a given observable is proportional to $\tau(\mathbf{q})$. This observation is true for any simulated system. Another important feature of the lipid bilayers is that the roughness of the bilayer and the out-of-plane fluctuations tend to increase with the lateral size of the membrane. It is our intention to maintain at the lowest possible level the influence of such undulations on the resulting properties of the bilayer. For this purpose we could increase the membrane tension or introduce a vertical confining potential, but this would potentially alter the outcome of the analysis and take us away from a natural equilibrium membrane state. Finally, we established the best sizes for the current project goals to be 512 lipids in total in the whole bilayer with symmetric leaflets consisting of equal number of lipids per leaflet. Simulation time for the pure DPPC and DLiPC were 10 μs while for the other pure lipid bilayers it was decided to go only up to 100 ns.

Let us discuss now in more detail the pure 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (**DPPC**, 16:0) bilayer. As was defined previously, in Fig. 3.3 one can see the mapping of DPPC lipid to SPICA FF, and find the mechanism of initial bilayer formation. All simulation systems were created as a bilayer sheet between two water slabs. Due to the PBC it is not completely right to separate water region into two slabs, so here mean it artificially with wrapped coordinates in mind.

Firstly, let us properly define a bilayer system as a combination of four fluids as shown in Fig. 4.1. Such representation gives us the possibility to calculate membrane properties of the whole bilayer or each leaflet separately. This approach can be justified if we look deeper into the theory of replicas (see Chapter 2). The main idea would be that in the absence of interactions between different leaflets these systems are basically a replication of themselves and so such properties as compressibility would be simply twice the properties of each leaflet. In our case, weak interaction is still present but we can neglect it as will be shown later.

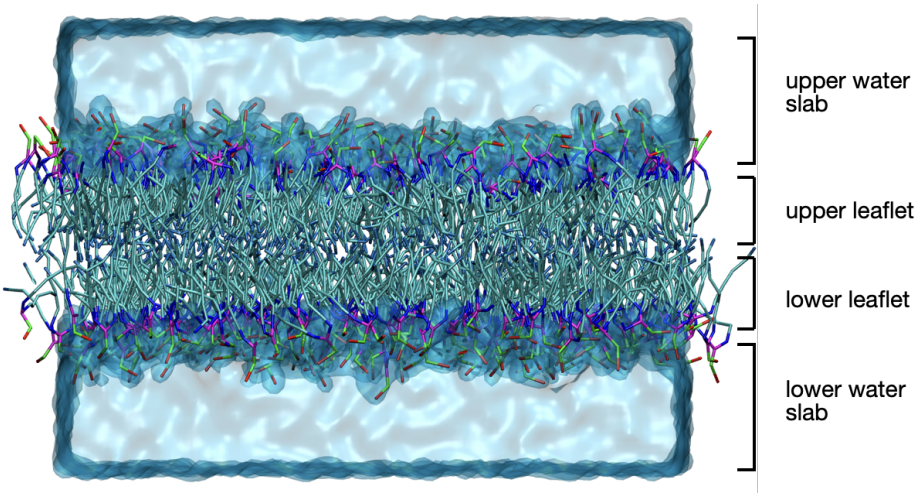


FIGURE 4.1: Bilayer system as a four-component fluid: two monolayers encapsulated into two water slabs in a form of a "burger".

Fig. 4.1 shows a snapshot of a 512 DPPC system at 298 K and 1 atm, hydrated with 8192 water beads or equivalently 16 water beads per lipid. In this situation two slabs of water in excess surround the bilayer. It is commonly admitted that 20-24 water molecules per lipid are sufficient to fully hydrate a DPPC lipid in the fluid state. Given the 3 to 1 SPICA coarse-graining level, it requires at least 8 water beads to reach full hydration so in the case of 16 SPICA water beads per lipid we are in a favorable situation of 48 water molecules per lipid.

Fig. 4.2 presents the number density of lipid beads as a function of their z position. One can observe the typical curves for the CG model bilayers with density profiles for all available beads of the DPPC molecule. For clarity and because the SPICA beads do not represent the same number of atoms, all histograms were weighted with a volume proportional to the cube of the Lenard-Jones distance parameter of the corresponding bead interaction potential, thus giving to each bead a realistic excluded volume. Even though the slab size was smaller than the average bead diameter, this quantity can be interpreted in a similar manner as a packing fraction. Density of the tails beads (CnA) is taken considering that each lipid has two tails (a pair of $sn1$ and $sn2$ carbon CnA and CnB beads), n stands for the bead number in the lipid tail, and is multiplied by two. In the following analysis C1n is equal to CnA representation and accounts only for a bead in one tail of each lipid molecule. This nomenclature is somewhat artificial and was beneficial at the early stages of the project to distinguish different lipid species. The central peak shows that there is indeed an overlap of the last tail beads, leading to an idea of a possibility of some coupling between different leaflets. We assume the effect of this to be comparably small. The number density profiles for all studied lipids are available in the Appendix C. One can also observe the total lipid density profile and the tail group density profiles.

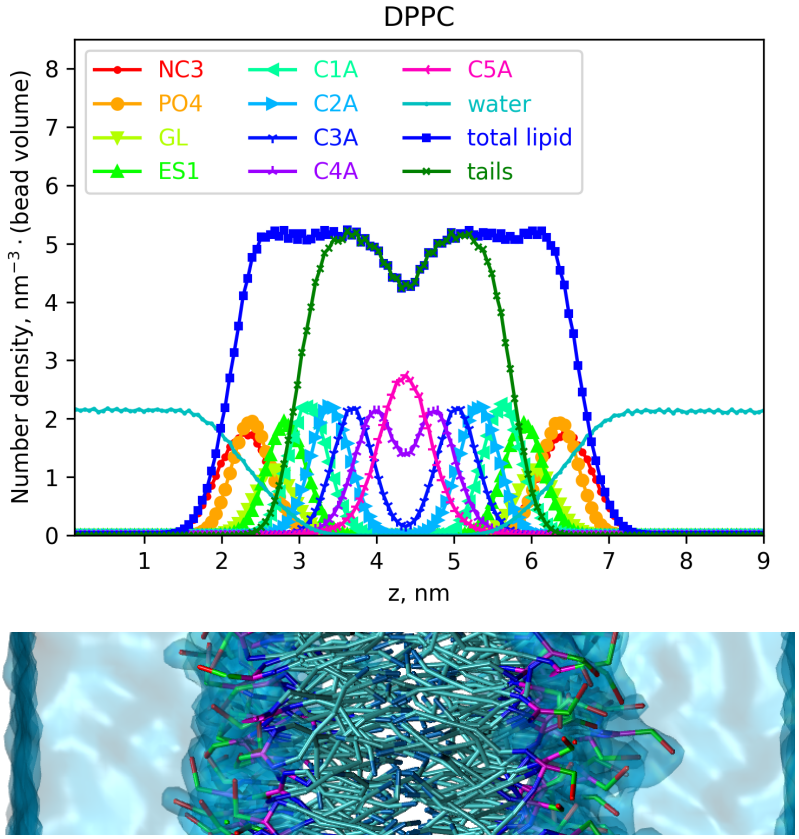


FIGURE 4.2: The number density profile along z axis for the pure DPPC bilayer of 512 lipid and 8192 water molecules. Snapshot of the DPPC pure bilayer is shown below the plot.

It can be seen that the permeability of water is quite high in the case of DPPC bilayer as some amount of the water beads are still present up until the ester bead. Density profiles can be used to estimate the thickness of studied system.

4.2 Reducing phospholipid molecules to a single pair of horizontal coordinates

The thermodynamic analysis approach in the current project is based on treating the lipids in a bilayer as if they were a simple two dimensional fluid. To do so it is necessary to assign to each lipid a representative (or reference) point $\mathbf{r} = (x, y)$. In supplement we disregard the vertical exploration of the lipid molecules along the normal direction z. This motion, sometimes called protrusion, is small compared to the thickness of a leaflet, because hydration forces are strong enough to prevent the phospholipids penetration in the solvent. The fact that lipids have a very low water solubility is also favorable for this situation. The overall projection procedure can be seen as an "extreme coarse-graining" approach. Even though this method is attractive from the computational perspective and

analytical simpleness, we loose the information on the lipid orientation (director vector **M** as defined in Chapter 2) inside the membrane.

Further analysis of the lipid systems shows us that separation for mono and bilayer description is not the only possibility to explore. Coarse-graining simplifies the structure of the studied molecules and so a few real atoms (for example, here, 3 carbon atoms for lipid tails and 3 water molecules for water) are mapped to one bead of the CG field. Thus, during the analysis of such molecules, we can choose to take different reference points to represent the whole structure. This question might not arise naturally for some readers but in the case of molecules with a complex structure like lipids, this choice might play a big role in such thermodynamic properties as structure factor and compressibility.

Let us introduce here possibilities we have explored during the current project:

$$\begin{aligned} &\text{real COM} && \mathbf{r} = \mathbf{R}_{\text{COM}} = \sum^{\text{beads}} \frac{\mathbf{r}_i m_i}{m_i} \\ &\text{pseudo COM} && \mathbf{r} = \mathbf{R}_{\text{pCOM}} = \sum \frac{\mathbf{r}_i m_i}{m_i} \overset{m_i=1}{=} \frac{\sum^{\text{beads}} \mathbf{r}_i}{N_{\text{bead}}} \\ &\text{bead} && \mathbf{r} = \mathbf{R}_{\text{bead}} \end{aligned} \tag{4.1}$$

where real COM stands for the real center of mass (COM) of the lipid molecule, pseudo COM - COM if all beads in a molecule had equal mass $m_i = 1$, bead stands for the coordinates of some specific beads of the lipid (*i.e.* GL, NC, EST), $\mathbf{r}_i$ is the ith bead coordinate, N_{beads} is the total number of the representative particles. The summation in eq. 4.1 is done over beads in one lipid molecule. To calculate real and pseudo COM values we used a Python script that identified all bead masses and corresponding positions, and then found the COM evolution of each molecule in the system. Lipid pseudo center of masses are easier to compute in SPICA, as realistic bead masses are included into this model. Preserving realistic masses is important as far as kinetic properties are concerned, but should be of a secondary importance when considering only equilibrium fluctuations. Classical (*i.e.* non quantum) thermodynamic properties cannot depend explicitly on the beads masses. After thoughtful investigation we concluded that there is little to gain to keep the realistic masses in our lipid representations.

Both COM and pseudo COM representations are "immaterial" in the sense that they are related to lipid beads positions but are not exactly present in the simulation system as a defined objects. It is very much possible for two distinct lipid molecules to have their center of mass (or pCOM) located at the same position, especially given that the information on the vertical component is not present anymore. This means that the effective interaction between center of masses can be considered "soft" (a feature already known from polymer physics). There is no rules forbidding the overlap between two different

molecules' COMs. On the contrary, choosing one bead such as GL (glycerol) as a representative of the whole molecule preserves the hard core short-range repulsion between neighboring lipids, even though the repulsive character is weakened as compared to the initial Lennard-Jones interactions. This justifies why an effective 2-body potential (if relevant) acting on the projected lipid positions can be different and can give rise to the different structure factors.

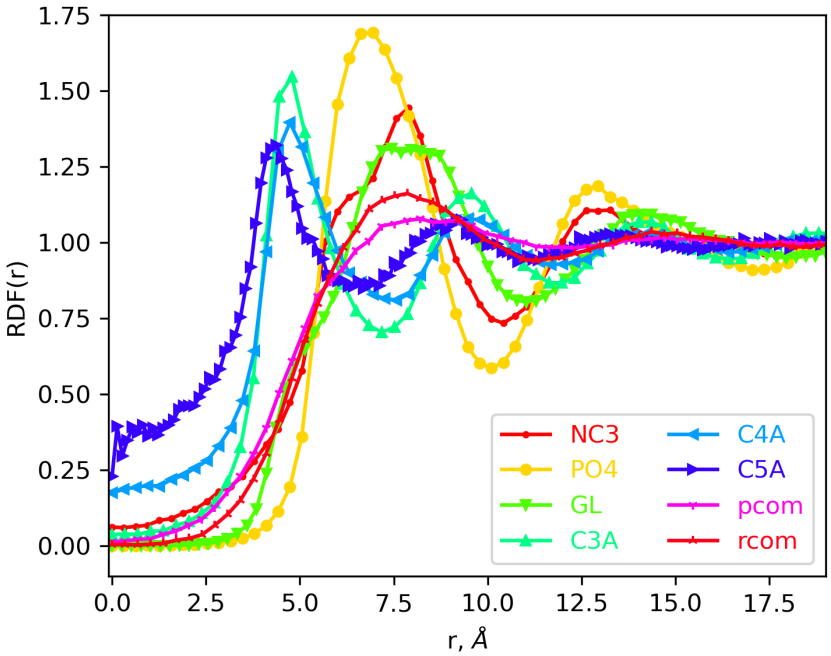
4.3 The real space radial distribution function

Let us begin our exploration by the real space property - **radial distribution function** (RDF). As was defined previously, RDF shows the probability to find specific atoms on the distance r from the particle in question. In Fig. 4.3 one can see pair distribution function calculated for both inter and intra leaflet interactions. As can be seen, the inter-leaflet interaction is very low compared to the intra-leaflet RDF, which proves that the assumption about replicated monolayers is fair. As is expected, the interaction between the last beads in the carbon tail is still significant so in further calculations this was taken into account and such values were used only for comparison reasons. In Fig. 4.3 peaks correspond to expected the values as stated by the SPICA FF for inter atom distances in the LJ parameters. Here we can also see how important is to choose the right representation of lipid for calculations as the resulting curves differ a lot. We will try to give an analytical explanation of this effect later throughout this chapter.

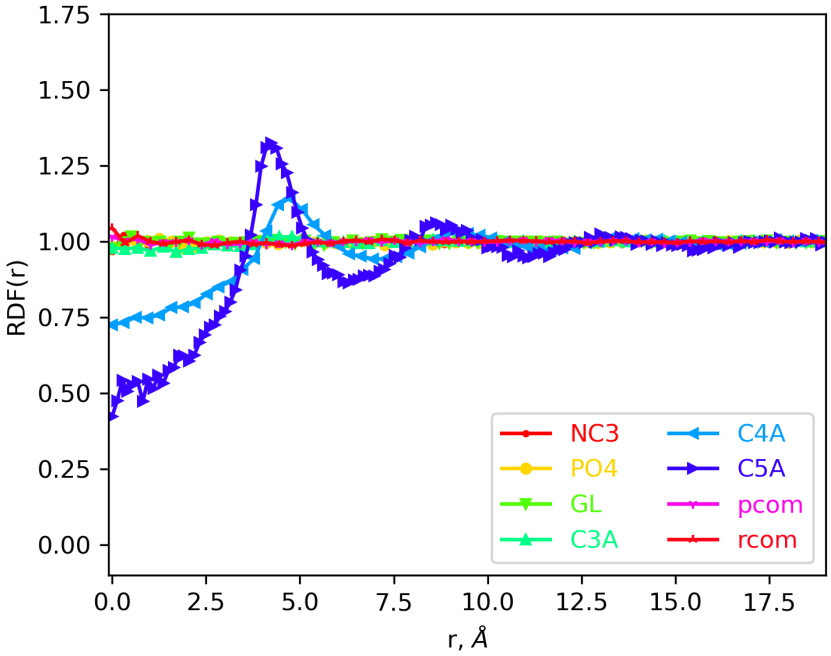
Let us discuss Fig. 4.3 (A) in more detail. It is important to remind here the two characteristic length scales for the studied system: the average distance r_0 between two DPPC molecules and a typical SPICA bead LJ σ parameter (see Appendix B). The first one can be defined as follows:

$$r_0 = a_0^{0.5} \tag{4.2}$$

where a_0 is the equilibrium area per lipid. In the case of DPPC $r_0 = (59.4\ \text{Å}^2)^{0.5} \approx 7.7$ Å. On the other hand, a bead diameter σ set into the LJ potential ranges from 3.8 to 5.8 Å. As can be seen, both are quite close values. The real space first correlation peak is quite visible for the NC, PH (PO4 on the graph) and GL beads. One can see that the best representative peak location is observed in GL, NC and COM cases, when the peak is located exactly around the r_0 value. The strongest correlations, *i.e.* the strongest deviation from 1, are the one of PH (phosphate). Then come NC (choline) and GL (glycerol). PH and NC are both charged which certainly favor a stronger structure and could be the reason behind a PH shift and a shape of the NC peak. When NC sits on the top of PH, this creates a vertical dipole which should interact repulsively with neighboring similar dipoles. One can see that the peaks are ordered in the same manner as they are positioned in the lipid



(A) Intra leaflet RDF.



(B) Inter leaflet RDF.

FIGURE 4.3: The radial distribution function for a DPPC bilayer system calculated inside one leaflet (intra) and between different leaflets (inter) for different beads and the center of mass of the lipid as a function of the distance r.

molecule: NC, PH, GL, with a small shift in PH bead RDF. Both real COM and pseudo COM show less correlated peaks at the r_0 location. This can point to an idea that they are subject to a relatively soft effective mutual interactions. Finally, the tail beads C3A, C4A and C5A show strong correlation peak at a shorter distance, fluctuating around their $\sigma = 4.5$ Å value. That is both due to the fact that there are two hydrocarbon chains present in each molecule and that the acyl chains are known to adopt a loose hexagonal order in the fluid phase [74].

In Fig. 4.3 (B) one can see the inter leafler RDF for the DPPC pure bilayer. It can be observed that the cross leaflet correlations are almost absent for all studied beads except for the C4A and C5A beads. This can indicate that the lipid beads pertaining to the opposite leaflets can be found at the same horizontally projected place. On the other hand, this is not always true for the last tail beads of the molecule. One can try to explain the latter by the fact that the acyl chain final beads of opposing leaflets mix uniformly and each of them tend to "forget" the original lipid to which it belongs. Thus, it can be concluded that the assumption about the absence of the coupling between different leaflets can be hold fair in most of the cases.

4.4 Properties of the 2d structure factors

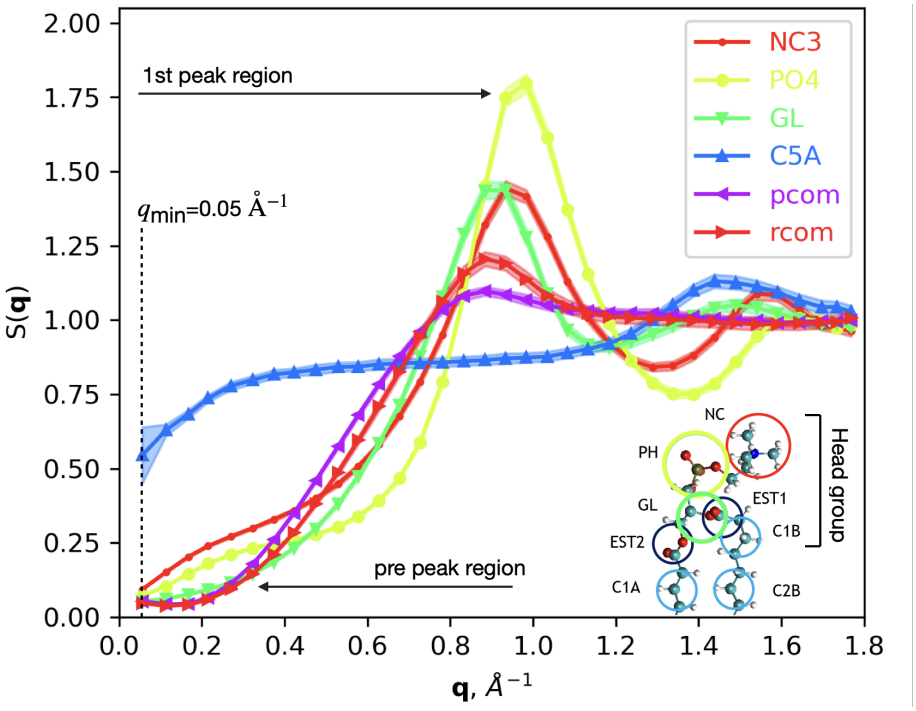


FIGURE 4.4: The structure factor of the whole DPPC pure bilayer calculated with respect to different beads, pseudo and real center of mass. The confidence interval is shown by transparent color around the main lines.

Next, we dive into the reciprocal space description of the system. In the case of the lipid bilayer, the lowest $q = \frac{2\pi}{L_x}$ wave-vector value available would depend highly on the simulation box size. In Appendix C in Fig. C.3 and Fig. C.4 one can see the **structure factor** $S(\mathbf{q})$ as was defined before for different representations of the lipid molecule for available model lipids: DPPC, DLiPC, DOPC, DOPE, SMM and PSM. We can directly see the density fluctuations of the system in the larger $\mathbf{r}$ scale (smaller $\mathbf{q}$ scale). In Fig. 4.4 one can see the structure factor of DPPC lipid with confidence intervals and peak regions as we define them in this work: the first peak region and a pre peak region.

Let us discuss in more detail the two regions one can see in Fig. 4.4 (see Chapter 2 eq. 2.4 and eq. 2.70). The **first peak region** governs peaks that are directly related to the peaks we see in the RDF. For example, peak for the phosphate group PH is approximately at $\mathbf{q} \approx 1.0$ Å^{-1} which corresponds to $\mathbf{r} \approx 6.28$ Å that can be seen in Fig. 4.3. As well as that, in both cases one can see that the peak for PH is the most pronounced one among all beads. The choline (NC) and glycerol (GL) peaks are visibly smoother and of similar amplitude. The height of these peaks can be a good estimator of the degree of structuring in the system. The higher and sharper peak means that system is more restricted than the one that has low and broad peaks in $S(\mathbf{q})$. One can see that the real COM display a mild peak while the pseudo COM has almost no peak present. The latter is in agreement with the results obtained using the RDF. The tail bead C5 shows different behavior compared to the other presented cases. The peak of this carbon bead is shifted in $\mathbf{q}$ to approximately 1.4 Å^{-1} that correspond to $\mathbf{r} \approx 4.48$ Å. That is in consistence with what was observed for the RDF. It is also interesting to note here that the ration between the PH and C5 peak wave vectors is close to $\sqrt{2} = 1.41$, which is to be expected as a consequence of the fact that there are two carbon chains in the tail group compared to only one head group and thus, two carbon beads are seen by the density fluctuations.

One can attempt to interpret the structure factors around the first peak region with the use of a form factor. As was proposed earlier, $S(\mathbf{q}) - 1 = (S_{\text{best}}(\mathbf{q}) - 1)|F(\mathbf{q})|^2$ and it is natural to assume that the PH bead seems to be the closest to the unknown "best" representation of the lipid molecule, with $|F_{\text{PH}}(\mathbf{q})| \simeq 1$. The two beads directly attached to PH (GL and NC) should then be subject to a small form factor $|F_{\text{GL}}(\mathbf{q})|^2 \simeq |F_{\text{NC}}(\mathbf{q})|^2$. At the same time, the latter form factor should be monotonously decreasing proportional to the inverse of the horizontal distance between the consecutive beads in SPICA representation (equal to a fraction of the bonding distance between the consecutive beads).

The rCOM and pCOM form factors seem to be much smaller than the ones discussed before. This reflects certainly the fact that the center of mass position of the lipids falls under the influence of all lipid beads, including the ones in the two carbon chains. The decay distance of the corresponding $|F(\mathbf{q})|^2$ must then be of the order of the inverse of the

lipid horizontal characteristic size, which is also the inter lipid distance $r_0 = 7.7$ Å defined previously. pCOM is the smoothest of all structure factors around 1 Å^{-1}. No attempt was made to compute numerically $S_{\text{best}}(\mathbf{q})$ and the $F(\mathbf{q})$ factors. The fact that the peaks position and intensity are ordered (pCOM, rCOM, GL$\simeq$NC, PH) as the inverse of their respective height is however going in the right direction with the above arguments.

At this point it is interesting to mention that based only on the analysis of the first peak region the PH bead seems to be the best reference point of the lipid molecule. At the same time, the pseudo COM turns out to have the smoothest features of all leading to the idea that it lacks any specific information about the internal structure of the lipid. rCOM certainly favors the head group beads which are heavier, and lies closer to the PH on average. Unweighted bead average position (pCOM) or weighted bead positions (rCOM) both seem to lead to a broader distribution of positions with respect to the PH, NC or GL beads. Both GL and NC could be good representative beads with the peak position in the expected r_0 position, although the height and shape of these peaks are definitely not as pronounced as for PH. Final tail bead C5 is certainly the worst possible locator of the lipid due to its large wandering motion in the fluid state.

On the other hand, the structure factor should normally behave as a $\mathbf{q}^2$ near $\mathbf{q} \to 0$, while here we see a **pre peak** formation in PH, NC, and C5 beads (as a small reminder, C5 corresponds to the 5th carbon bead in DPPC). The description of $S(\mathbf{q})$ in the low $\mathbf{q}$ region shows notable differences from what was said for the first peak region. The curves do not superimpose but they seem to point towards a similar extrapolated value $S(0)$. The only exception is C5, which definitely does not behave in a similar way as the other positions, most likely due to the same reasons already mentioned before. On the contrary to single beads, their collective representatives pCOM and rCOM display a minimum in $S(\mathbf{q})$ at finite $\mathbf{q}$. One can see that the GL bead has the smoothest curve near the zero limit wave vector and approaches minimal $\mathbf{q}$ with constant $S(\mathbf{q})$ value. Even though the NC and PH beads indeed approach the same $S(\mathbf{q} \to 0)$ value, they are exposed to the appearance of the pre peak. Opposite to the COMs behavior, peaks in the latter beads have positive curvature. It is interesting, that both PH and NC are charged particles and they seem to form some structures in the long-range $\mathbf{r}$ region. As well as that, PH in the pre peak region has a lower peak than NC bead. One can conclude now that the PH might not be the best representative bead in the low-$\mathbf{q}$ limit, but GL might be. So the question which representation is the best as a one reference point in the molecule still remains unanswered and depends on the characteristics one might be interested in.

One can see that the value of $S(\mathbf{q})$ has a strong dependence on the bead type that was used in the calculation in all studied cases (see Appendix C Fig. C.3 and Fig. C.4). As well as that, saturated lipids and sphingolipids seem to form more structured configurations

with less fluidity in the system. The latter can be assumed to be based on the height and shape difference of the corresponding peaks for each lipid. Also, it is worth mentioning here, that values for DPPC and DliPC lipids were obtained as an average of 10 μs runs while other systems were only run for 100 ns. We can see that the statistics is also better in the case of saturated lipids regardless of the simulation time. It was found that the pre peak is present in all available molecules and, thus, is a common feature for all studied systems. We attempt to explain the observed pre peak behavior with the use of the so-called Helfrich undulations [23] and the tilt of the lipid molecule.

4.4.1 Fit and extrapolation to zero of the structure factors

Our interpretation of the pre peak region is that it reports on the effects of tilt and undulations that act mainly in the low-$\mathbf{q}$ region. The PH and NC inflexion points are located around $\mathbf{q}_{\mathrm{PH}} = 0.42$ Å^{-1} and $\mathbf{q}_{\mathrm{NC}} = 0.35$ Å^{-1} or equivalently a length $\mathbf{r} = 2\pi/\mathbf{q} \approx 15 - 18$ Å. As was shown in the theory section, undulations are expected to modify $S(0)$ if the lipid locator departs significantly from the neutral surface. One can test this hypothesis by performing a fit of each of the structure factors by a parabolic $S(0) + C\mathbf{q}^2$ in the range of $\mathbf{q}_{\mathrm{min}} \ldots 0.2$ Å^{-1}. The actual value of $\mathbf{q}_{\mathrm{min}}$ varied by the system as it is highly dependent on the membrane lateral size, the typical value was around 0.05 Å^{-1}. A simple Python algorithm was developed for this purpose using NumPy and SciPy packages. This algorithm identified data points that should be included in the predefined $\mathbf{q}$ range and performed approximation and extrapolation to the $S(0)$.

In Appendix C in Fig. C.3 one can see the least squares approximation of obtained data points after the bootstrap algorithm. Even though the final curves are not as precise as one would want them to be, we consider our results consistent enough to proceed with further analysis. To obtain the observables that depend on the zero-value of the structure factor we used both zero-values from the extrapolation procedure and the raw data, extrapolated afterwords with a similar procedure as here. Then we compared the results and found small to no difference in the resulting values.

The resulting $S(0)$ values are reported in Fig. 4.5 for DPPC and other model lipids used in the current project. The curve shows a clear trend when $S(0)$ is plotted as a function of the reference bead, from the head group to the tail chain. All glycerophospholipid molecules display a parabolic dependence of the extrapolated $S(0)$ as a function of the "expected vertical position" of the associated beads. It is tempting to recognize here a law

$$S(0) = \rho_0 \frac{k_B T}{K_A} + \rho_0 \mathcal{D}^2 \frac{k_B T}{\kappa} \tag{4.3}$$

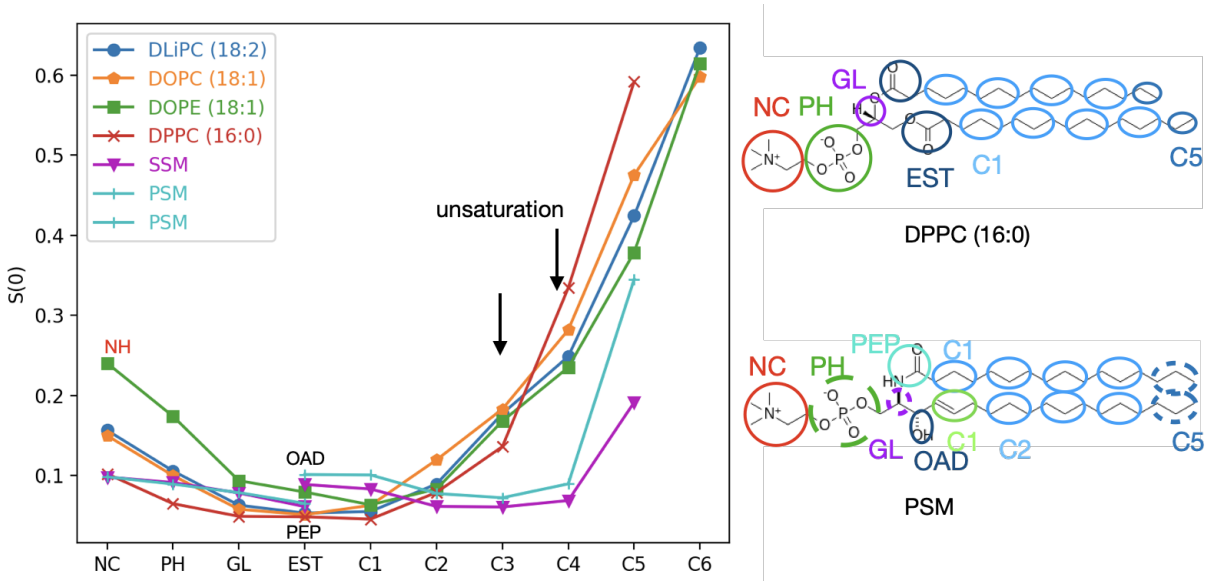


FIGURE 4.5: Zero limit structure factor $S(0)$ dependence on the bead type for model lipids in upper leaflet. On the right chemical structures and corresponding mapping to the force field of DPPC and PSM lipids are shown as a little reminder to the reader.

For phospholipids with the choline headgroups PC (DliPC, DPPC, DOPC) one can see a flat minimum around the glycerol (GL), ester (EST) and first carbon (C1) beads, showing that these points all predict a similar $S(0)$ value. One can connect those 3 beads to a vanishing or comparably small $\mathcal{D}$ value. All three beads should make decent predictions as far as in plane density fluctuation modes are concerned. A striking conclusion is also that PH, which was the best locator of lipid molecules in the first peak region, is also subject to some undulation corrections. Thus, the best decision would be to use the average over the GL, EST and C1 beads in case of calculations for DPPC, DLiPC and DOPC.

On the other hand, for DOPE lipid the situation is a bit different. Even though the overall trend is very similar for all studied phospholipids, the ethanolamine (PE) headgroup of the DOPE seems to be more prone to the out of plane fluctuations around the neutral surface. In the case of the PE lipid the minimum is clearly located next to the C1 bead, as if the neutral surface was below the one found in the PC lipid case. This can be explained if one looks at the masses of the PE and PC head group beads (see Appendix B). Since the NH bead is almost two times lighter than the NC bead, it can be assumed that the center of mass of such molecule is shifted to the carbon tail of the lipid. Thus, the best estimate of $S(0)$ for DOPE should be expected from the C1 or neighboring beads.

The distance parameters $\mathcal{D}_{\mathrm{PH}}$ and $\mathcal{D}_{\mathrm{NC}}$ seem to be significant for all studied phospholipids, with $\mathcal{D}_{\mathrm{NC}} > \mathcal{D}_{\mathrm{PH}}$. Generally speaking, this trend should continue up to the neutral surface, where $\mathcal{D} = 0$. Symmetrically, the $\mathcal{D}$ parameter increases as one goes down the lipid chain groups $\mathcal{D}_{\mathrm{C1}} < \mathcal{D}_{\mathrm{C2}} < \mathcal{D}_{\mathrm{C3}} < \mathcal{D}_{\mathrm{C4}} < \mathcal{D}_{\mathrm{C5}} < \mathcal{D}_{\mathrm{C6}}$. It can be seen that the tail

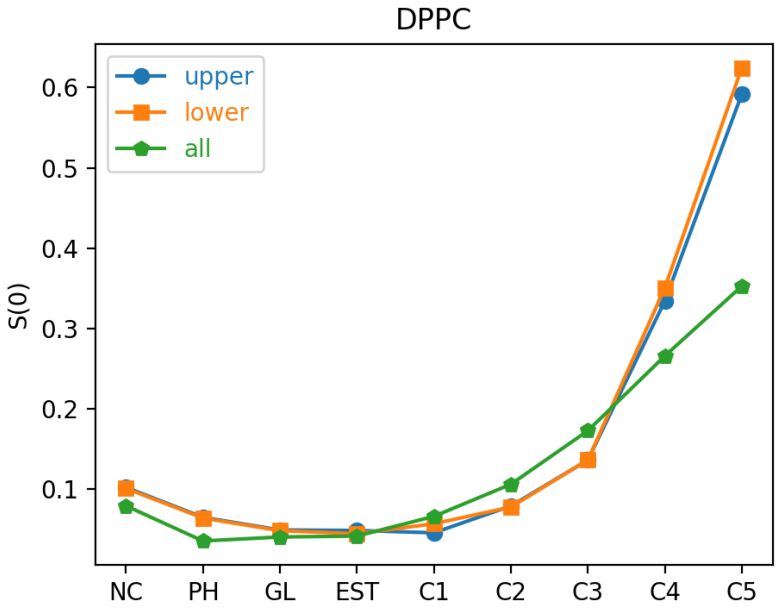


FIGURE 4.6: Zero limit structure factor $S(0)$ dependence on the bead type for model DPPC lipid in upper and lower leaflets, and in the full bilayer.

beads are strongly affected by the undulations as could be expected. We conclude that for PC molecules, the neutral surface should pass approximately across the EST group position, while for the PE lipid the neutral surface is shifted to the tail group due to the effect of the lighter head group.

On the other hand, sphingolipid molecules PSM and SSM show different behavior from the phospholipids. They have more pronounced, sharper peaks (see Appendix C Fig. C.4) which is known to be due to their higher rigidity in the tails. The weak dependence of $S(0)$ in the nature of the bead suggests a reduced influence of tilt and undulations probably due to a larger value of the bending modulus κ. It can be concluded, that in the case of PSM and SSM one can use the average over the structure factor calculated for all beads (except the first and last beads). The latter can be understood using similar arguments as before as some tilt is still present in the system and it can influence the most distant beads from the neutral surface.

In Fig. 4.6 one can see the dependence of $S(0)$ on the bead type for the DPPC lipid calculated for the upper, lower leaflets and the whole bilayer. Values for the separate monolayers are consistent and both show higher bead dependence than the whole bilayer. It can be stressed here that the undulation effects seem to cancel out at some point in the case of the membrane calculations. Even though that might be the case, the influence of tilt and undulations is still present, although less strongly than for the separate bilayer leaflets.

4.5 Comparison of alternative measures of compressibility

TABLE 4.1: Structural properties of pure bilayers: area per lipid (APL, nm^2), thickness (z, nm), and stretching modulus (K_A, mN/m).

	APL	z	K_A^{fluct}	K_A^{str}			K_A^{eqos}	K_A^{exp}
				upper×2	lower×2	bilayer		
DPPC[1]	0.594	4.15	258.6±64.4	248.4±8.4	226.0±11.2	248.6±14.4	275.1±10.4	231±20
DLiPC[1]	0.679	3.97	281.2±57.2	216.6±5.5	215.0±5.1	279.8±9.1	319.7±9.6	247±21
DOPC[1]	0.670	4.06	305.1±118.9	213.5±4.1	222.7±4.9	273.6±9.2	-	265±18
DOPE[2]	0.629	4.07	370.0±68.4	197.3±7.6	199.6±7.5	271.8±11.0	-	-
PSM[3]	0.518	4.29	411.9±164.5	227.3±11.4	220.5±10.0	263.1±13.1	-	-
SSM[3]	0.523	4.49	297.9±152.9	221.5±8.9	225.5±8.8	242.5±12.0	-	-

[fluct] volume fluctuations, [str] zero limit structure factor, [eqos] equation of state, [exp] experiment data from literature [9, 40, 64]

[1] $T = 298$ K, average over GL, EST and C1 beads

[2] $T = 298$ K, average over EST, C1 and C2 beads

[3] $T = 318$ K, average over PH, GL, PEP, OAD, C1, C2, C3, C4 beads

Table 4.1 presents the area per lipid (APL), membrane thickness z and the area elasticity coefficient K_A of a number of lipid molecules available in the SPICA world. Area per lipid is defined as the equilibrium box area for a vanishing tension $\sigma = 0$ divided by the number of lipids per monolayer. Membrane thickness corresponds to an average distance between the upper and lower phosphate (PH) beads performed over simulation time and lipid molecules.

We have calculated the compressibility of pure DPPC and DLiPC bilayers with the three methods mentioned in Chapter 2. First is the fluctuation approach which corresponds to an equilibrium box fluctuation method using a barostat with an anisotropic coupling in the z and x, y directions. The barostat provided by LAMMPS enforces a constant pressure $P_x = P_y = P_z = 1$ atm, zero tension $\sigma = 0$ mN/m and user-defined temperature T. After the equilibration of the box size to some fluctuation value, one can extract K_A^{fluct} from the area fluctuations data. In this work we have used second half of the NPT run for this purpose. Second method is the equation of state approach. This method consists in enforcing a finite tension $\sigma \neq 0$ and obtaining the sampled average area $\langle A \rangle$. One can compute directly the derivative $\mathrm{d}\langle A \rangle / \mathrm{d}\sigma$ to get K_A^{eqos}. In this project, an average over five independent configurations under $P_x = P_y = [-6, -5 \ldots 5, 6]$ atm was used. The third method involves the structure factor zero limit. The extrapolated $S(\mathbf{q} \to 0)$ value of the pure system can be used to obtain the area stretching modulus $K_A^{\text{str}} = \rho k_B T / S(0)$.

It can be seen that the compressibility obtained by the area fluctuation calculation has higher error values but is roughly consistent with result from the equation of state method in both cases. On the other hand, if we use the zero-structure factor to get the stretching modulus of the studied systems, there is a significant difference between the final values depending on the reference subsystem choice. By the reference subsystem here we mean

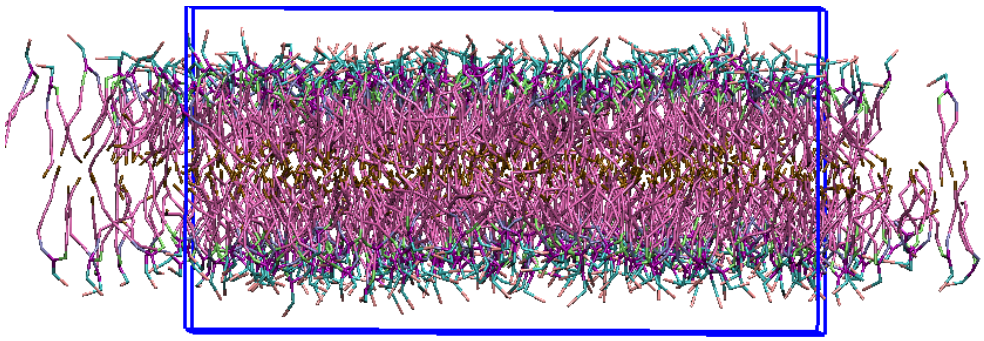
the possibility to perform the calculation on the monolayers or the whole bilayer all-together (see Fig. 4.1).

From the Tab. 4.1 one can see that the agreement between K_A^{fluct} and K_A^{str} seems better with the whole bilayer than with the separate leaflets. It is important to mention here that the calculation of the compressibility from the fluctuation of the area was performed for the whole bilayer. In addition, the separate leaflets appear to be softer than the whole system together, which means that the low $\mathbf{q}$ compression modes have higher amplitude. In other words, one can say that the structure factor is larger in this case. It can be assumed that the undulations affect more the low wave length regime of separate monolayers than the bilayer as a whole. As was mentioned in Chapter 2, one expects the tilt and undulation displacements to act in an opposite manner on the two bilayer leaflets. For instance, if one leaflet is curved with a local "splay" expansion of the director, the opposite leaflet is then curved in the opposite direction with a compression of the director. The difference between leaflets and bilayer K_A^{str} values certainly reflects the effect of the non vanishing $\mathcal{D}$ coupling to the undulation modes.

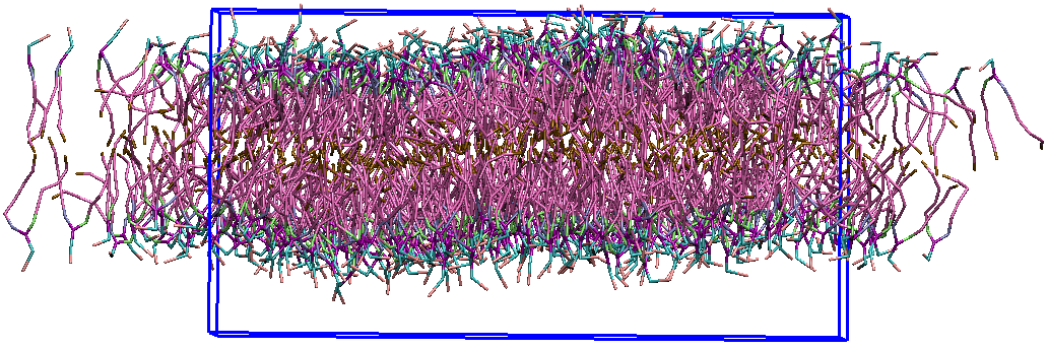
The other studied model lipids show similar trends in the stretching modulus calculations. In Tab. 4.1 we introduced the results for phospholipids and sphingolipids. It can be seen that the error for the fluctuation method is roughly 20-30% for phospholipids and almost 50% for sphingomyelin. This is unfortunate but can be understood as a consequence of the fact that both these lipids require higher temperature to remain in the fluid phase and so might still be subject to some residual structuring. It is interesting however that the structure factor approach does not show the same feature. Overall we conclude that there is a rough consistency between the prediction of the fluctuation method and the structure factor method.

4.6 Linear response in density

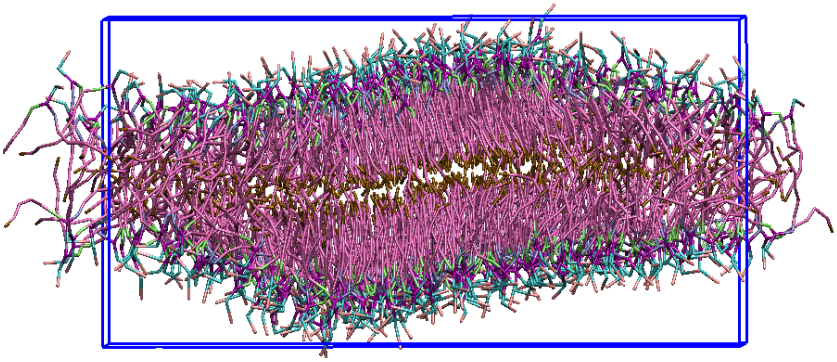
Let us present now the results of an alternative way to calculate the structure factor. As was introduced in the Chapter 2, we can apply an external field (here, periodic cosine) to the system and obtain the density fluctuations as a response to the external potential. The tricky part here was to find the appropriate force field strength so the response in linear but still detectable. In Fig. 4.7 one can see an example of the external potential of different amplitude influence on the system. It can be seen how a strong field deforms the bilayer structure and the response can even be seen by visual inspection (see Fig. 4.7 (C)). The more dense region is observed for the high amplitude which creates clusters of higher density and thus, changes overall shape of the bilayer. In this case we are no longer in the linear response regime and so forces in the magnitude of $U_0 = 0.5$ kcal/(mol·Å) are



(A) $U_0 = 0.00$ kcal/(mol·Å).



(B) $U_0 = 0.05$ kcal/(mol·Å).



(C) $U_0 = 5.00$ kcal/(mol·Å).

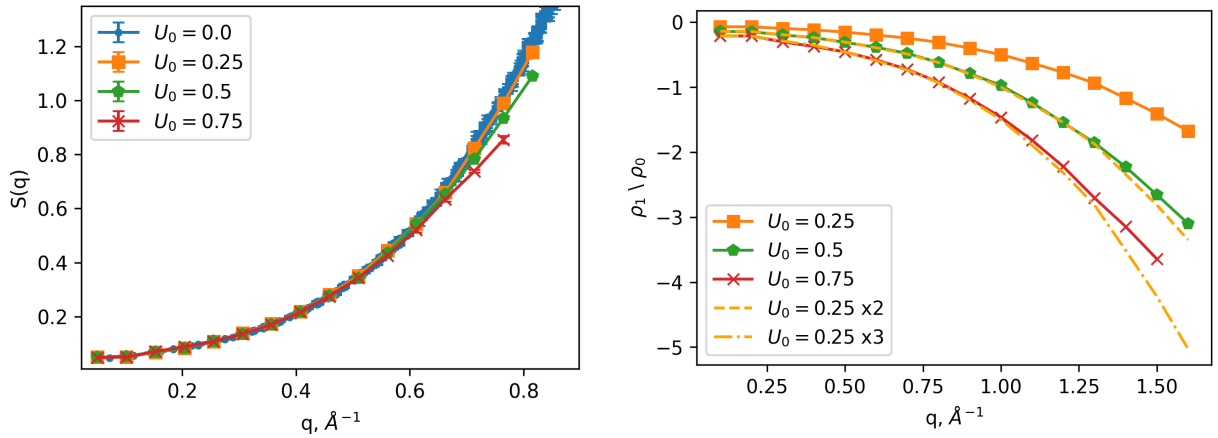
FIGURE 4.7: Snapshots of the DPPC bilayer showing how the different external fields impact the system. Water is not shown.

more suitable for our purpose. As well as that, the force was normalized by the number of lipids and/or number of beads. One can see that Fig. 4.7 (A-B) are almost identical meaning that the external potential does not disturb the systems' structure. We present the comparison between three U_0 tested: 0.25, 0.5, 0.75 kcal/(mol·Å).

In Fig. 4.8 one can see structure factor with and without the external field. It can be seen that smaller force fields give better agreement with the unbiased data. On the other hand, smaller amplitude requires longer simulation times to equilibrate and gain enough statistics as it is almost equal to zero-force situation. One can see from Fig. 4.8 (B) that relative density ρ_1/ρ_0 is consistent for three cases up to $\mathbf{q} \approx 0.9$ Å^{-1}. However, if we look

on the $S(\mathbf{q})$ for such systems, we can already see some inconsistency even in the small-$\mathbf{q}$ region. To be more precise, the size of this system was $L_x = 12.3$ nm so $q_x = 0.7$ Å^{-1} would mean we are looking on density fluctuations at $r_x \approx 0.9$ nm which is almost a half of the cut-off of the LJ potential between lipid beads. Thus, it is possible to assume that at smaller distances interactions between particles are not strong enough anymore and at some point the force plays bigger role to rearrange the system structure. Finally, we can conclude that the linear response method is efficient in the case of small external forces and wave-vectors.

It is satisfactory to see that the structure factor predicted by the density fluctuation approach is fully consistent with the one derived from applying a sine potential to the lipid and reading out the magnitude of the resulting sine density wave. In addition, one is free to choose any reference bead, or center of mass position to compute the equilibrium structure factor. It is important to keep consistency between the way a potential is applied and the analysis is performed: the force has to be deployed exactly on the same bead or reference point as when computing the biased equilibrium density wave amplitude.



(A) Structure factor.

(B) Relative density.

FIGURE 4.8: (A) Structure factor as a function of wave-vector $\mathbf{q}$ calculated in the presence of an external field. (B) Relative density obtained in the external potential. In both cases, U_0 is the amplitude of the force field, $\times 2$, $\times 3$ means that the results for $U_0 = 0.25$ kcal/(mol·Å) were multiplied twice or three times correspondingly to check the linearity of the response. Results shown for the GL bead.

The possibility of applying a tailored force on each beads in a system was one of the main reasons for choosing LAMMPS software and SPICA force field. The possibility to apply user-defined forces to molecules or particle during the simulation is not always present in other MD packages. For example, the biasing methods of GROMACS are more restricted. The initial idea here was to explore the linear response route to compute the low $\mathbf{q}$ values of $S(\mathbf{q})$, the only ones required for the thermodynamic parameters of interest.

It usually is more efficient and statistically more significant to compute an average $\langle \hat{n}_{\mathbf{q}} \rangle$ than a variance $\langle \hat{n}_{\mathbf{q}} \hat{n}_{-\mathbf{q}} \rangle$. However, it turned out that the biasing method required more testing (finding the right amplitude to stay in the linear regime, the right reference point for biasing) and that the signal to noise ratio was not so favorable in the low bias amplitude region. As well as that, the required simulation time was found to be longer than expected initially. Regardless, further experimentation with this method was carried out in the case of phospholipid mixtures and is discussed in the following chapter.

Summary

Let us conclude the study of the pure lipid bilayers. In this chapter we have shown the results for available SPICA model lipids: DPPC, DLiPC, DOPC, DOPE, SMM and PSM. The DPPC pure lipid bilayer was used to establish the simulation protocol and was overall very instructive. Firstly, we discussed the different ways of reducing a lipid molecule to a single point and found that the extreme coarse-graining to the 2d fluid is fair in the thermodynamic limit. The examination of the real space radial distribution function shows that opposite leaflets can be considered spatially uncorrelated with an exception for the last hydrocarbon tail beads. Based on the obtained RDFs it was possible to establish a hierarchy of beads depending on the structure of their first peak. It was found that peaks are spread by position and intensity as follows: PH, NC, GL, rCOM, pCOM. Second step was to explore the 2d structure factor of the studied systems. Two different wave vector regimes were discovered for the pure bilayer system: the first peak and the pre peak regions. The first region is the one of the main correlation peak and reflects the hierarchy of the lipid representations mentioned for the RDF. At this point it could be assumed that the PH bead is the best locator of the lipid molecule as far as local lipid order and packing are concerned. Overall, the first peak features seem to be reasonably rationalized by a the existence of a Debye-Waller or form factor. On the other hand, the low $\mathbf{q}$ or the pre peak region follows completely different rules. The latter confirms the picture of undulation perturbations of the density modes discussed in the Chapter 2. As was expected, different lipid representations are not equally affected by the undulations, and our conclusion is that the best locators are rather GL, EST or C1 in the glycerophospholipids case. For PE lipid the situation was somewhat different with the best locators EST, C1 or C2 beads. Sphyngomyelin showed good persistence to the undulations and can be represented by any bead, except for the first and the last ones. Bunching both leaflets together into a single bilayer system also seems to give better results, possibly due to the fact that undulations tend to cancel out in this situation. Finally, linear response using biasing potential turns out to be consistent with the equilibrium calculations, but does not provide any significant computational gain.

CHAPTER 5

Binary lipid mixtures

THE main goal of this chapter is to study the mixing properties of various binary lipid mixtures. The approach relies on finding the value B of the regular solution binary term, that here is referred as a mixing parameter. This is done by resorting to the expressions 2.30, 2.31. Considering the fact that the Kirkwood-Buff theory only provides the derivative μ_{12} and that the theory of regular solutions is only an assumption, it is important to test the applicability of the present approach to binary mixtures of lipids. The latter is done by changing the molar fraction x of the binary system and checking whether the computed B depends on x. There is a limit however to the extent on which one can vary x. As a matter of fact, with x too small or too close to 1, the minority compound is diluted, thus, it is difficult to obtain a correct statistics for the density fluctuations $\langle \hat{n}_{\alpha,\mathbf{q}} \hat{n}_{\beta,-\mathbf{q}} \rangle$. Due to less particles present in the system, the diluted component has a poor statistics which influences the final consistency of the results. For these reasons, in this study concentrations of $x = 0.25$, 0.5 and 0.75 were considered for phospholipid mixtures and $x = 0.25$, 0.5 for mixtures with Cholesterol.

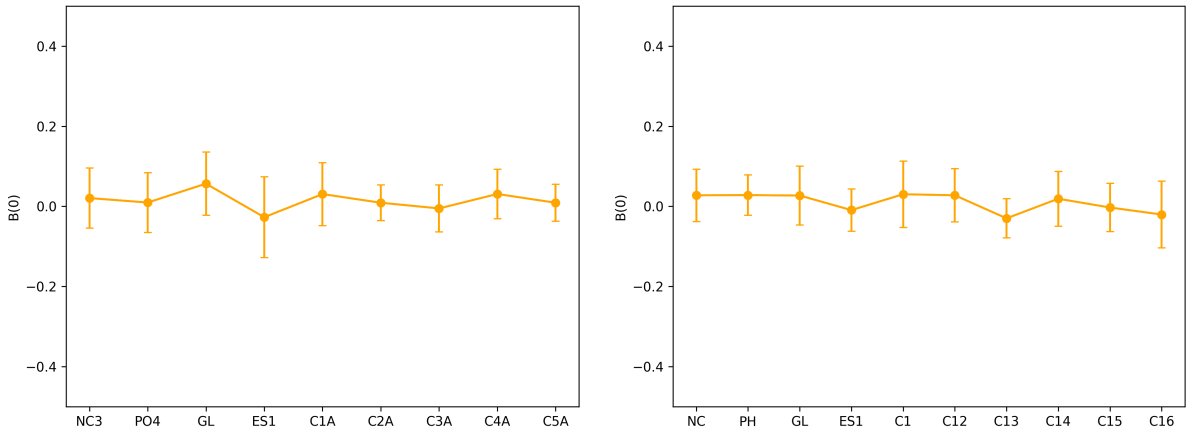
5.1 Binary mixtures of phospholipids

Let us discuss here the binary mixture of phospholipids. Mainly we will focus on the DPPC and DLiPC mixtures. We introduce the concept of "fake" mixtures and then proceed to real solution of two species of lipid molecules. Systems of 512 and 1024 lipids in total in 8192 water beads solution at $T = 298$ K and $P = 1$ atm were simulated for 10 μs

in a fixed volume ensemble to obtain static structure factors. Fluctuation methods were applied on simulation in fixed pressure ensemble to calculate area stretching modulus.

5.1.1 Fake mixtures

To test the validity of the methods proposed in this work, firstly we have decided to create a so-called "fake" mixture of lipids. This was done by simple labeling of same species as different type. With the developed algorithm it is possible to emulate and analyze any ratio of lipids "1" and lipids "2" from one original trajectory of a pure lipid bilayer.



(A) DPPC 50%:50%.

(B) DLiPC 50%:50%.

FIGURE 5.1: The mixing parameter as a function of the bead type for fake mixtures corresponding to two model lipids: DPPC and DLiPC.

An ideal binary mixture is one system where two molecules of different type can be exchanged without any free-energy cost. This is obvious in the case of two artificially labelled molecules with exactly the same shape and interaction parameters. What makes the fake mixture ideal in this favorable case is the identity $\tilde{h}_{11}(\mathbf{q}) = \tilde{h}_{22}(\mathbf{q}) = \tilde{h}_{12}(\mathbf{q})$. This ensures that the Δ factor from eq. 2.25 vanishes for all $\mathbf{q}$ and in particular for $\mathbf{q} = 0$. In other words, the KB integrals G_{11}, G_{22} and G_{12} (eq. 2.7) coincide. It entails that $B = 0$ irrespective of the compound concentration x.

As can be seen from Fig. 5.1, the mixing parameter calculated for such fake systems is zero on average for any bead type for both DPPC and DLiPC lipids. In the presented plot 1:1 systems were considered, meaning that in the DPPC pure bilayer half of the lipids from both leaflets were labeled as lipids "1" and the other half as lipids "2". The choice was made randomly but so that each leaflet has equal concentration of the corresponding molecules (the membrane is symmetric). The full bilayer is considered when computing the density autocorrelation functions. One can see from Fig. 5.1 that the prediction is well verified and can also estimate the precision of the predicted value of the mixing

parameter. A weak dependence of B on the choice of the representative bead points to the fact that the undulation effects cancel out in the presented calculation.

5.1.2 Real mixtures

Let us discuss here the results for the binary phospholipid mixtures. As was stated in previous chapters, we have performed analysis of DPPC/DOPC and DPPC/DLiPC lipid solutions. The molecules of the first pair proved to be too similar and so the resulting mixing properties were close to that of the ideal mixture. Thus, here we will present results for DPPC and DLiPC bilayers. We have studied two system sizes to check for the system size effects and found the results to be consistent and have no or low dependency on the size of the system. Systems of 512 and 1024 lipids in total were chosen as the most suitable for current research. We have developed bilayers of 1:3, 1:1 and 3:1 concentrations, meaning that for 1:3 we will have 128 DLiPC lipids and 384 DPPC lipids, making 512 lipids in total (or 256 DLiPC and 768 DPPC for systems of 1024 lipids).

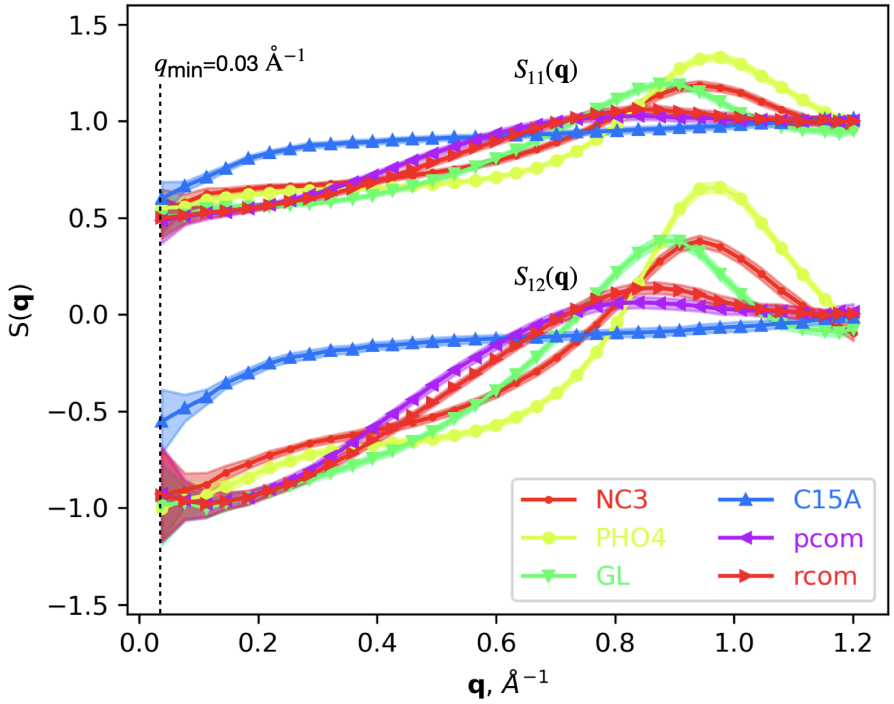
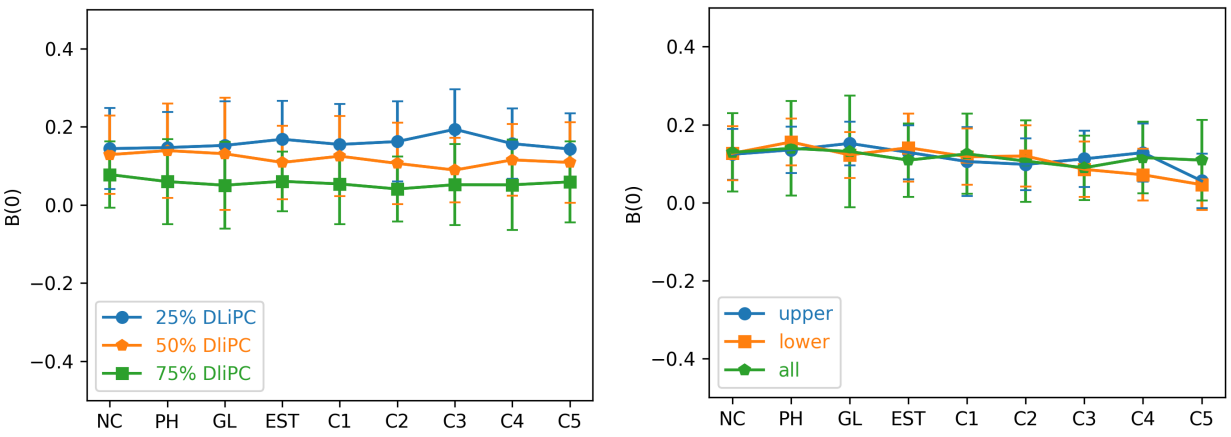


FIGURE 5.2: The structure factors of the 1:1 DPPC/DLiPC mixture as a function of wave-vector **q** for the upper leaflet of the bilayer. S_{11} here corresponds to DLiPC lipid and S_{12} to the cross-term of DLiPC and DPPC. Data labels correspond to the bead names as defined by the SPICA FF mapping.

In Fig. 5.2 one can see the structure factor for a 1:1 mixture of the DPPC and DLiPC calculated for different choices of representative bead in each lipid (see Appendix C Fig. C.6 for the full structure factor). Pairs for calculations were chosen in similar manner as for the pure bilayers. Let us trace the process for the NC bead. Both DLiPC and DPPC have

this bead in common and so we can use NC_{DLiPC} and NC_{DPPC} correspondingly to get S_{11} and S_{22} terms for the structure factor. For the cross-term S_{12} we use both NC_{DLiPC} and NC_{DPPC} in the definition of partial structure factor as was defined in Chapter 2. A similar pre peak behavior is observed and the reason behind it is assumed as in the previous section to be caused by the Helfrich undulations and lipid tilt. As can be expected, the confidence intervals close to the zero-limit of the wave-vector are comparably big and so one can see some uncertainty in the data points. One can also see how the pre-peak is more present in the DLiPC structure factor part (S_{11}). This might be due to the fact that the unsaturated lipids are more exposed to the kinks and molecule bending while DPPC is more rigid. However, we can still assume a consistency in the results with proper extrapolation to the zero values. Same analysis protocol was applied as for pure mixtures to get $S(\mathbf{q} \to 0)$ values, assuming $\mathbf{q}^2$ behavior close to the low wave-vector regime. Afterwards, the obtained zero-value structure factors were used for calculations of the mixing parameter B as defined in Chapter 2 eq. 2.31.



(A) (B)

FIGURE 5.3: The mixing parameter of DPPC/DLiPC mixture as a function of lipid bead type calculated for (A) the total bilayer for different concentrations of phospholipids and (B) the leaflets and the whole bilayer. Systems of 1024 lipids and 8192 water beads.

As can be seen from Fig. 5.3, the mixing parameter has a weak to no dependency on the bead type. This confirms our previous assumption that the choice of a reference point in a lipid has little to no influence on the mixing behavior of the system. More generally, we can conclude that mixing is a global system property instead of a local molecular characteristic. We have observed the same trend regardless of system size and composition of the mixture (see Fig. 5.3 (A)). More than that, the choice whether to treat systems of separate leaflets or bilayer as a whole, had also almost no influence on the final mixing parameter (see Fig. 5.3 (B)). In Fig. 5.4 one can see the final mixing parameters averaged over beads, leaflets and system sizes. As can be observed, there seem to be a linear trend

so it can be concluded, that B depends on the composition of phospholipids in the framework of the current study. Of course, we can always do an average over those values as they are overlapping up to the error, but more concentrations are required to surely establish or disregard this trend.

Although values for the mixing parameter are quite low, a slight demixing trend can be observed in such systems. One can see formations of 'clusters" in Fig. 3.7 meaning that mixing is not ideal and so a positive value of B is expected. Lipid pairs of DPPC and DOPC, DLiPC and DOPC were also studied and found to show similar trends, although the overall B values are smaller and harder to extract due to higher similarity of the molecules in those lipid pairs. Thus, our results confirm the slight demixing trend in the phospholipid mixtures of species with different saturation level.

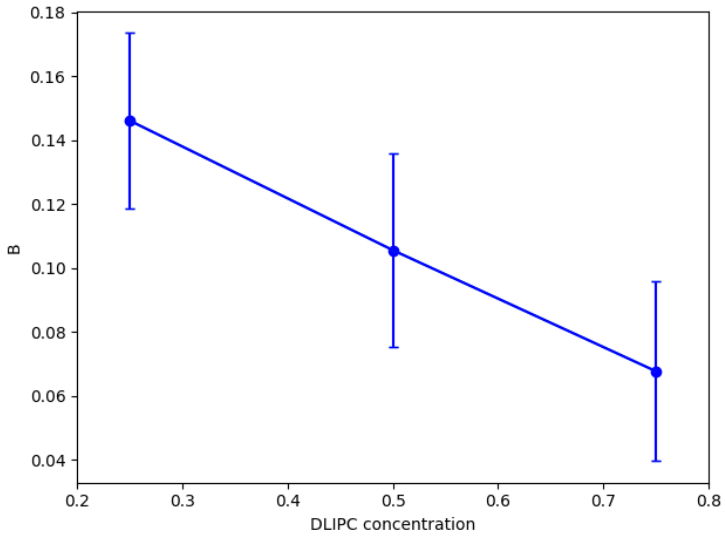


FIGURE 5.4: The mixing parameter of DPPC/DLiPC mixture as a function of DLiPC concentration. Value is averaged over systems of 512 lipids, 1024 lipids and 8192 water beads, and over bilayer representation: separate or both leaflets.

5.1.3 Compressibility of the lipid mixtures

Table 5.1 summarizes the findings regarding the compressibility of the binary mixtures of phospholipids. Furthermore, one can see detailed data for DLiPC and DPPC mixture at different concentrations: 25%, 50% and 75%. Two system sizes were studied and compared to verify for the system size effects that might be present. Results for the false mixtures of DPPC and DLiPC are shown. In the latter case, values for thickness, APL, K_A^{fluct} and K_A^{eqos} are the same as for the pure bilayers, but values of K_A^{str} were calculated as for the real binary mixtures using the KB relations as was defined in Chapter 2 eq. 2.25. One

TABLE 5.1: Structural properties of the bilayer mixtures: area per lipid (APL, nm^2), thickness (z, nm) and stretching modulus (K_A, mN/m).

	APL	z	K_A^{fluct}	K_A^{str}			K_A^{eqos}
				upper×2	lower×2	bilayer	
DPPC/DPPC*	0.594	4.16	258.6±64.4	253.5±7.0	256.6±6.7	275.8±10.3	276.2±9.8
DLiPC/DLiPC*	0.679	3.97	281.2±57.2	219.6±15.5	220.4±14.9	278.6±18.0	314.2±12.4
DLiPC/DPPC[1] 1:3	0.619	4.09	341.2±73.3	242.7±4.2	242.9±4.1	293.1±7.4	297.2±11.5
DLiPC/DPPC[1] 1:1	0.643	4.03	234.1±50.2	232.6±2.5	231.0±9.8	280.8±9.1	301.5±12.3
DLiPC/DPPC[1] 3:1	0.664	3.99	348.8±106.7	222.0±3.9	224.6±3.0	279.3±9.0	308.2±10.2
DLiPC/DPPC[2] 1:3	0.619	4.12	238.6±58.6	200.0±10.2	191.2±3.7	294.0±12.2	307.6±12.3
DLiPC/DPPC[2] 1:1	0.642	4.03	260.2±70.8	249.8±6.4	261.3±6.1	292.9±12.0	302.9±10.1
DLiPC/DPPC[2] 3:1	0.662	3.99	275.0±32.9	242.8±7.3	261.2±7.9	296.3±11.8	307.6±12.3

[fluct] volume fluctuations; [str] KB relations; [eqos] equation of state

* fake mixtures of 50% 512 lipids; [1] 512 lipids; [2] 1024 lipids

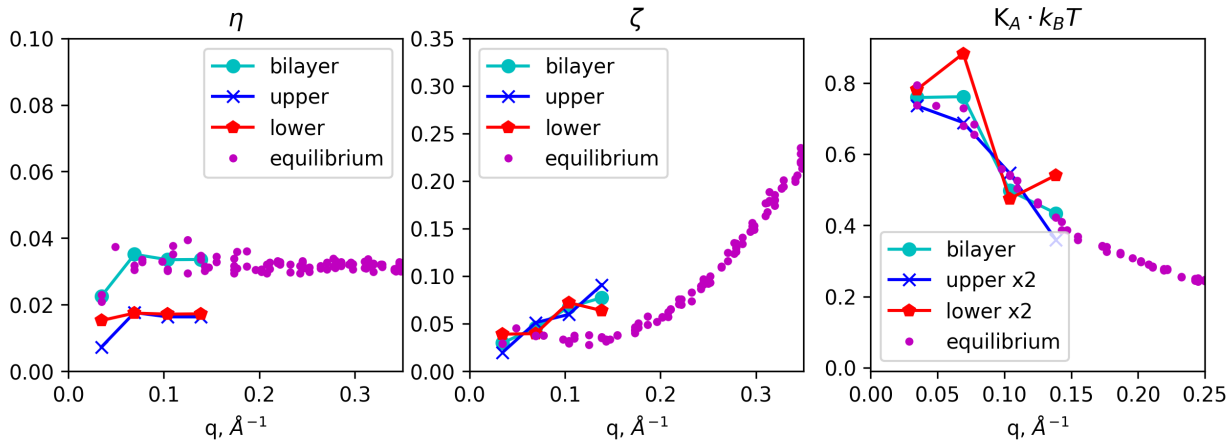
can see that for false mixtures the convergence of the three methods can be considered good.

On the other hand, in the real mixture case the convergence of the crossed density correlation functions is slower than in the pure case, thus, the resulting error from the fluctuation method here goes up to 30%. It can be seen that for all cases the agreement between the leaflet calculation is good, while the full bilayer calculation of K_A^{str} is typically consistent with fluctuation and equation of state approaches. The overall trend is that separete leaflets are softer than the whole bilayer as was observed in the pure bilayer case. One can see that there is a reasonable agreement between two systems of 512 and 1024 lipids. It can be concluded, that this table demonstrates that the values of the compressibility deduced from the structure factor with the Kirkwood-Buff integrals are reasonable and that the convergence of the partial structure factors is satisfactory.

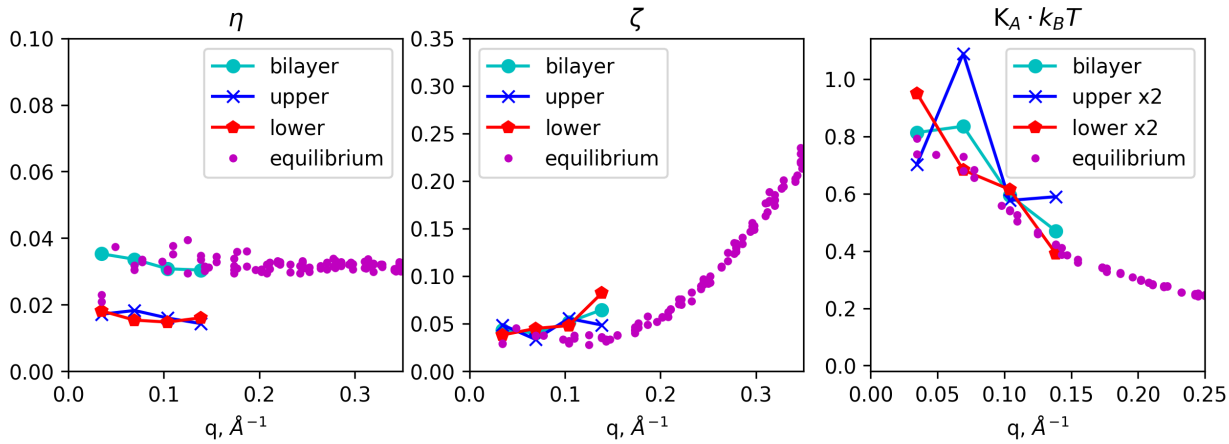
One can also see from Tab. 5.1 the structural characteristics of the studied mixtures: thickness and area per lipid. It is interesting to observe that both values change linearly with the concentration of components in the system. DLiPC lipid seem to increase the membrane area by 12% in the highest studied concentration of 75%. On the other hand, the thickness of such bilayers is decreasing, guarding the overall membrane volume approximately constant. This behavior is consistent with the expectation and both thickness and area per lipid of studied systems are not dependent on the system size.

5.1.4 Linear response in density

The next step was to apply the linear response approach to phospholipid mixtures. As was defined in Chapter 2 one can calculate directly the Kirkwood-Buff parameters η and ζ required for finding chemical potential derivatives, mixing parameter, etc. (see Appendix A eqs. A.35 and A.34). As can be seen in Fig. 5.5, similarly to the pure bilayer



(A) $U_0 = 0.025$ kcal/(mol·Å).



(B) $U_0 = 0.05$ kcal/(mol·Å).

FIGURE 5.5: Dependence of Kirkwood-Buff parameters η and ζ on the external force, using a bias potential and linear response.

case we have tested two force amplitudes: 0.025 kcal/(mol·Å) and 0.05 kcal/(mol·Å). On the plots one can see the comparison of the data from the equilibrium and biased simulations. It can be observed that there is good agreement for η and compressibility but ζ is only consistent at low-**q** values. Furthermore, this results also prove the idea about the replicas: our system can be treated as a system of two uncoupled monolayers. The best illustration for that argument is the η dependence on **q**. It can be seen that values for the whole bilayer (cyan line) and upper or lower leaflets (blue and red lines correspondingly) agree up to a prefactor of two. Values for ζ and K_A are already multiplied by two for monolayers and also illustrate the good consistency between the three approaches.

Though promising at first, the linear response method for calculating thermodynamic properties using the Kirkwood-Buff theory did not show much improvement in simulation time or accuracy. Even though we have found that it was possible to exert an

external force directly with LAMMPS without the use of Python-LAMMPS communication, the time required for a system to equilibrate in a biased state was not smaller than the one needed for an equilibration calculation. This can be due to a too small applied force field, not strong enough to cause an immediate response of the density modes. Going up to higher amplitudes results in distorted systems with a response exceeding the linear regime boundaries. We therefore abandoned this route for effectively calculating the mixing parameter B

5.2 Binary mixtures of lipid and Cholesterol

Let us present now the results for the mixing of lipids and Cholesterol. We have studied systems of 512 molecules in total (excluding water) with 0.25 and 0.5 molar fraction of Cholesterol. This molecule is known to be rigid and so it should be less important which representative bead one chooses to use. Here we present results as if Cholesterol molecule was somewhat similar to the lipid one, so Cholesterol beads are paired with the tail beads of a given lipid. It was seen from the simulation snapshots that sterol prefers to localise in the hydrophobic region of the lipid carbon tails, thus, such choice is justified.

As was discussed in Chapter 1, Cholesterol may induce domain formations of denser regions in the phospholipid mixtures. The typical size of our simulation patches was too small to observe this phenomena. Although we could not detect any phase separation in the studied bilayer systems, the overall trend of stiffening of the systems with higher sterol concentration was still confirmed.

To compare the changes induced by Cholesterol molecules on the lipid bilayer, one can use the number density profiles along the z axis (lipid bilayer normal). As can be seen in Fig. 5.6 and Fig. 5.7, the presence of the sterol lipid has a slight effect on the bilayer structure. The height of the profile is due to the number of lipids present in the system and is left without normalization so the changes in the shape of the density profile are easier to observe. A more detailed set of density profiles for the studied systems can be found in the Appendix C. One can see that the peaks tend to shift to the outer edge of the bilayer, which signifies some amount of thickness change. In systems with sphingomyelin the observed sharper peaks can signal the existence of ordered phases. That is expected as such lipids are known to be in the gel or liquid ordered phase at $T = 298$ K, thus, oscillations in PSM and SSM could point to some amount of crystallisation present in the system. A similar reasoning can be used for the DPPC bilayer to explain the formation of more pronounced peaks in both tail and head regions. On the other hand, unsaturated phospholipids as DLiPC, DOPC, DOPE seem to show a flattening of the profile. Thus, it is possible to assume that such lipids show a more homogeneous spread of the beads

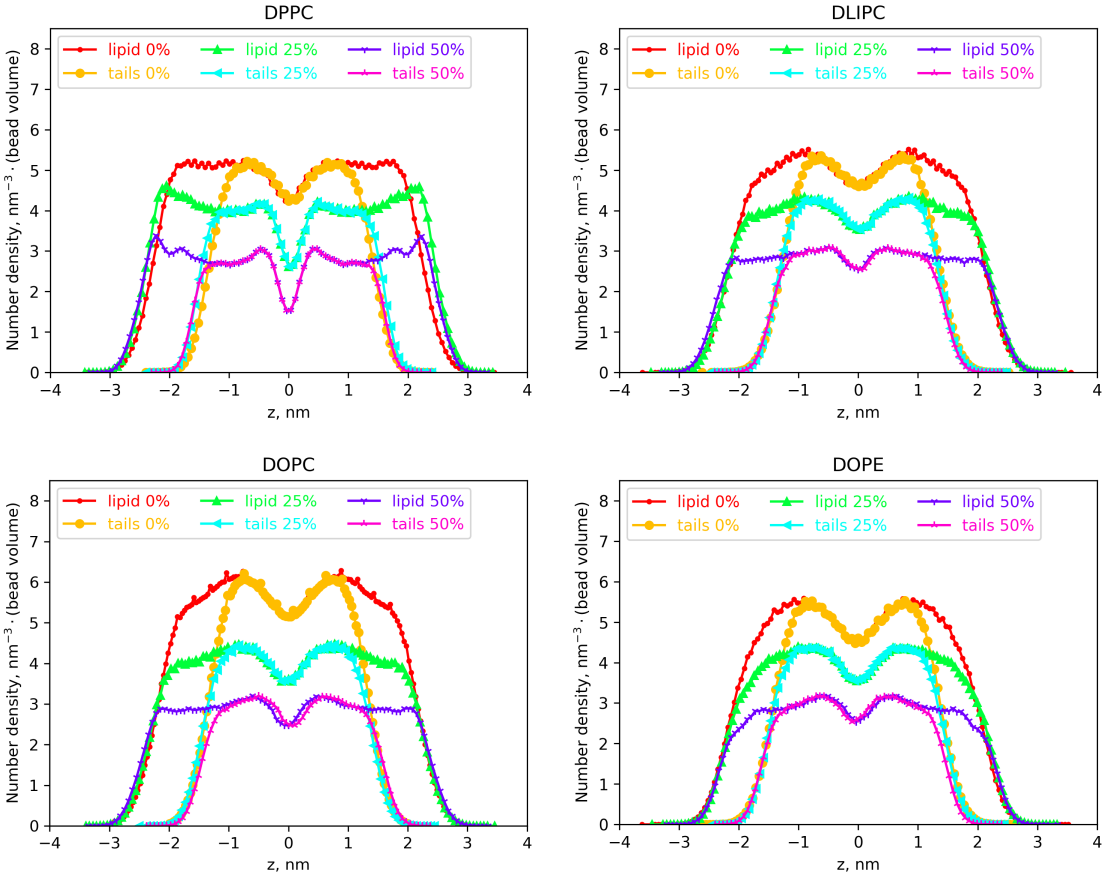


FIGURE 5.6: The number density profiles of the model lipids: DPPC, DLiPC, DOPC, DOPE, for different Cholesterol concentrations. Number density is normalized by the inverse of the bead volume for each lipid. Here, lipid and tails correspond to the profile of the whole lipid molecule and only the carbon tails.

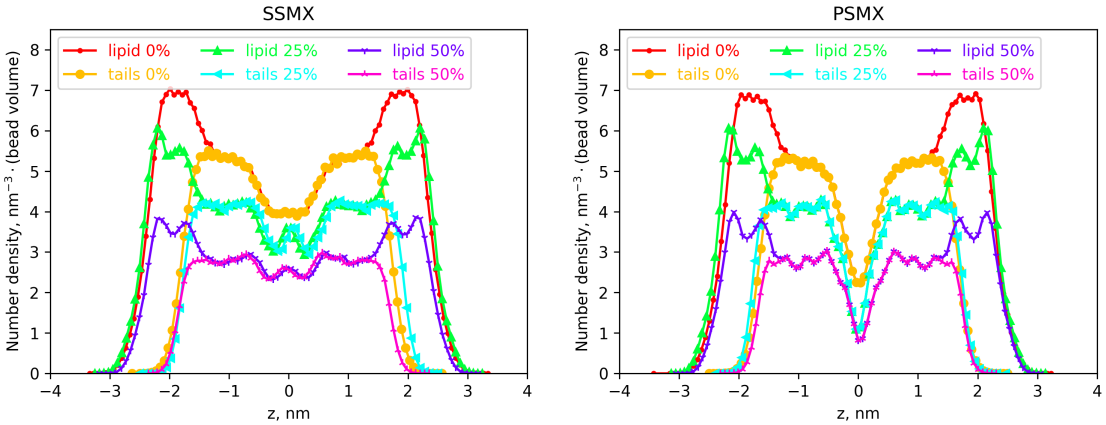
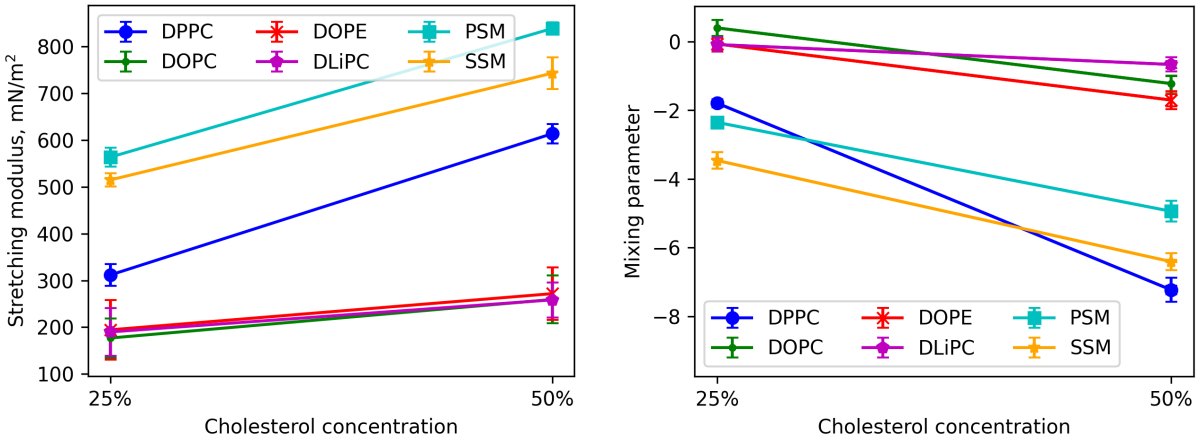


FIGURE 5.7: The number density profiles of the model sphingolipids: SSM, PSM, for different Cholesterol concentrations. Number density is normalized by the inverse of the bead volume for each lipid. Here, lipid and tails correspond to the profile of the whole lipid molecule and only the carbon tails.

along the z direction, meaning Cholesterol is probably "stretching" the tails of the lipid molecules and making a more rigid bilayer. The latter is a fact that is true for all studied systems.

Cholesterol is known to mix preferentially with saturated lipid molecules such as DPPC, PSM, SSM. There is a maximal solubility fraction of Cholesterol, as pure Cholesterol bilayers do not exist [76]. We show here the results that obtained for both PC/PE lipids and sphingolipids. In all cases one can see a trend of decreasing B with increase of Cholesterol in the system (see Fig. 5.8). Similar trend was observed for phospholipid mixture in the previous Chapter: with the increase of DLiPC we observed a decrease of the mixing parameter. Negative B values are interpreted as a strong pro-mixing trend between lipid and Cholesterol.



(A) (B)

FIGURE 5.8: (A) The stretching modulus and the mixing parameter B as a function of the Cholesterol concentration in Cholesterol/lipid mixtures averaged over leaflet representation: upper, lower leaflet and the whole bilayer. (B) The mixing parameter of the lipid mixtures with Cholesterol as a function of the Cholesterol concentration.

The structural properties of the studied bilayer systems change with Cholesterol concentration. One can see how the thickness and area per lipid evolve for the model membranes in Tab. 5.2. The thickness of such systems is visibly increasing, while the area per lipid decreases. This trend is true for all model glycerophospholipids we investigated in this work. On the other hand, sphingomyelin shows trend of decrease of both thickness and area per lipid. This probably signifies that the system is getting denser so the packing of the molecules increases. This can certainly be a result of the Cholesterol "condensing effect".

One can see from Tab. 5.2 the compressibility calculated for mixtures of lipids and Cholesterol. We found that the constant σ fluctuation methods gives extremely large error bars

TABLE 5.2: Structural characteristics (APL - area per lipid, nm^2, z - thickness, nm, K_A - area stretching modulus, mN/m) and mixing parameter (B) for lipid mixtures with Cholesterol.

	APL	z	B	K_A^{fluct}	K_A^{str}		
					upper×2	lower×2	bilayer
25% Cholesterol							
DPPC	0.476	4.58	-2.61±0.11	529.6±433.4	273.0±10.8	279.8±8.5	322.1±12.4
DOPC	0.564	4.28	0.29±0.02	352.5±83.4	144.3±5.4	143.3±6.5	235.9±12.7
DLiPC	0.573	4.18	-0.25±0.03	346.1±121.4	153.4±6.4	154.0±6.8	297.5±12.8
DOPE	0.538	4.24	-0.37±0.06	332.1±117.7	148.6±5.2	150.8±5.2	304.5±11.5
SSM	0.438	4.70	-2.52±0.32	1577.9±1542.8	517.1±13.7	496.2±13.0	531.3±10.6
PSM	0.437	4.52	-1.81±0.13	1559.7±4887.5	550.8±15.8	554.0±15.8	592.7±19.8
50% Cholesterol							
DPPC	0.421	4.75	-6.71±0.35	1703.5±2659.4	584.9±9.9	630.9±10.9	625.5±11.7
DOPC	0.469	4.57	-1.68±0.11	510.5±266.7	220.5±6.6	226.1±6.4	331.3±10.1
DLiPC	0.476	4.45	-1.02±0.09	511.2±232.5	231.5±5.2	232.3±5.3	311.6±10.9
DOPE	0.458	4.41	-1.87±0.24	815.1±557.3	237.3±5.8	240.0±5.5	374.0±8.6
SSM	0.414	4.61	-4.69±0.15	2381.6±3306.8	699.3±18.9	748.5±20.4	781.3±19.4
PSM	0.411	4.50	-3.96±0.09	2043.5± 4599.0	829.1±18.9	829.0±18.6	857.7±17.6

$^{\text{fluct}}$ volume fluctuations, $^{\text{str}}$ zero limit structure factor

in the case of mixtures containing DPPC, PSM and SSM. This is likely due to both large spatial and temporal heterogeneities caused by the strong affinity of cholesterol for saturated lipid chains. The density fluctuation approach at constant volume seems to give more reliable estimates of the elastic properties of the bilayer. One can see from Fig. 5.8 that the ability of the system to compress is strongly influenced by the Cholesterol for saturated lipid (DPPC) and sphingolipids (PSM, SSM), while the unsaturated lipids (DOPC, DOPE, DLiPC) experience relatively small changes. Although it is clear that the impact of the sterol lipid is not of the same amount, all studied lipids showed tendency to increase the stretching modulus in the mixture with an increase of the Cholesterol concentration. Thus, Cholesterol increases the rigidity of the membrane as expected. It is interesting to mention that DPPC seems to be the most responsive lipid to the presence of Cholesterol in this study.

In Fig. 5.8 (B) one can see that systems with sphingomyelin (PSM, SSM) have much more negative values of the mixing parameter than the ones with phospholipids. On the other hand, unsaturated phospholipids have close to negative zero values of B which means such systems are most likely in a state that is similar to an ideal mixed solution. With the exception of DOPC lipid, all studied system show similar trend to form super mixed structures with Cholesterol. Considering that the size of the bilayer in this work was comparably small (maximal lateral size was approximately 19 nm), it was not possible to observe lipid rafts or similar formations. However, it was found that Cholesterol is indeed placed in the hydrophobic region of the bilayer sheet, located between the tail groups of

the lipids. It can be seen from Fig. 5.8 that sphingolipids are strongly influenced by the Cholesterol concentration. We explain this phenomena by the fact that both PSM and SSM are saturated lipids. On the other hand, it can be seen that the decrease of B in the mixture of Cholesterol and DPPC is much larger than in the case of other lipids. DPPC seems to mix even better with Cholesterol than sphingolipids, having the lowest mixing parameter value as can be seen from Tab. 5.2. Our study thus nicely confirms that Cholesterol mixes preferably with saturated lipids DPPC, PSM, SSM and this is supported by a strongly negative B value. Unsaturated lipids still mix well with Cholesterol as supported by a slightly negative B value, with the only exception of DOPC which seems to "dislike" Cholesterol with a relatively low value of the mixing parameter $B = 0.29$. The strong pro-mixing behavior of saturated lipids is conform to the established consensus. A number of theoretical models support this idea, such as "umbrella effect" [25] or "condensed complexes" [62].

It is interesting to mention here that the slope of the $B(x_{\text{Chol}})$ is similar for DOPC and DOPE lipids, so it is likely that the head group of the phospholipid does not play a crucial role in the mixing trend with the concentration, while influencing the overall B value. Comparing DOPC and DOPE mixing parameters it can be seen that the PE lipid shows lower B values meaning it has more intense bonding with the Cholesterol molecules. Interestingly enough, all studied lipids with the exception of DLiPC and DPPC show similar concentration dependency. It is still hard to assume a reason behind such behavior. DLiPC is the most fluid model lipid in this work as it has a double unsaturation on each tail, while DOPC and DOPE have only one unsaturation per tail, and DPPC, PSM and SSM are saturated lipids. The latter argument can also be used to explain that the mixing parameter is overall lower for saturated lipids than the one for the unsaturated ones (see Fig. 5.8). Furthermore, both pairs - SSM, PSM and DOPC, DOPE - have the same tail groups (SSM has one extra bead compared to PSM) and show similar behavior with solution concentration.

It can be concluded that mixtures containing Cholesterol tend to form well-mixed solutions with negative mixing parameter B and high stretching modulus. Both mixing parameter and stretching modulus do not depend or have only a slight dependency on the bead type (NC, PH, GL, etc). The bilayer representation choices (upper, lower leaflet or whole bilayer) give consistent results. Saturated lipids tend to mix more easily with Cholesterol, having on average four times lower B that unsaturated ones while Cholesterol seems to increase the stretching modulus values in these systems. In all cases the results are in qualitative agreement with predictions and experimental results. The strong negative values of B certainly points to strong positional correlations between Cholesterol and saturated phospholipids. It means that the low $\mathbf{q}$ Cholesterol and phospholipid

density modes cannot fluctuate independently from each other. This goes in the same direction as a "condensed complex" picture which for instance associates locally one Cholesterol and one saturated lipid molecule in 1:1 manner. Independent density fluctuations would break this kind of ratio over large length scales and are therefore unfavorable.

Summary

Let us conclude the binary mixture chapter. It was possible to compute the density modes cross-correlations in the low $\mathbf{q}$ regime for a number of binary mixtures: DLiPC/DPPC mixture and mixtures of model SPICA lipids with Cholesterol. First of all, the concept of "fake" mixtures was introduced to provide a first level of control of the overall procedures. The so-called fake mixtures were created on the base of pure DPPC and DLiPC bilayers. As expected, such mixtures are found to be ideal with the mixing parameter B equal to zero. Furthermore, consistent compressibility values were recovered from the false mixtures. Next step was to apply the given algorithm to the real phospholipid mixture, in the scope of this project we presented DLiPC and DPPC mixture in different compositions and system sizes. As in the pure bilayer case, the compressibility falls in the expected value range. Mixing parameter was obtained for this situation and it was found that on average $B = 0.1$, which demonstrates that these phospholipid molecules mix well though non ideally. Moreover, the obtained results seem to be immune to the undulation effects, especially if the full bilayer is considered for calculation. Unfortunately, the application of bias onto the binary mixtures and the linear response approach did not work as efficiently as was expected initially. Finally, the same methodology was applied to the mixtures of phospholipids and Cholesterol in two concentrations: 25% and 50% Cholesterol. These systems showed consistent results and we obtained a confirmation of the preference of Cholesterol molecule for saturated lipid compound. In the highest studied concentration of the sterol molecule, all lipids were found to mix well with negative B values.

Conclusions and perspectives

Conclusions

In the current project we used Molecular Dynamics to study lipid bilayers. Lipid molecules were modeled with the SPICA coarse-grained force field. Systems of pure 1,2-dipalmitoyl-sn-glycero-3-phosphatidylcholine (DPPC), 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), 1,2-dilinoleoyl-sn-glycero-3-phosphocholine (DLiPC), 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE), Palmitoyl sphingomyelin (PSM) and Stearoyl sphingomyelin (SSM) bilayers were studied. Typical bilayers consisting of 512 lipids and 8192 SPICA water beads were simulated in the fluid phase (298 K for glycerophospholipids and 318 K for sphingomyelin). It was found that sphingolipids and DPPC form thicker bilayers with smaller area per lipid than unsaturated lipids (DOPC, DOPE, DLiPC). As well as that, the stretching modulus was calculated using three methods: area fluctuations, zero-limit of structure factor and direct stretching of membrane with an external pressure. The second method also allowed to compare results for separate leaflets and full bilayer. It was found that these methods agree regarding the magnitude of the obtained values.

A linear response approach was introduced and checked using a periodic external potential. We found that the theory can be applied to bilayer systems with a limitation on the maximal wave-vector and force amplitude, beyond which the response in the density modulation was found to be no longer linear.

During these calculations we questioned the choice of the lipid representation and the possibility to reduce a lipid molecule to an explicit point. Three possibilities were compared (center of mass of the lipid, pseudo center of mass of the lipid, *i.e.* neglecting mass

differences and weighing beads equally, and selecting a representative bead for the whole lipid molecule). If in a usual fluid the representation does not matter in the macroscopic limit, the membrane undulations were shown to affect in a non trivial way the thermodynamic properties of lipid bilayers. In the end, we found that combining full bilayer statistics and optimal choice of the representative beads can reduce the impact of tilt and undulation corrections.

Our investigations culminated with the study of binary mixtures. To start with, the consistency of the Kirkwood-Buff theory coupled to the density mode correlations analysis was checked with the use of false mixtures. False mixtures were created from pure bilayer systems by picking up random lipid molecules and labelling them as if they were different species. Such approach allowed us to apply our thermodynamic analysis to bilayers that were ideal by construction and derive the apparent mixing parameters in a controlled case. For this purpose false mixtures of DLiPC and DPPC were created with 1:3, 1:1 and 3:1 compositions. It was found that for such mixtures the mixing parameter was indeed approximately zero irrespective of the lipid representation that was considered. These findings allowed us to conclude that the choice of the reference point had little to no influence on the value of the mixing parameter and that the whole approach was a viable route for studying the mixing behavior phospholipid mixtures.

After establishing the protocol for the false mixture case, a binary mixture of DPPC and DLiPC was studied. The same compositions of 1:3, 1:1 and 3:1 were created for two system sizes: 512 and 1024 lipids in total (whole bilayer). Such bilayers were created with a symmetrical composition of the upper and lower leaflets. During this step we adapted and tested a special instance of replicated fluids theory, each leaflet representing a replica. The compressibility and the mixing parameter were calculated for the upper leaflet, the lower leaflet, the whole bilayer and compared afterwards. It was found that the stretching modulus depends slightly on the representation of the bilayer and is consistent in magnitude with expected values. The mixing parameter of this phospholipid mixture was found to be slightly positive and dependent on the concentration of the components. This suggests that a theory of regular solutions with quadratic interactions does not fit exactly our simulations.

Next step was to study mixtures of the available model lipids with Cholesterol. For this purpose, systems of 25% and 50% Cholesterol were created, with a total of 512 lipids in both leaflets. This part of the study revealed that the sterol molecule is subject to strong mixing in almost all studied systems, especially in the case of saturated lipids (DPPC, PSM, SSM). The compressibility of these mixtures increased with the Cholesterol concentration, while the mixing parameter B decreased. Unsaturated model lipids (DOPC,

DOPE, DLiPC) showed less susceptibility to the Cholesterol presence in the system, however the same trend was still observed as in the case of the saturated lipids.

To conclude, we have introduced and thoroughly tested a methodology for studying the thermodynamics of mixing of lipid binary mixtures.

Perspectives

Let us discuss now a few ideas for the future developments that could be inspired by the current project. The dependence of the mixing parameter on the molar concentration of the components should be investigated systematically, for all the lipid pairs available in the SPICA CG force-field. We came across some indirect consequences of the membrane undulations. It would be interesting to find a way to reduce their influence, for example by applying an external tension to the lipid membrane, or enforcing a harmonic constraint in the normal z direction. Both methods could be compared with the current unconstrained bilayer approach.

Another exciting project would be to compare the SPICA and MARTINI models. The latter could provide us with faster simulations and perhaps better statistics (MARTINI does not require long-range electrostatics calculations). The MARTINI force field is simpler to use and more lipid molecules are available. The consistence between both models would be a very interesting fact to check for, and a more comprehensive lipid-lipid interaction map could be built.

Parallelization and computational efficiency tends now to be in favor of GPU based molecular dynamics engines. Further studies should be in priority based upon such highly efficient MD codes. These codes could make it possible to extend the current approach to the more realistic and computational extensive all-atoms model systems.

The long-range electrostatics of the SPICA force field is inconvenient as far as computational efficiency and parallelization is concerned, but it could provide a framework for studying the effect of added salts on the systems. It could either be used to simulate charged lipids, such as phosphatidylglycerol (PG) or cationic lipids, and to question the influence of the electrostatic charge on the thermodynamics of mixing. For example, 1,2-dipalmitoyl-sn-glycero-3-phosphoglycerol, sodium salt (DPPG) is one of the main lung surfactants alongside DPPC. Studying their cross-interactions could improve the understanding of biochemical processes inside lung membranes.

Generally speaking, the method described in this thesis can be applied to any binary mixture and our approach is able to determine quantitatively the derivative μ_{12} of the chemical potential with respect to the concentration. One could imagine improving upon

the theory of regular solutions and possibly finding more sophisticated expressions of the chemical potential in non-ideal mixtures. For instance alternative Cholesterol-lipid mixing models [88] could provide better insights on the condensed complexes and phase separation phenomena.

It would also be interesting to extend our approach to some non lipid amphiphilic molecules, such as fatty alcohols and cationic surfactants which have pharmaceutical or industrial applications [54]. Membrane spanning molecules, the so-called bolalipids, typically used to increase the robustness of industrial membranes would also be worth investigating [8].

As was briefly mentioned above, one could also create mixtures of higher complexity, meaning ternary or more components fluids. Our approach can be certainly generalized to such cases and applied to the thermodynamic properties of complex lipid solutions such as the one used to vectorize the mRNA vaccines (phospholipids, cholesterol, cationic and pegylated lipids). More generally, a better understanding of mutual lipid interactions will benefit from computational approaches such as ours, and will be beneficial to the whole biochemichal and biophysical community.

Résumé en français

Mélanges non-idéaux dans les bicouches phospholipidiques: approche par dynamique moléculaire

Introduction

Les membranes cellulaires contiennent un grand nombre de types de lipides différents. Les phospholipides sont des éléments constitutifs essentiels des membranes des cellules vivantes qui permettent aux organismes d'exister. Un phospholipide typique se compose d'un groupe "tête" et d'un groupe "queue" qui sont respectivement hydrophiles et hydrophobes. La variété des longueurs de queue, des types de groupes de tête et du degré d'insaturation conduit à un grand nombre de molécules différentes possibles. Compte tenu de l'importance biologique évidente de l'étude de ces systèmes, ils sont l'objet d'une grande attention [60, 67].

En même temps, l'accroissement de la puissance de calcul disponible et l'amélioration des modèles moléculaires fait des simulations de dynamique moléculaire (MD) une approche prometteuse. À ce jour, la dynamique moléculaire a été largement utilisée pour l'étude de systèmes protéines-membranes réalistes [37, 38], de la diffusion et du mélange [84] et des propriétés structurelles de membranes modèles [20, 51, 55]. Des champs de forces (FF) sont disponibles avec une résolution atomique (tout-atome, AA) tels CHARMM-36 [32] ou plus grossière (Coarse-grained, CG) comme MARTINI [39], et la qualité de ces modèles est en constante amélioration.

Dans ce travail, le champ de force SPICA [51, 69] a été utilisé. Ce modèle CG a été construit comme une amélioration du modèle SDK par W. Shinoda et al. [70, 71]. SPICA peut reproduire avec précision les propriétés structurales expérimentales et les distributions spatiales obtenues à partir des simulations AA-MD. Ce champ de force est moins grossier que MARTINI (actuellement le plus populaire pour les membranes) et devrait donc donner un aperçu plus réaliste de la dynamique et de la structure de la bicouche. Couplé au moteur de simulation LAMMPS, il offre de plus larges possibilités de simulation à l'équilibre et hors-équilibre.

Ici, notre intérêt se porte sur les propriétés thermodynamiques des membranes, leur propriétés de mélange et d'étirement en particulier, que nous voulons déterminer quantitativement à l'aide de simulations de dynamique moléculaire de modèles simples, tels que la dipalmitoylphosphatidylcholine (DPPC, 16:0, lipide saturé) et la dilinéoylphosphatidylcholine (DLiPC, 18:2, lipide insaturé). La liste complète des lipides étudiés et leur décomposition en gros-grains est précisée sur la Fig. 7.1.

L'étude par simulation de membrane biologiques permet d'étudier leur structure à une échelle spatiale trop petite ou temporelle trop transitoire pour être accessible aux techniques expérimentales actuelles. Nous présentons dans cette thèse les résultats de simulations de dynamique moléculaire à l'équilibre et hors équilibre de systèmes modèles de membranes lipidiques. Nous nous efforçons d'établir un lien quantitatif entre les observables de la dynamique moléculaire d'une part (facteur de structure, modes de fluctuations de densité) et des quantités thermodynamiques d'autre part (énergie libre de mélange en particulier). Nous comparons diverses prescriptions pour aller de la représentation moléculaire aux propriétés macroscopiques et sondons directement la réponse linéaire d'une bicouche soumise à une modulation spatiale des potentiels chimiques des lipides. Nous étudions également la cinétique des fluctuations des modes de densité. Un mélange modèle de DPPC (dipalmitoylphosphatidylcholine) et DLiPC (lineoylphosphatidylcholine) a été choisi pour l'évaluation initiale, avant d'appliquer la procédure à un ensemble plus large de mélanges lipidiques pertinents, y compris avec du cholestérol.

Méthodologie

Configuration du système

Comme nous l'avons déjà mentionné, les molécules lipides amphiphiles et le solvant (eau) ont été modélisés à l'aide du champ de force CG SPICA: [70, 71]. Le grossissement est ici effectué de la manière suivante: une eau SPICA correspond à trois molécules d'eau réelles et une bille SPICA à trois atomes de carbone. Les interactions non liantes

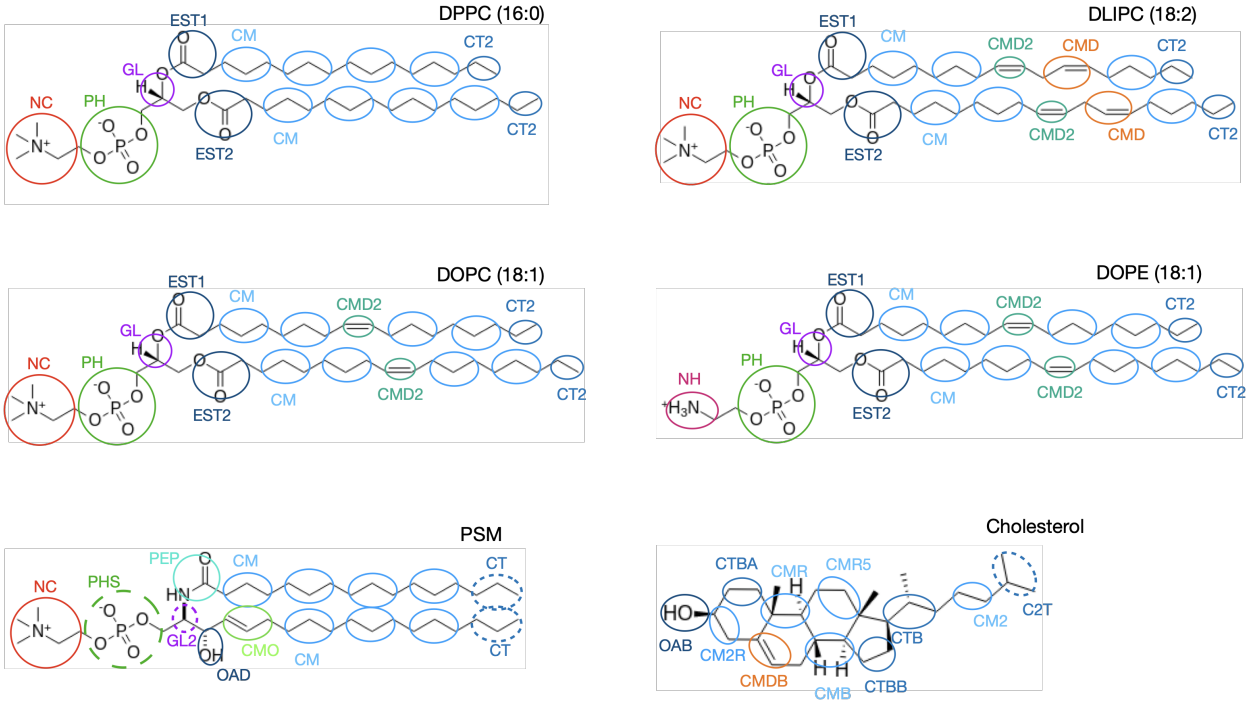


FIGURE 7.1: Représentation de la structure atomique des lipides modèles avec SPICA. Les noms des billes suit la topologie du modèle SPICA, et la taille des billes est proportionnelle à leur masse.

TABLE 7.1: Temps de relaxation typiques des systèmes étudiés à partir des fluctuations de l'aire.

	τ, ns
DPPC	0.65
DLiPC	0.4
DLiPC/DPPC 1:3	0.55
DLiPC/DPPC 1:1	0.50
DLiPC/DPPC 3:1	0.45
DLiPC/DOPC 1:3	0.75
DLiPC/DOPC 1:1	0.42
DLiPC/DOPC 3:1	0.44

de Lennard-Jones (LJ) sont tronquées à 1,5 nm et les interactions de Coulomb sont calculées par sommation d'Ewald en présence de particules chargées dans les groupes de tête: [71]. La constante diélectrique a été fixée à $\epsilon_r = 80$ pour rendre compte de la propriété diélectrique statique de l'eau. Les potentiels intraparticulaires sont représentés par les potentiels d'étirement de liaison harmonique et de flexion d'angle pour les paires liées 1-2 et 1-2-3, tandis que toutes les autres paires interagissent via des forces non liantes. Cette force est donnée par le potentiel (12-4) LJ pour les paires impliquant l'eau et par le potentiel (9-6) LJ pour toutes les autres paires. Toutes les billes ont des masses distinctes qui correspondent à la somme des masses atomiques réelles des composants de chacune.

Les configurations initiales ont été créées par placement aléatoire en utilisant le logiciel PACKMOL [41, 44] et en contraignant les positions du groupe de tête et les dernières billes du groupe de queue des lipides dans des plans horizontaux xy de sorte qu'une membrane soit formée initialement avec des molécules lipides orientées le long de l'axe z.

Des simulations d'équilibre ont été réalisées à l'aide du logiciel LAMMPS [59] pour un certain nombre de systèmes purs (DPPC, DOPC, DLiPC, DOPE, PSM, SSM) et de mélanges binaires (DPPC/DLiPC, DPPC/DOPC, DPPC/Chol, DOPC/Chol, DLiPC/Chol, DOPE/-Chol, PSM/Chol, SSM/Chol). La taille des systèmes simulés a été variée (256-512-1024 lipides) afin de sonder les effets de taille finie et d'ondulation de Helfrich. D'autres simulations ont été réalisées en présence d'une contrainte anisotrope (compressibilité), d'une modulation du potentiel chimique (réponse linéaire en densité) ou d'une force de traction du feuillet (friction interne de la membrane). Il a été observé que les temps de relaxation des systèmes étudiés étaient de l'ordre de 0,5 ns (voir Tableau 7.1). Il est intéressant de mentionner ici que le temps de relaxation des lipides saturés était un peu plus élevé que celui des insaturés, ce qui, selon nous, peut être dû à la plus grande fluidité de ces derniers.

L'analyse des données obtenues a été réalisée à l'aide de scripts Python avec l'aide des modules scientifiques MDAnalysis [19, 50, 80] et MDTraj [45]. L'analyse visuelle, les instantanés du système et le prétraitement ont été effectués à l'aide du logiciel Visual Molecular Dynamics (VMD) [26, 34].

Contexte théorique

Nous présentons maintenant la théorie qui sous-tend les calculs effectués et certaines méthodes d'analyse que nous avons utilisées. La section suivante est écrite en s'inspirant des fluides tridimensionnels (3d). L'adaptation aux fluides bidimensionnels (2d) et aux bicouches lipidiques est naturelle, grâce à la correspondance $V \leftrightarrow A$ et $P \leftrightarrow -\sigma$, où V est

le volume, A est la surface d'un système et P, σ sont la pression et la tension de surface correspondantes.

Théorie des solutions régulières

Les membranes lipides sont des mélanges de composition complexe. La thermodynamique de tels mélanges, si une seule phase est présente, est contrôlée par $G(P, T, n_1, ..., n_p)$ - l'énergie libre de Gibbs, avec une pression P, une température T, une composition molaire $n_1, ..., n_p$, et p le nombre de composants. Dans le cadre d'une évolution isotherme-isobare, le changement de la fonction de Gibbs G est directement sous le contrôle des première et deuxième lois de la thermodynamique.

Pour les phases condensées, l'énergie libre de Helmholtz

$$F(V, T, n_1, ..., n_p) = G - PV \tag{7.1}$$

peut être utilisée à la place de G. Il est important de noter que la fonction de Helmholtz est directement liée à l'approche fonctionnelle de la densité et au formalisme des fonctions de corrélation directe. Les transformations à volume constant et à pression constante se révèlent assez similaires pour les phases condensées dans des conditions de pression normales.

Les potentiels chimiques μ_a sont donc dérivés de F et G comme suit :

$$\mu_a = \left(\frac{\partial G}{\partial n_a}\right)_{P,T,n_{b\neq a}} = \left(\frac{\partial F}{\partial n_a}\right)_{V,T,n_{b\neq a}} \tag{7.2}$$

où $\mu_a(P, T, n_1, .., n_p) = \mu_a(V, T, n_1, .., n_p)$ hérite d'une dépendance dans les variables du potentiel utilisé dans la dérivation.

Soit $n = n_1 + ... + n_p$ le nombre total de moles et $x_a = n_a/n$ avec $x_1 + ... + x_p = 1$ les fractions molaires. L'énergie libre (G_{mix}) et l'enthalpie de mélange (H_{mix}) sont définies comme la différence respective d'énergie libre et d'enthalpie entre les composants entièrement séparés et entièrement mélangés, en l'absence de changement d'état discontinu. Le cas le plus favorable se présente lorsque les composés sont si similaires en termes de forme et d'interactions qu'ils peuvent être échangés sans grande conséquence. Dans ce cas, le seul effet du mélange est d'augmenter l'entropie, ce qui signifie que $H_{\text{mix}} = 0$ qui est le postulat de la théorie des solutions idéales. Cette observation est à l'origine de l'intéressant paradoxe de Gibbs concernant le mélange de molécules équivalentes *vs.* non équivalentes.

Pour ce dernier cas, la thermodynamique des interactions lipidiques peut être traitée avec une théorie des solutions régulières [57]. Dans cette approche, dans la limite thermodynamique, l'enthalpie de mélange n'est plus considérée comme nulle mais supposée être une expression quadratique. L'énergie libre de Gibbs du mélange s'exprime alors comme suit

$$\frac{G_{\text{mix}}}{Nk_BT} = x\ln x + (1-x)\ln(1-x) + Bx(1-x) \tag{7.3}$$

où N est le nombre total de particules, k_B est la constante de Boltzmann, $x = x_1, 1-x = x_2$ et B un coefficient sans dimension - paramètre de mélange du mélange binaire donné.

Dans le cadre des solutions régulières, les coefficients additifs par paire B peuvent être combinés pour traiter la thermodynamique des mélanges binaires, au moyen d'une matrice d'interaction quadratique $\sum_{ij} B_{ij}x_ix_j/2$. Un certain nombre de tentatives ont été faites pour déterminer expérimentalement les paramètres d'interaction B_{ij}. Ces approches expérimentales ont été examinées et les valeurs des paramètres d'interaction ont été compilées par P.F.F. Almeida [35].

Relations de Kirkwood-Buff

La théorie de Kirkwood-Buff (KB) fournit le lien souhaité entre l'énergie libre de mélange et la structure moléculaire. Elle est bien établie pour les solutions tridimensionnelles et ne fait pas d'hypothèses sur la nature des interactions entre les molécules. Notre objectif est d'étendre cette approche aux mélanges de lipides sur des membranes fluctuantes. La théorie KB [31] donne une interprétation thermodynamique aux "intégrales KB" (eq. 7.4). Elle résume toutes les connexions entre les fluctuations d'équilibre et les susceptibilités dans le cas de mélanges fluides homogènes monophasés. La théorie des solutions de KB donne une connexion entre les quantités thermodynamiques et les fonctions de distribution moléculaire pour les systèmes multicomposants dans le cas général dans l'ensemble grand-canonique T, V, μ [5].

D'une manière générale, la théorie de Kirkwood-Buff établit une relation directe entre les propriétés thermodynamiques telles que la compressibilité, les volumes molaires partiels et les dérivées des potentiels chimiques en termes d'intégrales dites KB, qui sont définies comme suit:

$$\tilde{h}_{ab}(0) = \int_0^\infty [g_{ab}(\mathbf{r}) - 1]4\pi r^2 d\mathbf{r} \tag{7.4}$$

où $g_{ab}(\mathbf{r})$ est la fonction de corrélation des paires ou fonction de distribution radiale (RDF), $g_{ab}(\mathbf{r}) = \frac{1}{\rho}\langle\sum_{a\neq 0}\delta(\mathbf{r}_a - \mathbf{r}_b)\rangle$, ρ est la densité du système, $\mathbf{r}_i$ est le rayon vecteur de la particule i.

La comparaison de la dérivée des potentiels chimiques du mélange binaire par rapport à la composition dans la théorie KB [5] et dans la théorie des solutions régulières [57] donne l'expression suivante pour le paramètre B apparaissant dans l'eq. 7.3 :

$$2B = \frac{\rho[\tilde{h}_{aa}(0) + \tilde{h}_{bb}(0) - 2\tilde{h}_{ab}(0)]}{1 + \rho x(1-x)[\tilde{h}_{aa}(0) + \tilde{h}_{bb}(0) - 2\tilde{h}_{ab}(0)]} \tag{7.5}$$

Notre objectif est de tester la validité de la théorie des solutions régulières pour les bicouches lipidiques et de déterminer numériquement les paramètres d'interaction B_{ij} et la compressibilité isotherme χ_T associés à une sélection de lipides représentatifs.

Les simulations MD fournissent une estimation directe des fluctuations de densité $\hat{n}_{\mathbf{q}}$ dans l'espace réciproque, à condition que le vecteur de l'espace réciproque $\mathbf{q}$ soit commensurable à la boîte de simulation. Les fluctuations de densité sont directement liées aux facteurs de structure.

Considérons deux espèces $\alpha = 1, 2$ avec N_1, N_2 molécules respectivement, dans un système 2d $\mathbf{r} = x\mathbf{e}_x + y\mathbf{e}_y$ avec des tailles latérales L_x, L_y et des conditions aux limites périodiques appliquées [2]. Les modes de densité peuvent alors être définis comme

$$\hat{n}_{\alpha,\mathbf{q}} = \sum_{j=1}^{N_\alpha} e^{-i\mathbf{q}.\mathbf{r_j}} \tag{7.6}$$

avec $\mathbf{r_j}$ étant la position de la molécule j (*e.g.* centre de masse ou atome de référence) de l'espèce $\alpha = 1, 2$ et $\mathbf{q}$ est un vecteur d'onde qui est restreint aux vecteurs réciproques du réseau $\left(\frac{2\pi}{L_x}m, \frac{2\pi}{L_y}n\right)$ avec $n = -n_{\max}, \ldots, n_{\max}$ et $m = 0, \ldots, n_{\max}$, $n_{\max}$ nombre entier quelconque. Ainsi, nous parcourons la moitié des vecteurs d'onde possibles $\mathbf{q}$, l'autre étant déterminé par la relation de conjugaison $\hat{n}_{-\mathrm{q}} = \hat{n}^*_{\mathrm{q}}$.

Par définition du facteur de structure statique [81, 91] dans le cas des mélanges :

$$S_{\alpha\beta}(\mathbf{q}) = \frac{1}{\sqrt{N_\alpha N_\beta}} \langle \hat{n}_{\alpha,\mathbf{q}} \hat{n}_{\beta,-\mathbf{q}} \rangle \tag{7.7}$$

Nous pouvons maintenant définir la fonction de corrélation dans l'espace réciproque :

$$\begin{aligned} \tilde{h}_{\alpha\alpha}(\mathbf{q}) &= \frac{1}{\rho_\alpha}(S_{\alpha\alpha}(\mathbf{q}) - 1) \\ \tilde{h}_{\alpha\beta}(\mathbf{q}) &= \frac{S_{\alpha\beta}(\mathbf{q})}{\sqrt{\rho_\alpha \rho_\beta}} \end{aligned} \tag{7.8}$$

avec $\rho = \rho_i + \rho_j$ - densité numérique totale.

TABLE 7.2: Propriétés structurelles des bicouches pures: surface par lipide (APL), épaisseur (z) et module d'étirement (K_A).

	APL	z	K_A^{fluct}	K_A^{str}			K_A^{eqos}	K_A^{exp}
				upper×2	lower×2	bilayer		
DPPC[1]	0.594	4.15	258.6±64.4	248.4±8.4	226.0±11.2	248.6±14.4	275.1±10.4	231±20
DLiPC[1]	0.679	3.97	281.2±57.2	216.6±5.5	215.0±5.1	279.8±9.1	319.7±9.6	247±21
DOPC[1]	0.670	4.06	305.1±118.9	213.5±4.1	222.7±4.9	273.6±9.2	-	265±18
DOPE[2]	0.629	4.07	370.0±68.4	197.3±7.6	199.6±7.5	271.8±11.0	-	-
PSM[3]	0.518	4.29	411.9±164.5	227.3±11.4	220.5±10.0	263.1±13.1	-	-
SSM[3]	0.523	4.49	297.9±152.9	221.5±8.9	225.5±8.8	242.5±12.0	-	-

[fluct] fluctuations de volume, [str] facteur de structure a la limite zero, [eqos] équation d'état, [exp] expérience [9, 40, 64]

[1] $T = 298$ K, moyenne sur les perles GL, EST et C1

[2] $T = 298$ K, moyenne sur les perles EST, C1 et C2

[3] $T = 318$ K, moyenne sur les perles PH, GL, PEP, OAD, C1, C2, C3, C4

Prescription pour le passage de moléculaire à macroscopique

Comme l'ensemble grand-canonique n'est pas directement accessible lorsqu'on simule des systèmes moléculaires complexes, nous analysons plutôt ici les fluctuations de densité finies de $\mathbf{q}$, de grande longueur d'onde, des feuillets de la bicouche lipidique. Ces modes de densité ne modifient pas la composition globale du système et devraient devenir indépendants de l'ensemble dans la limite thermodynamique. Les coefficients thermodynamiques d'intérêt (compressibilité, dérivée des potentiels chimiques) sont récupérés en extrapolant les facteurs de structure $S_{\alpha\beta}(\mathbf{q})$ à $q = 0$. Notre approche suppose qu'un feuillet de bicouche lipidique peut être traité comme un fluide ordinaire 2d. Il est donc nécessaire de projeter les conformations moléculaires lipidiques sur des positions uniques dans le plan 2d xy. Nous discutons et comparons un certain nombre de méthodes possibles. Nous observons également que les ondulations de Helfrich hors du plan ont une influence directe sur les statistiques des modes de densité.

Résultats et discussion

Les observables typiques d'intérêt pour une membrane - surface par lipide et épaisseur - ont été calculés. Comme on peut le voir dans les Tables 7.2, 7.3 nos données sont cohérentes avec les résultats expérimentaux correspondants et les données rapportées par les simulations MD. On peut voir que l'épaisseur et la surface par lipide présentent une tendance linéaire vers les valeurs des lipides purs correspondants avec le changement de ratio.

La compressibilité du système a été obtenue par trois méthodes indépendantes: fluctuation de la taille de la boîte à l'équilibre, simulation dans des conditions de contrainte

TABLE 7.3: Propriétés structurelles des bicouches de mélanges de lipides: surface par lipide (APL), épaisseur (z) et module d'étirement (K_A).

	APL	z	K_A^{fluct}	K_A^{str}			K_A^{eqos}
				upper×2	lower×2	bilayer	
DPPC/DPPC*	0.594	4.16	258.6±64.4	253.5±7.0	256.6±6.7	275.8±10.3	276.2±9.8
DLiPC/DLiPC*	0.679	3.97	281.2±57.2	219.6±15.5	220.4±14.9	278.6±18.0	314.2±12.4
DLiPC/DPPC[1] 1:3	0.619	4.09	341.2±73.3	242.7±4.2	242.9±4.1	293.1±7.4	297.2±11.5
DLiPC/DPPC[1] 1:1	0.643	4.03	234.1±50.2	232.6±2.5	231.0±9.8	280.8±9.1	301.5±12.3
DLiPC/DPPC[1] 3:1	0.664	3.99	348.8±106.7	222.0±3.9	224.6±3.0	279.3±9.0	308.2±10.2
DLiPC/DPPC[2] 1:3	0.619	4.12	238.6±58.6	200.0±10.2	191.2±3.7	294.0±12.2	307.6±12.3
DLiPC/DPPC[2] 1:1	0.642	4.03	260.2±70.8	249.8±6.4	261.3±6.1	292.9±12.0	302.9±10.1
DLiPC/DPPC[2] 3:1	0.662	3.99	275.0±32.9	242.8±7.3	261.2±7.9	296.3±11.8	307.6±12.3

[fluct] fluctuations de volume, [str] relations KB, [eqos] équation d'état

* calculs du faux mélange 50% 512 lipids; [1] 512 lipids; [2] 1024 lipids

anisotrope, extrapolation des valeurs du facteur de structure à $q = 0$. Nous avons également considéré les facteurs de structure des feuillets séparés et du système complet. Les résultats pour les systèmes bicouches purs sont présentés dans la Table 7.2 et pour les mélanges de lipides dans la Table 7.3. Il est connu que les ondulations de Helfrich donnent lieu à une dépendance de taille finie [55] de χ_T que nous avons également observée jusqu'à des systèmes relativement petits. Même si les barres d'erreur que nous avons obtenues sont encore significatives, dans l'ensemble, nous avons trouvé que les résultats étaient relativement cohérents.

Le facteur de structure pour le système bicouche pur et le mélange a été calculé (voir Fig. 7.2, 7.3). Trois descriptions possibles des lipides ont été implémentées en utilisant soit le centre de masse (COM) soit des billes représentatives pour localiser chaque molécule lipidique :

$$\begin{aligned} &\text{real COM} && \mathbf{r} = \mathbf{R}_{\text{COM}} = \sum \frac{\mathbf{r}_i m_i}{m_i} \\ &\text{pseudo COM} && \mathbf{r} = \mathbf{R}_{\text{pCOM}} = \sum \frac{\mathbf{r}_i m_i}{m_i} \overset{m_i=1}{=} \frac{\sum \mathbf{r}_i m_i}{N_{\text{beads}}} \\ &\text{bead} && \mathbf{r} = \mathbf{R}_{\text{bead}} \end{aligned} \tag{7.9}$$

Cette dernière procédure nous a permis d'étudier l'influence de la représentation de la molécule sur les propriétés thermodynamiques finales d'intérêt. On s'attend à ce que différentes approches de la description du lipide convergent dans la limite $q \to 0$ à la fois pour les bicouches pures et mélangées. Cela est dû au fait que, dans la limite macroscopique, la façon dont nous décrivons le lipide ne devrait pas avoir d'importance. Par conséquent, nous pouvons utiliser des calculs pour différentes billes, COM et pCOM pour vérifier la cohérence de nos méthodes et résultats.

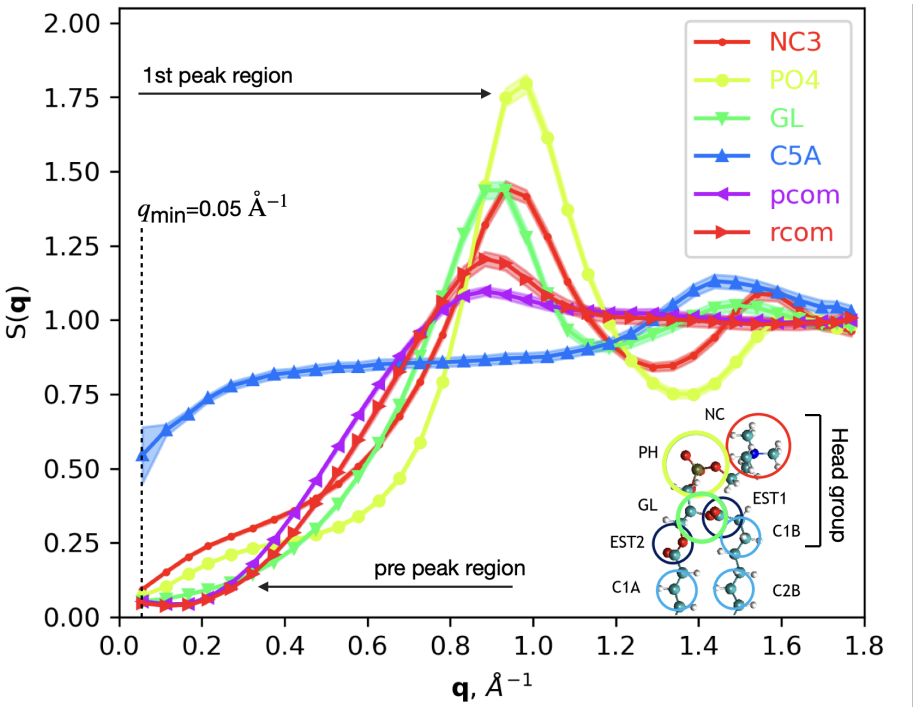


FIGURE 7.2: Facteur de structure de la bicouche pure de DPPC calculé par rapport à différentes billes, pseudo et vrai centre de masse. L'intervalle de confiance est indiqué par une couleur transparente autour des lignes principales.

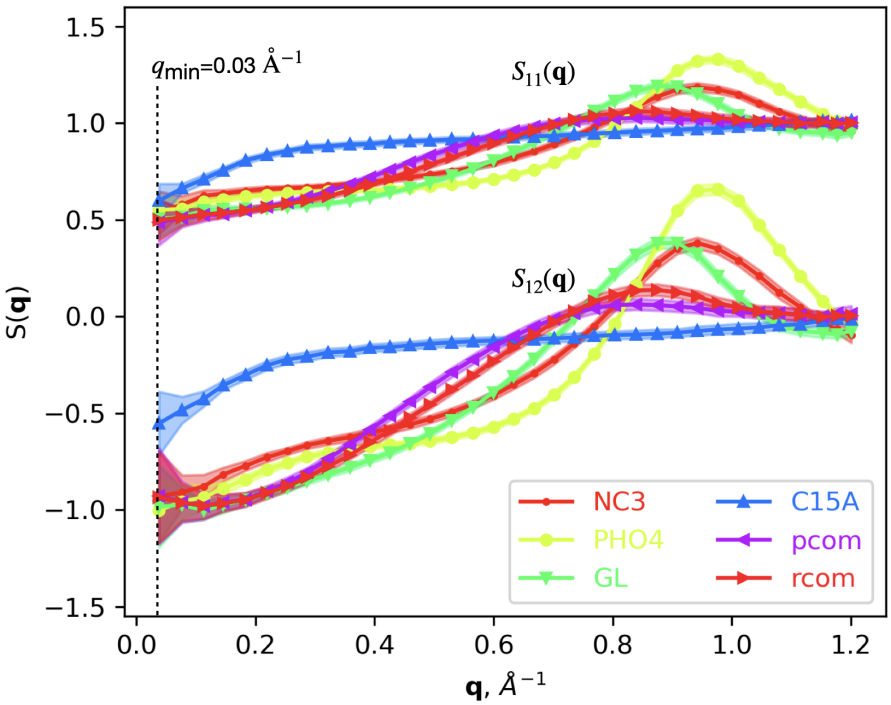
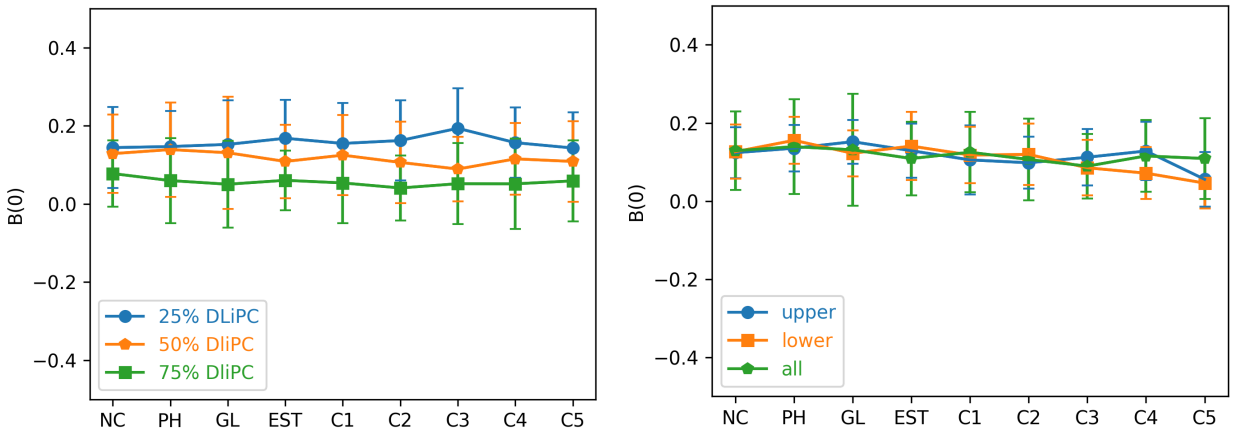


FIGURE 7.3: Facteur de structure pour le système 1:1 de DLiPC/DPPC pour différentes billes. Le regroupement final est indiqué par des lignes, l'intervalle de confiance est en couleur transparente. Un binning est effectué pour clarifier les données. Les composants $S_{11}(q)$ (au-dessus) et S_{12} (en dessous) sont représentés ici.

Dans le régime de faible vecteur d'onde q, nous pouvons constater la présence d'un pré-pic. Des pré-pics similaires ont été observés dans des systèmes bicouches DPPC: les auteurs expliquent cet effet comme étant purement induit par la queue, alors que nous avons également constaté une forte influence des billes du groupe de tête. Notre interprétation est que le pré-pic est dû à un couplage avec des modes d'ondulation et d'inclinaison des lipides indépendants de l'espèce.



(A) (B)

FIGURE 7.4: Paramètre de mélange pour un système de mélange DLiPC et DPPC en fonction des différentes perles pour (A) différentes concentrations de phospholipides and (B) les feuillets. Systèmes de 1024 lipides et 8192 perles d'eau.

Le principal défi de ce travail était de calculer le paramètre de mélange B à partir des relations de Kirwook-Buff (voir eq. 7.5). Dans le cas de mélanges réels de lipides DLiPC et DPPC, une tendance claire de $B(\mathbf{q})$ peut être observée à $q \to 0$ (voir Fig. 7.4). Nous avons observé que le paramètre de mélange B est constant jusqu'à une certaine erreur pour les tailles et concentrations des systèmes étudiés: $B = 0.12 \pm 0.02$. Ceci est en accord avec l'hypothèse théorique faite plus haut que B ne dépend pas de la taille d'un système étudié. Nous pouvons également voir que les calculs pour les feuillets non couplés, la bicouche entière et le centre de masse (CDM) sont cohérents au régime de faible q malgré le bruit pour toutes les tailles de système et les valeurs de concentration. Il y a une préférence des molécules de DLiPC pour se regrouper les unes autour des autres, plutôt qu'un mélange parfait avec le DPPC. Malgré le fait que cet effet soit comparativement faible, il peut être observé et estimé avec précision.

Le même protocole a été appliqué aux systèmes avec du cholestérol. Nous avons étudié des mélanges de DOPE (1,2-dioléoyl-sn-glycéro-3- phosphoethanolamine), DOPC (1,2-dioléoyl-sn-glycéro-3- phosphocholine), SSM (N-stearoyl sphingomyelin) et PSM (palmitoyl sphingomyelin) avec du cholestérol dans des rapports 1:3 et 1:1. Les résultats pour le paramètre de mélange sont présentés dans la Table 7.4. Il a été constaté que ces mélanges

TABLE 7.4: Propriétés structurelles (APL - surface par lipide, nm^2, z - épaisseur, nm, K_A - module d'étirement, mN/m) et paramètre de mélange (B) pour des compositions avec du Cholestérol.

	APL	z	B	K_A^{fluct}	K_A^{str}		
					upper×2	lower×2	bilayer
25% Cholesterol							
DPPC	0.476	4.58	-2.61±0.11	529.6±433.4	273.0±10.8	279.8±8.5	322.1±12.4
DOPC	0.564	4.28	0.29±0.02	352.5±83.4	144.3±5.4	143.3±6.5	235.9±12.7
DLiPC	0.573	4.18	-0.25±0.03	346.1±121.4	153.4±6.4	154.0±6.8	297.5±12.8
DOPE	0.538	4.24	-0.37±0.06	332.1±117.7	148.6±5.2	150.8±5.2	304.5±11.5
SSM	0.438	4.70	-2.52±0.32	1577.9±1542.8	517.1±13.7	496.2±13.0	531.3±10.6
PSM	0.437	4.52	-1.81±0.13	1559.7±4887.5	550.8±15.8	554.0±15.8	592.7±19.8
50% Cholesterol							
DPPC	0.421	4.75	-6.71±0.35	1703.5±2659.4	584.9±9.9	630.9±10.9	625.5±11.7
DOPC	0.469	4.57	-1.68±0.11	510.5±266.7	220.5±6.6	226.1±6.4	331.3±10.1
DLiPC	0.476	4.45	-1.02±0.09	511.2±232.5	231.5±5.2	232.3±5.3	311.6±10.9
DOPE	0.458	4.41	-1.87±0.24	815.1±557.3	237.3±5.8	240.0±5.5	374.0±8.6
SSM	0.414	4.61	-4.69±0.15	2381.6±3306.8	699.3±18.9	748.5±20.4	781.3±19.4
PSM	0.411	4.50	-3.96±0.09	2043.5± 4599.0	829.1±18.9	829.0±18.6	857.7±17.6

$^{\text{fluct}}$ fluctuations de volume, $^{\text{str}}$ facteur de structure a la limite zero

présentent une forte tendance à la pro-mixtion, ce qui est en accord avec les observations expérimentales et la théorie des radeaux lipidiques. Il est également intéressant de noter que le cholestérol induit des phases liquides ordonnées à l'intérieur de la bicouche et préfère se mélanger avec les sphingolipides plutôt qu'avec les phospholipides insaturés, tout en se mélangeant fortement avec le DPPC.

Conclusions

Dans le présent travail, les phospholipides saturés (DPPC), insaturés (DOPC, DOPE, DLiPC) et les sphingolipides (PSM, SSM) ont été étudiés. De plus, des mélanges avec des rapports 1:3, 1:1 et 3:1 de DPPC et DLiPC ont été développés et étudiés. L'effet de la taille du système s'est avéré faible dans les résultats donnés. Nous avons calculé le module d'étirement des systèmes donnés à l'aide de trois méthodes et avons trouvé un accord jusqu'à une erreur dans le module d'étirement et le paramètre de mélange pour les calculs de feuillets couplés et non couplés. Il a été montré que les lipides DLiPC et DPPC ont une légère tendance à la démixtion et que le paramètre de mélange ne dépend pas de la taille du système étudié. En revanche, les mélanges cholestérol-lipides présentent une forte tendance à la pro-mixtion quelle que soit l'espèce lipidique étudiée. Nous avons donc démontré que la thermodynamique du mélange de lipides peut être abordée par des techniques de simulation de dynamique moléculaire et avons introduit une méthode pratique à cet effet.

APPENDIX A

Technical details

Let us discuss here in more detail some technical specifics of the discussed material.

A.1 Fourier transforms notations and conventions

Fourier transforms in unbounded space are noted with a tilde $\tilde{\cdot}$, $\mathbf{r} = (x, y)$, $\mathbf{q} = (q_x, q_y)$, $\mathrm{d}\mathbf{r} = \mathrm{d}x\,\mathrm{d}y$ and $\mathrm{d}\mathbf{q} = \mathrm{d}q_x\,\mathrm{d}q_y$.

$$\begin{aligned} \tilde{z}(\mathbf{q}) &= \int_{-\infty}^{+\infty} \mathrm{d}x \int_{-\infty}^{+\infty} \mathrm{d}y\, z(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}} \\ &= \int \mathrm{d}\mathbf{r}\, z(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}} \\ z(\mathbf{r}) &= \frac{1}{(2\pi)^2} \int \mathrm{d}\mathbf{q}\, \tilde{z}(\mathbf{q}) e^{-i\mathbf{q}\cdot\mathbf{r}} \end{aligned} \tag{A.1}$$

In simulations, one deals with finite boxes, usually combined with periodic boundary conditions. In this work, we restrict ourselves to rectangular boxes with sizes (L_x, L_y, L_z) and area $A = L_x L_y$. We therefore denote $\mathcal{S} = [0, L_x] \times [0, L_y]$ the (x, y) domain associated with a single unit cell. The natural Fourier transform in this case in the Fourier series, denoted with a hat $\hat{\cdot}$. The only admissible wave vectors form a discrete set $q_x = 2\pi n_x / L_x$, $q_y = 2\pi n_y / L_y$, where n_x and n_y span all the possible integer values, negative and positive ($\mathbb{Z}$).

For instance

$$\begin{aligned} \hat{z}_{\mathbf{q}} &= \int_0^{L_x} \mathrm{d}x \int_0^{L_x} \mathrm{d}y\, z(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}} \\ &= \int_{\mathcal{S}} \mathrm{d}\mathbf{r}\, z(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}} \\ z(\mathbf{r}) &= \frac{1}{L_x L_y} \sum_{q_x,q_y} \hat{z}_{\mathbf{q}} e^{-i\mathbf{q}\cdot\mathbf{r}} \\ &= \frac{1}{A} \sum_{\mathbf{q}} \hat{z}_{\mathbf{q}} e^{-i\mathbf{q}\cdot\mathbf{r}} \end{aligned} \tag{A.2}$$

Fourier transform and Fourier series of a real valued functions lead to complex valued transforms with relations $\tilde{z}(-\mathbf{q}) = \tilde{z}(\mathbf{q})^*$, $\hat{z}_{-\mathbf{q}} = \hat{z}^*_{\mathbf{q}}$. In practice, only half of the admissible wavevectors $\mathbf{q}$ needs to be calculated.

We make a systematic use of the Fourier series of the discrete density operator $n(\mathbf{r}) = \sum_j^N \delta(\mathbf{r} - \mathbf{r}_j)$ of N points in a periodic box located at positions $\mathbf{r}_j$. This leads to

$$\begin{aligned} \hat{n}_{\mathbf{q}} &= \sum_{j=1}^{N} \exp(i\mathbf{q}\cdot\mathbf{r}_j) \\ &= \left(\sum_{j=1}^{N} \cos(\mathbf{q}\cdot\mathbf{r}_j)\right) + i\left(\sum_{j=1}^{N} \sin(\mathbf{q}\cdot\mathbf{r}_j)\right) \end{aligned} \tag{A.3}$$

Equilibrium fluctuations of the $\hat{n}_{\mathbf{q}}$ observables are mapped to sampled time averages of molecular dynamics simulations, by means of the ergodic hypothesis. These equilibrium fluctuations provide us with numerical approximations of the structure factors of the fluid, which are then given a thermodynamic interpretation.

A.2 Grand-canonical ensemble

Let us define the grand-canonical partition function [21]

$$\begin{aligned} \Xi([\phi(\mathbf{r}) - \mu] &= \sum_{N=0}^{\infty} \int_{\mathrm{PhaseSpace}} \mathrm{d}\mu_{\mathrm{can}} \exp\{-(E_{\mathrm{tot}} - \mu N)/k_B T\} \\ \mathrm{d}\mu_{\mathrm{can}} &= \frac{\Pi_{i=1}^{N} \mathrm{d}\mathbf{r}_i}{N!\Lambda(1/k_B T)} \end{aligned} \tag{A.4}$$

with T the temperature k_B the Boltzmann constant, $E_{\mathrm{tot}} = U_{\mathrm{int}} + U_{\mathrm{ext}}$ the total energy of the microstate, U_{int} the configurational energy of N particles in the absence of the external field, $\mathrm{d}\mu_{\mathrm{can}}$ is the canonical integration measure, $\Lambda(1/k_B T)$ is the de Broglie thermal wavelength and d is the dimension of the system. It is interesting to note here that in the

absence of external potential ($U_{\text{ext}} = 0$), the grand-canonical potential is directly related to the fluid pressure: $k_B T \ln \Xi = -PV$. Let us apply a mild $U_{\text{ext}} = \frac{1}{k_B T}\phi(\mathbf{r})$. Then the fluid density of the systems changes as follows:

$$\delta(-k_B T \ln \Xi) = -k_B T \frac{\delta \Xi}{\Xi} = \int \mathrm{d}\mathbf{r}\, \rho(\mathbf{r})\phi(\mathbf{r}) \tag{A.5}$$

Considering eqs. 2.35, 2.37, A.4 we can interpret the last formula as a functional derivative

$$\rho(\mathbf{r}) = \langle n(\mathbf{r}) \rangle = k_B T \frac{\delta \ln \Xi}{\delta(\mu - \phi(\mathbf{r}))} \tag{A.6}$$

The second derivative in this case would give us the so-called connected correlation functions. The 1-point correlation function is just the average density $\rho(\mathbf{r})$ 2.37, while the 2-point correlation function is $\rho_2(\mathbf{r}, \mathbf{r}') = \rho(\mathbf{r})\rho(\mathbf{r}')g_2(\mathbf{r}, \mathbf{r}')$:

$$(k_B T)^2 \frac{\delta^2 \ln \Xi}{\delta(\mu - \phi(\mathbf{r}))\, \delta(\mu - \phi(\mathbf{r}'))} = \langle n(\mathbf{r}) n(\mathbf{r}') \rangle - \langle n(\mathbf{r}) \rangle \langle n(\mathbf{r}') \rangle \tag{A.7}$$

$$= \rho(\mathbf{r})\, \delta(\mathbf{r} - \mathbf{r}') + \rho(\mathbf{r})\rho(\mathbf{r}')[g_2(\mathbf{r}, \mathbf{r}') - 1] \tag{A.8}$$

$g_2(\mathbf{r}, \mathbf{r}')$ is a pair correlation function and for homogeneous fluids $g_2(\mathbf{r}, \mathbf{r}') \to g(r = \|\mathbf{r} - \mathbf{r}'\|)$ is the radial distribution function (RDF). Here we can define the direct correlation function $h(r)$ as

$$h(r) = g(r) - 1 \tag{A.9}$$

Interestingly, $h(r) \to 0$ as $r \to \infty$. In the case of an ideal gas the exact expression can be found for the partition function in grand-canonical ensemble. For the 2-dimensional (2d) case it can be expressed as follows:

$$\Xi = \exp\left\{ \Lambda^{-2} \int \mathrm{d}\mathbf{r} \exp\left\{ \frac{1}{k_B T}(\mu - \phi(\mathbf{r})) \right\} \right\} \tag{A.10}$$

leading to the average density of an ideal gas

$$\rho(\mathbf{r}) = \exp\left\{ -\frac{1}{k_B T}\phi(\mathbf{r}) + \frac{1}{k_B T}\mu - \ln \Lambda^{-2} \right\} \tag{A.11}$$

In case of mild potentials $\frac{1}{k_BT}\phi(\mathbf{r}) \ll 1$, we can expand the last equation into

$$\rho(\mathbf{r}) = \rho_0 \exp\left\{-\frac{1}{k_BT}\phi(\mathbf{r})\right\} \simeq \rho_0 - \frac{1}{k_BT}\rho_0\phi(\mathbf{r}) \tag{A.12}$$

where ρ_0 is the grand-canonical density in the absence of the external field. The right hand side of eq. A.12 gives us an explicit form of the linear response contribution to the average fluid density: $\delta\rho = -\frac{1}{k_BT}\rho_0\phi(\mathbf{r})$.

Now let us remind the reader that unfortunately we don't work with an ideal gas here but with a set of interacting particles. To get the generalized expression of eq. A.12 we need to use eq. A.4 and expand it to the first order in $\phi(\mathbf{r})$:

$$\begin{aligned}
\rho(\mathbf{r}) &= \sum_{N=0}^{\infty} \int_{\text{PhaseSpace}} \mathrm{d}\mu_{\text{can}} \exp\left\{-\frac{1}{k_BT}(U_{\text{int}} - \mu N)\right\} \frac{\exp\left\{-\frac{1}{k_BT}\int \mathrm{d}\mathbf{r}' n(\mathbf{r}')\phi(\mathbf{r}')\right\}}{\Xi(\mu - \phi(\mathbf{r}))} \\
&= \sum_{N=0}^{\infty} \int_{\text{PhaseSpace}} \mathrm{d}\mu_{\text{can}} \exp\left\{-\frac{1}{k_BT}(U_{\text{int}} - \mu N)\right\} \frac{1 - \frac{1}{k_BT}\int \mathrm{d}\mathbf{r}' n(\mathbf{r}')\phi(\mathbf{r}') + \mathcal{O}(\phi^2)}{\Xi(\mu - \phi(\mathbf{r}))} \\
\frac{1}{\Xi(\mu - \phi(\mathbf{r}))} &= \frac{1}{\Xi(\mu)} - \frac{\int \mathrm{d}\mathbf{r}' \rho(\mathbf{r}')\phi(\mathbf{r}')}{\Xi(\mu)^2} + \mathcal{O}\left(\phi^2\right)
\end{aligned} \tag{A.13}$$

Finally, we have

$$\begin{aligned}
\rho(\mathbf{r}) &= \rho_0 - \int \mathrm{d}\mathbf{r}' [\langle n(\mathbf{r}) n(\mathbf{r}')\rangle - \langle n(\mathbf{r})\rangle\langle(\mathbf{r}')\rangle] \frac{1}{k_BT}\phi(\mathbf{r}') \\
&= \rho_0 - \rho_0 \frac{1}{k_BT}\phi(\mathbf{r}) - \rho_0^2 \int \mathrm{d}\mathbf{r}' (g_2(\mathbf{r}, \mathbf{r}') - 1) \frac{1}{k_BT}\phi(\mathbf{r}')
\end{aligned} \tag{A.14}$$

To summarize, the above equation shows a non local density change at $\mathbf{r}$ in response to the external field at $\mathbf{r}'$, and assumes a general form:

$$\delta\rho(\mathbf{r}) = \rho(\mathbf{r}) - \rho_0 = \int \mathrm{d}\mathbf{r}' R(\mathbf{r}, \mathbf{r}')\phi(\mathbf{r}') \tag{A.15}$$

where the **susceptibility** $R(\mathbf{r}, \mathbf{r}')$ is defined as

$$
\begin{aligned}
R(\mathbf{r},\mathbf{r}') &= -\frac{1}{k_B T}[\langle n(\mathbf{r})n(\mathbf{r}')\rangle - \langle n(\mathbf{r})\rangle\langle n(\mathbf{r}')\rangle] \\
&= -\frac{1}{k_B T}\rho(\mathbf{r})\,\delta(\mathbf{r}-\mathbf{r}') - \frac{1}{k_B T}\rho(\mathbf{r})\rho(\mathbf{r}')[g_2(\mathbf{r},\mathbf{r}')-1] \\
R(r) &= -\frac{1}{k_B T}\rho_0\,\delta(r) - \frac{1}{k_B T}\rho_0(g(r)-1))
\end{aligned} \tag{A.16}
$$

The latter relation between response and centered fluctuations of the density is typical result from a linear response theory in statistical physics. The first term is associated with ideal gas behavior, while the second term is only present in case of interacting particles. To get expression for $R(r)$ we use the fluid homogeneity when no external field is present. It is important to note here that the linear response susceptibility can only be determined once the density ρ_0 and the pair correlation function $g(r)$ are known.

We note that these expression are also valid for molecular fluids, provided each molecule is identified properly by a position vector $\mathbf{r}_i$. Linear response expression have a general validity.

Similar reasoning leads to a generalization of the above expression to the case of mixtures

$$
\delta\rho_\alpha(\mathbf{r}) = \rho_\alpha(\mathbf{r}) - \rho_{\alpha,0} = \sum_\beta \int \mathrm{d}\mathbf{r}' R_{\alpha\beta}(\mathbf{r},\mathbf{r}')\phi_\beta(\mathbf{r}') \tag{A.17}
$$

where an external potential $\phi_\beta(\mathbf{r})$ is applied exclusively on molecules of type β. The susceptibility is then given by

$$
\begin{aligned}
R_{\alpha\beta}(\mathbf{r},\mathbf{r}') &= -\frac{1}{k_B T}[\langle n_\alpha(\mathbf{r})n_\beta(\mathbf{r}')\rangle - \langle n_\alpha(\mathbf{r})\rangle\langle n_\beta(\mathbf{r}')\rangle] \\
&= -\frac{1}{k_B T}\left[\rho_\alpha(\mathbf{r})\delta_{\alpha\beta}\,\delta(\mathbf{r}-\mathbf{r}') + \rho_\alpha(\mathbf{r})\rho_\beta(\mathbf{r}')[g_{2,\alpha\beta}(\mathbf{r},\mathbf{r}')-1]\right] \\
R_{\alpha\beta}(r) &= -\frac{1}{k_B T}\left[\rho_{0,\alpha}\delta_{\alpha\beta}\,\delta(r) + \rho_{0,\alpha}\rho_{0,\beta}(g_{\alpha\beta}(r)-1))\right]
\end{aligned} \tag{A.18}
$$

The Fourier transform of the susceptibility is then

$$
\tilde{R}_{\alpha\beta}(q) = \frac{1}{k_B T}\left[\rho_{0,\alpha}\delta_{\alpha\beta} + \rho_{0,\alpha}\rho_{0,\beta}\tilde{h}_{\alpha\beta}(q)\right] \tag{A.19}
$$

which show directly the connection with the density correlations $\tilde{h}_{\alpha\beta}(q)$.

A.3 Replicas

As was discussed in Chapter 2, p-component bilayers can be treated as $2p$ component fluids. We remind the notation $u, l = 1, 2$ for the leaflet indices, $\alpha, \beta = 1 \dots p$ for the component indices and $a, b = 1 \dots 2p$ for the extended number of component indices.

Bilayer mixtures can be treated in the grand-canonical or canonical ensembles, and the general relation

$$\sum_{c=1}^{2p} \frac{\partial N_a}{\partial \mu_c} \frac{\partial \mu_c}{\partial N_b} = \delta_{ab} \tag{A.20}$$

holds. The derivatives $\partial N_a / \partial \mu_c$ correspond to the grand-canonical ensemble where particles numbers are function of chemical potentials, area A and temperature T. They are in particular related to the KB integrals as follows:

$$\frac{\partial N_a}{\partial \mu_b} = \frac{1}{k_B T} \left(N_a \delta_{ab} + \frac{N_a N_b}{A} G_{ab} \right) \tag{A.21}$$

The derivatives $\partial \mu_c / \partial N_b$ correspond to the canonical ensemble. It is there useful to introduces the intensive densities $\rho_a = N_a / A$ and Helmholtz free-energy $\mathcal{F}(N_a, A, T) = A f(\rho_a, T)$.

$$\frac{\partial \mu_a}{\partial N_b} = \left. \frac{\partial^2 \mathcal{F}}{\partial N_a \partial N_b} \right|_{A,T} = \frac{1}{A} \frac{\partial^2 f}{\partial \rho_a \partial \rho_b} \tag{A.22}$$

With these new relations

$$\begin{aligned} \sum_{c=1}^{2p} \frac{\partial N_a}{\partial \mu_c} \frac{\partial \mu_c}{\partial N_b} &= \sum_{c=1}^{2p} \frac{1}{k_B T} \left(N_a \delta_{ac} + \frac{N_a N_c}{A} G_{ac} \right) \frac{1}{A} \frac{\partial^2 f}{\partial \rho_c \partial \rho_b} \\ &= \frac{1}{k_B T} \sum_{c=1}^{2p} \left(\rho_a \delta_{ac} + \rho_a \rho_c G_{ac} \right) \frac{\partial^2 f}{\partial \rho_c \partial \rho_b} \\ &= \delta_{ab} \end{aligned} \tag{A.23}$$

Therefore, we have an inverse matrix connection between the KB terms G_{ab} and the second derivative of f. One can show that all the interesting thermodynamic quantities can be expressed in terms of $\partial^2 f / \partial \rho_a \partial_b$ as follows:

$$\frac{1}{\chi_T} = \sum_{a,b=1}^{2p} \rho_a \rho_b \frac{\partial^2 f}{\partial \rho_a \partial \rho_b} \tag{A.24}$$

$$\mathcal{A}_a = \chi_T \sum_{b=1}^{2p} \rho_b \frac{\partial^2 f}{\partial \rho_a \partial \rho_b} \tag{A.25}$$

$$\mu_{ab} = \frac{1}{A}\left[\frac{\partial^2 f}{\partial \rho_a \partial \rho_b} - \frac{\mathcal{A}_a \mathcal{A}_b}{\chi_T}\right] \tag{A.26}$$

with $\mathcal{A}_a$ the partial molar area of species a and μ_{ab} the derivative $\partial \mu_a / \partial N_b$ at constant tension σ and temperature T. Combining A.23 and A.26 makes it possible, in principle, to express all thermodynamic quantities in terms of the KB integrals G_{ab}.

Let us apply now this formalism to a pure bilayer, $p = 1$ and $a = 1, 2$ depending on the leaflet. One has $G_{11} = G_{22}$ the integrated pair correlation function and $G_{12} = G_{21} = G'$ the integrated crossed correlation function. In the absence of leaflet interaction $G' = 0$; Eq. A.23 reduces to

$$\begin{aligned}\begin{pmatrix} \frac{\partial^2 f}{\partial \rho_1^2} & \frac{\partial^2 f}{\partial \rho_1 \partial \rho_2} \\ \frac{\partial^2 f}{\partial \rho_2 \partial \rho_1} & \frac{\partial^2 f}{\partial \rho_2^2} \end{pmatrix} &= k_B T \begin{pmatrix} \rho_m + \rho_m^2 G_{11} & 0 \\ 0 & \rho_m + \rho_m^2 G_{11} \end{pmatrix}^{-1} \\ &= \begin{pmatrix} \frac{k_B T}{\rho_m + \rho_m^2 G_{11}} & 0 \\ 0 & \frac{k_B T}{\rho_m + \rho_m^2 G_{11}} \end{pmatrix}\end{aligned} \tag{A.27}$$

with ρ_m the monolayer density. One deduces immediately

$$\frac{1}{\chi_T} = 2\rho_m^2 \frac{k_B T}{\rho_m + \rho_m^2 G_{11}} = 2\frac{\rho_m k_B T}{1 + \rho_m G_{11}} \tag{A.28}$$

$$\chi_T = \frac{1}{2}\left[\frac{1 + \rho_m G_{11}}{\rho_m k_B T}\right] \tag{A.29}$$

Therefore the compressibility of the bilayer is half of the compressibility of the monolayer $(1 + \rho_m G_{11})/(\rho_m k_B T)$.

A.4 Linear response for binary mixtures

Let us assume a system of binary mixture of composition x_1, x_2 and add two external potentials ϕ_1, ϕ_2 acting respectively on species 1 and 2.

An interesting choice is

$$\left\{ \begin{array}{rcc} \phi_1 & = & U_0 \cos(\mathbf{q}\cdot\mathbf{r})/x_1 \\ \phi_2 & = & 0 \end{array} \right. \tag{A.30}$$

leading to a density modulation of amplitude $\rho_1(\mathbf{r}) = \rho_{1,0} + \rho_{1,1}\cos(\mathbf{q}\cdot\mathbf{r})$, $\rho_2(\mathbf{r}) = \rho_{2,0} + \rho_{2,1}\cos(\mathbf{q}\cdot\mathbf{r})$, $\rho_0 = \rho_{0,1} + \rho_{0,2}$.

$$\left(\begin{array}{c} \rho_{1,1} \\ \rho_{2,1} \end{array} \right) = -\rho_0 \frac{U_0}{k_B T} \left(\begin{array}{c} 1 + \rho_0 x_1 \tilde{h}_{11} \\ \rho x_2 \tilde{h}_{12} \end{array} \right) \tag{A.31}$$

and

$$\left\{ \begin{array}{rcc} \phi_1' & = & 0 \\ \phi_2' & = & U_0 \cos(\mathbf{q}\cdot\mathbf{r})/x_2 \end{array} \right. \tag{A.32}$$

leading to $\rho_1(\mathbf{r}) = \rho_{1,0} + \rho'_{1,1}\cos(\mathbf{q}\cdot\mathbf{r})$, $\rho_2(\mathbf{r}) = \rho_{2,0} + \rho'_{2,1}\cos(\mathbf{q}\cdot\mathbf{r})$

$$\left(\begin{array}{c} \rho'_{1,1} \\ \rho'_{2,1} \end{array} \right) = -\rho_0 \frac{U_0}{k_B T} \left(\begin{array}{c} \rho_0 x_1 \tilde{h}_{12} \\ 1 + \rho x_2 \tilde{h}_{22} \end{array} \right). \tag{A.33}$$

It is then possible to express the coefficients ζ and η (2.25) as

$$\zeta = (k_B T)^2 \frac{\rho_{1,1}\rho'_{2,1} - \rho'_{1,1}\rho_{2,1}}{\rho_0^2 U_0^2} \tag{A.34}$$

$$\eta = k_B T \frac{x_1\rho_{2,1} + x_2\rho'_{1,1} - x_2\rho_{1,1} - x_1\rho'_{2,1}}{U_0}. \tag{A.35}$$

A.5 Director and membrane fluctuations

A.5.1 Helfrich hamiltonian

Let $z(\mathbf{r})$ be the elevation of a membrane over the horizontal plane in a periodic box of size L_x, L_y. The curvature Helfrich hamiltonian reads

$$\mathcal{H} = \int_0^{L_x} \mathrm{d}x \int_0^{L_y} \mathrm{d}y \left\{ \frac{\kappa}{2} (\Delta z)^2 \right\}. \tag{A.36}$$

Inserting the Fourier expansion $\hat{z}_{\mathbf{q}}$ gives:

$$\begin{aligned} \mathcal{H} &= \frac{1}{(L_x L_y)^2} \int_{\mathcal{S}} \mathrm{d}\mathbf{r} \sum_{\mathbf{q}} \sum_{\mathbf{q}'} \frac{\kappa}{2} (-q^2) \hat{z}_{\mathbf{q}} (-q'^2) \hat{z}_{\mathbf{q}'} e^{-i\mathbf{q}\cdot\mathbf{r} - i\mathbf{q}'\cdot\mathbf{r}} \\ &= \frac{1}{L_x L_y} \sum_{\mathbf{q}} \left[\frac{\kappa}{2} q^4 \right] \hat{z}_{\mathbf{q}} \hat{z}_{-\mathbf{q}}. \end{aligned} \tag{A.37}$$

As $z(\mathbf{r})$ is real, so $\hat{z}_{-\mathbf{q}} = \hat{z}^*_{\mathbf{q}}$. Modes $\mathbf{q}$ et $-\mathbf{q}$ are redundant. The real $\mathrm{Re}(\hat{z}_{\mathbf{q}})$ and imaginary $\mathrm{Im}(\hat{z}_{\mathbf{q}})$ parts of the Fourier modes $\hat{z}_{\mathbf{q}}$ fluctuate independently.

$$\mathcal{H} = \frac{1}{L_x L_y} \sum_{\mathbf{q}} \left[\frac{\kappa}{2} q^4\right] \left(\mathrm{Re}(\hat{z}_{\mathbf{q}})^2 + \mathrm{Im}(\hat{z}_{\mathbf{q}})^2\right). \tag{A.38}$$

The zero mode $\hat{z}_{\mathbf{0}}$ must be treated separately. In MD it is quasi-constant because the center of mass of the simulation is fixed and water permeation is very low. One can set $\hat{z}_{\mathbf{0}} = 0$ and exclude this mode from the sum $\sum_{\mathbf{q}}$.

It is convenient to define a subset $\mathcal{Q}$ of non-redundant independent modes

$$\mathcal{Q} = \left\{(q_x, q_y), q_x > 0 \,\text{or}\, (q_x = 0 \,\text{and}\, q_y > 0)\right\}. \tag{A.39}$$

The Helfrich partition function is

$$\mathcal{Z} = \int \prod_{\mathbf{q}\in\mathcal{Q}} \mathrm{dRe}(\hat{z}_{\mathbf{q}})\, \mathrm{dIm}(\hat{z}_{\mathbf{q}})\, \exp\left(-\frac{1}{k_B T}\mathcal{H}\right). \tag{A.40}$$

A.5.2 Energy equipartition, quadratic fluctuations

The canonical average $\langle \hat{z}_{\mathbf{q}} \hat{z}_{-\mathbf{q}} \rangle$ can be obtained from the following trick

$$\begin{aligned}
0 &= \int \prod_{\mathbf{q}'\in\mathcal{Q}} \mathrm{dRe}(\hat{z}_{\mathbf{q}'})\, \mathrm{dIm}(\hat{z}_{\mathbf{q}'})\, \frac{\partial}{\partial \mathrm{Re}(\hat{z}_{\mathbf{q}})} \left\{ \mathrm{Re}(\hat{z}_{\mathbf{q}}) \exp\left(-\frac{1}{k_B T}\mathcal{H}\right)\right\} \\
&= \int \prod_{\mathbf{q}'\in\mathcal{Q}} \mathrm{dRe}(\hat{z}_{\mathbf{q}'})\, \mathrm{dIm}(\hat{z}_{\mathbf{q}'}) \left(1 - \frac{1}{k_B T} \mathrm{Re}(\hat{z}_{\mathbf{q}}) \frac{\partial \mathcal{H}}{\partial \mathrm{Re}(\hat{z}_{\mathbf{q}})}\right) \exp\left(-\frac{1}{k_B T}\mathcal{H}\right) \\
&= \mathcal{Z}\left(1 - \frac{1}{k_B T}\left\langle \mathrm{Re}(\hat{z}_{\mathbf{q}}) \frac{\partial \mathcal{H}}{\partial \mathrm{Re}(\hat{z}_{\mathbf{q}})} \right\rangle\right).
\end{aligned} \tag{A.41}$$

Because in the sum (A.38) $\mathrm{Re}(\hat{z}_{\mathbf{q}})$ appears twice, one for $\mathbf{q}$, the other for $-\mathbf{q}$, one obtains:

$$\left\langle \mathrm{Re}(\hat{z}_{\mathbf{q}}) \frac{\partial \mathcal{H}}{\partial \mathrm{Re}(\hat{z}_{\mathbf{q}})} \right\rangle = \frac{2}{L_x L_y} \times \left\langle \mathrm{Re}(\hat{z}_{\mathbf{q}}) \left[\frac{\kappa q^4}{2}\right] 2\mathrm{Re}(\hat{z}_{\mathbf{q}}) \right\rangle = k_B T \tag{A.42}$$

and with the same reasoning for $\mathrm{Im}(\hat{z}_{\mathbf{q}})$,

$$\left\langle \mathrm{Re}(\hat{z}_{\mathbf{q}})^2 \right\rangle = \left\langle \mathrm{Im}(\hat{z}_{\mathbf{q}})^2 \right\rangle = \frac{k_B T L_x L_y}{2} \frac{1}{\kappa q^4} \tag{A.43}$$

leading to the desired result

$$\langle \hat{z}_{\mathbf{q}} \hat{z}_{-\mathbf{q}} \rangle = \frac{k_B T L_x L_y}{\kappa q^4}. \tag{A.44}$$

A.5.3 Tilt fluctuations

The calculation of the fluctuations of the director follows the same path. The Hamiltonian is now

$$\begin{aligned}\mathcal{H} &= \int_0^{L_x} \mathrm{d}x \int_0^{L_y} \mathrm{d}y \left\{ \frac{K_0}{2}\mathbf{M}_\parallel^2 + \frac{K_1}{2}\left[\left(\frac{\partial \mathbf{M}_{\parallel,x}}{\partial x}\right)^2 + \left(\frac{\partial \mathbf{M}_{\parallel,x}}{\partial y}\right)^2 + \left(\frac{\partial \mathbf{M}_{\parallel,y}}{\partial x}\right)^2 + \left(\frac{\partial \mathbf{M}_{\parallel,y}}{\partial y}\right)^2 \right] \right\} \\ &= \frac{1}{L_x L_y} \sum_{\mathbf{q}} \frac{1}{2}[K_0 + K_1 q^2] \hat{\mathbf{M}}_{\parallel,\mathbf{q}} \hat{\mathbf{M}}_{\parallel,-\mathbf{q}} \end{aligned} \tag{A.45}$$

The result is then

$$\left\langle \hat{\mathbf{M}}_{\parallel,\mathbf{q}} \hat{\mathbf{M}}_{\parallel,-\mathbf{q}} \right\rangle = \frac{k_B T L_x L_y}{K_0 + K_1 q^2} \tag{A.46}$$

APPENDIX B

Spica force field parameters

A full list of the available parameters for the SPICA force field can be found at the SPICA webpage and a paper by Wataru Shinoda, Russell DeVane and Michael Klein [69].

TABLE B.1: SPICA force field CG mapping to the all-atom structure.

CG-segment	all-atom	mass, g/mol
W	$(H_2O)_3$	54.046
CT	$CH_3CH_2CH_2$-	43.088
CM	-$CH_2CH_2CH_2$-	42.080
CT2	CH_3CH_2-	29.061
OA	$HOCH_2$-	30.026
NC	-CH_2CH_2-N-$(CH_3)_3$(+1)	87.164
NH	-CH_2CH_2-NH_3(+1)	45.084
PH	-PO_4-(-1)	94.971
GL	-CH_2CH-CH_2-	41.072
EST1	-CH_2CO_2- (in the sn-2 chain)	58.035
EST2	-H_2-CO_2- (in the sn-1 chain)	58.035
CMD2	-HC=CH- (cis)	26.037
CM2	-CH_2CH_2-	28.054
C2T	$(CH_3)_2CH$-	43.089

TABLE B.2: SPICA force field bond stretching parameters.

bond pair		k_b, kcal/mol/ Å^2	r_0, Å
NC	PH	4.80	4.25
PH	CM	12.00	3.69
PH	GL	8.90	3.52
GL	EST1	30.00	2.88
GL	EST2	8.40	3.48
EST1	CM	4.70	3.55
EST2	CM	5.10	3.61
NH	PH	9.40	3.60
GL	OA	70.00	2.29
GL	PEP	89.00	2.29
CM	OA	27.5	3.05
CM	PEP	89.00	2.29
CM	CM	6.16	3.64
CM	CT	6.16	3.65
CT2	CM	9.00	3.13
CMD2	CM	8.00	3.09
CT	CMD2	8.00	3.09

TABLE B.3: SPICA force field angle bending parameters.

angle pair			k_θ, kcal/mol/rad^2	θ_0, degree
NC	PH	CM	3.300	112.0
PH	CM	CM	1.100	178.0
NC	PH	GL	3.100	112.0
PH	GL	EST1	1.400	124.0
PH	GL	EST2	2.000	138.0
GL	EST1	CM	0.800	168.0
GL	EST2	CM	0.800	172.0
OA	GL	PH	25.600	95.5
CM	OA	GL	35.500	107.5
PEP	GL	PH	3.500	145.5
OA	GL	PEP	70.000	124.1
CM	PEP	GL	7.000	166.0
CM	CM	PEP	2.000	147.5
EST1	GL	EST2	1.000	95.0
EST1	CM	CM	1.000	178.0
EST2	CM	CM	1.000	178.0
NH	PH	GL	4.000	102.0
CM	CM	CM	1.190	173.0
CM	CM	CT	1.190	175.0
CT2	CM	CM	1.600	172.0
CT2	CM	CT	1.600	172.0
CT2	CM	CT2	1.700	173.0
CT	CM	CT	1.093	175.5
CT	CMD2	CT	7.700	116.0
CT2	CMD2	CT2	12.000	110.0
CMD2	CM	CM	1.900	161.0
CM	CMD2	CM	6.000	110.0

TABLE B.4: SPICA force field non bonded interaction parameters.

pair		LJ-type	ϵ, kcal/mol	σ, Å	pair		LJ-type	ϵ, kcal/mol	σ, Å
PH	CM	LJ9-6	0.300	4.9530	NH	CMD2	LJ9-6	0.300	4.3025
PH	CT	LJ9-6	0.320	4.9925	NH	CT2	LJ9-6	0.320	4.4105
NC	CM	LJ9-6	0.400	5.1280	NH	CT	LJ9-6	0.340	4.5925
NC	CT	LJ9-6	0.420	5.1675	NH	NH	LJ9-6	1.100	4.6000
PH	W	LJ12-4	1.000	4.0300	NH	PHE	LJ9-6	1.200	3.8000
NC	W	LJ12-4	0.900	4.6100	NH	GL	LJ9-6	0.750	4.1900
PH	PH	LJ9-6	1.400	5.4000	NH	EST1	LJ9-6	0.850	4.1100
NC	NC	LJ9-6	0.700	5.7500	NH	EST2	LJ9-6	0.850	4.1100
NC	PH	LJ9-6	1.150	4.2000	NH	W	LJ12-4	0.800	3.9500
GL	GL	LJ9-6	0.420	4.5060	PH	PHE	LJ9-6	1.400	5.0000
GL	EST1	LJ9-6	0.470	4.4030	NH	PH	LJ9-6	1.200	3.8000
GL	EST2	LJ9-6	0.470	4.4030	NC	PHE	LJ9-6	1.150	4.2000
GL	CM	LJ9-6	0.420	4.5060	NC	NH	LJ9-6	0.880	5.1750
GL	CT	LJ9-6	0.444	4.5455	CT	CT	LJ9-6	0.469	4.5850
GL	W	LJ12-4	0.640	4.4385	CT	CM	LJ9-6	0.444	4.5455
GL	CT2	LJ9-6	0.362	4.3635	CM	CM	LJ9-6	0.420	4.5060
PH	CT2	LJ9-6	0.280	4.8105	CMD2	CMD2	LJ9-6	0.232	4.0050
NC	CT2	LJ9-6	0.320	4.9855	CMD2	CT	LJ9-6	0.330	4.2950
CMD2	GL	LJ9-6	0.312	4.2555	CMD2	CT2	LJ9-6	0.269	4.1130
CMD2	EST1	LJ9-6	0.440	4.0050	CMD2	CM	LJ9-6	0.312	4.2555
CMD2	EST2	LJ9-6	0.440	4.0050	CMD2	W	LJ12-4	0.270	4.1880
PH	CMD2	LJ9-6	0.300	4.7025	EST1	CM	LJ9-6	0.470	4.4030
NC	CMD2	LJ9-6	0.350	4.8775	EST2	CM	LJ9-6	0.470	4.4030
GL	NC	LJ9-6	0.650	4.6200	EST1	CT	LJ9-6	0.470	4.4425
GL	PH	LJ9-6	0.300	4.7500	EST2	CT	LJ9-6	0.470	4.4425
EST1	NC	LJ9-6	0.750	4.4750	EST1	CT2	LJ9-6	0.390	4.2605
EST2	NC	LJ9-6	0.750	4.4750	EST2	CT2	LJ9-6	0.390	4.2605
EST1	PH	LJ9-6	0.500	4.5500	EST1	EST1	LJ9-6	0.495	4.3000
EST2	PH	LJ9-6	0.500	4.5500	EST1	EST2	LJ9-6	0.495	4.3000
PHE	PHE	LJ9-6	1.400	4.6000	EST2	EST2	LJ9-6	0.495	4.3000
PHE	CM	LJ9-6	0.300	4.9530	EST1	W	LJ12-4	0.820	4.2900
PHE	CMD2	LJ9-6	0.300	4.7025	EST2	W	LJ12-4	0.820	4.2900
PHE	CT	LJ9-6	0.320	4.9925	W	CT	LJ12-4	0.360	4.4780
PHE	CT2	LJ9-6	0.280	4.8105	W	CM	LJ12-4	0.340	4.4385
PHE	GL	LJ9-6	0.300	4.7500	CT2	CT2	LJ9-6	0.312	4.2210
PHE	EST1	LJ9-6	0.500	4.5500	CT2	CM	LJ9-6	0.362	4.3635
PHE	EST2	LJ9-6	0.500	4.5500	W	CT2	LJ12-4	0.290	4.2960
PHE	W	LJ12-4	1.000	4.0300	CT2	CT	LJ9-6	0.383	4.4030
NH	CM	LJ9-6	0.330	4.5530	W	W	LJ12-4	0.895	4.3710
NC	OA	LJ9-6	0.280	3.7852	NC	PEP	LJ9-6	0.546	4.2075
PH	OA	LJ9-6	0.317	4.3287	PH	PEP	LJ9-6	0.540	4.2975
GL	OA	LJ9-6	0.700	4.2020	GL	PEP	LJ9-6	0.414	4.3280
OA	OA	LJ9-6	0.718	4.0843	OA	PEP	LJ9-6	0.503	3.9315
PEP	PEP	LJ9-6	1.400	4.1500	CT	PEP	LJ9-6	0.3915	4.3675
OA	W	LJ12-4	0.380	4.1475	PEP	W	LJ12-4	0.600	4.2605

APPENDIX C

Structural properties of the model lipids

C.1 Number density of the pure lipid bilayers

The number density was calculated as follows:

$$\rho_N = \frac{N_{\text{particles}}}{V_{\text{slab}}} \cdot V_{\text{particle}} \tag{C.1}$$

where $N_{\text{particles}}$ is the total number of reference beads, $V_{\text{slab}} = L_x L_y dz$ is the volume of the slab used for the density calculation, $dz = L_z / n_{\text{bin}}$, is the slab thickness, n_{bin} is the number of bins along z axis, $V_{\text{particle}} = \pi D^3/6$ is the volume of the reference bead, D is the diameter of the reference bead as per the LJ potential (see Appendix B Tab. B.4).

In Fig. C.1 on can see the density profiles of the pure phospholipid bilayers of 512 lipids and 8192 water beads at 298 K and 1 atm. The number density for studied sphingolipids is shown in Fig. C.2 for systems of 512 lipids and 8192 water beads at 318 K and 1 atm. Peaks correspond to the density of the beads along z axis and so parallel to the bilayer normal.

One can see from Figs. C.1 and C.2 that all studied systems show water propagation into the bilayer. As well that, one broad peak can be seen for the last tail bead in all cases except PSM. This can show that in most cases there is a coupling between different leaflets. On the other hand, this only shows that there is no gap between the monolayers

which is generally logical. Thus, such profiles could help approximate the thickness of the studied systems and show structuring of the lipid along z direction.

A more pronounced shape for the sphingomyelin is assumed to be a consequence of its special structure. Since these molecules have higher melting temperature and thus, are more rigid in their nature, the peaks on Fig. C.2 show that the position of head and tail groups for such lipids is more restricted. As well as that one can see that saturated lipid DPPC and shpingolipids have higher thickness, which can be a consequence of the rigidity of these molecules. Double bonds allow unsaturated lipids to form kinks in the tails and so a cross-leaflet permeation is more likely to occur.

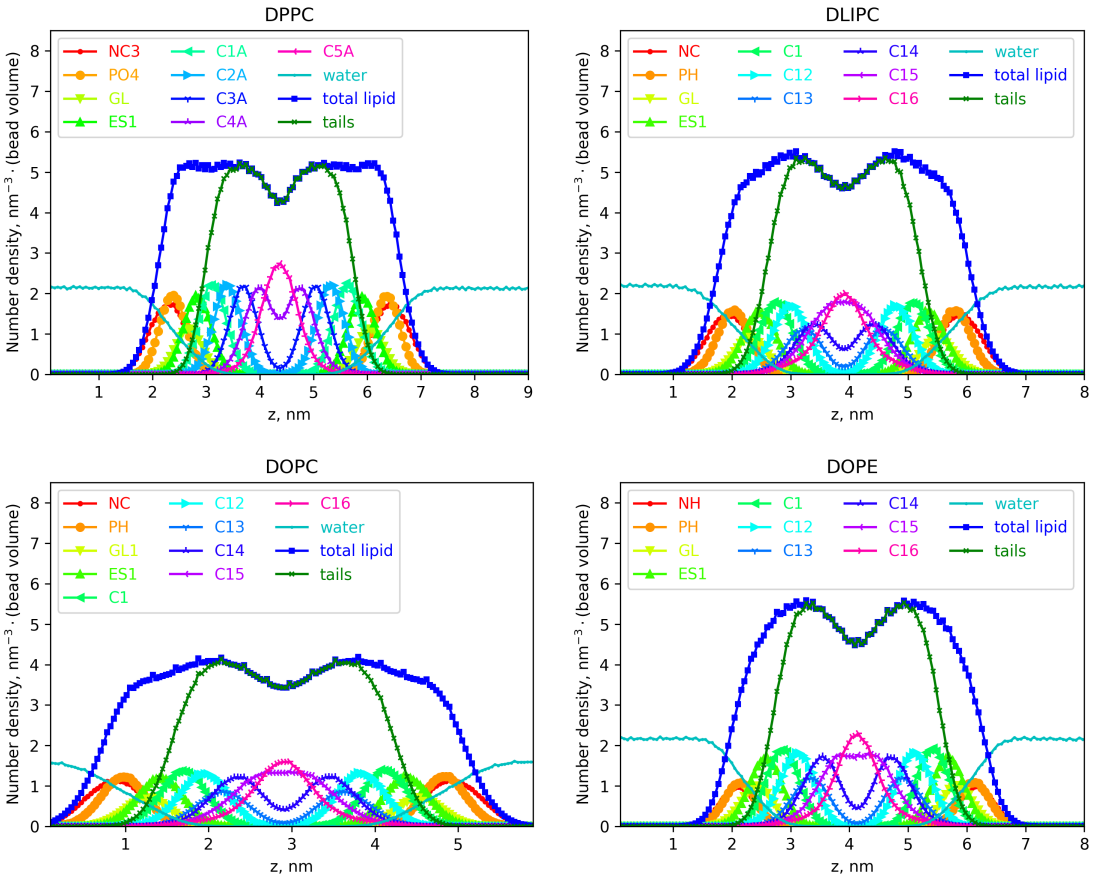


FIGURE C.1: The number density profiles of the model lipids: DPPC, DLiPC, DOPC, DOPE, in pure bilayer systems, for all available beads in the molecule, whole lipid and carbon tails.

C.2 Structure factor of the pure model bilayers

Figures in this section show the structure factor for pure bilayer systems of model phospholipids: DPPC, DLiPC, DOPC, DOPE C.3, and sphingolipids: SSM and PSM C.4. All systems have 512 lipid and 8192 molecules and are equilibrated at 1 atm. Phospholipid

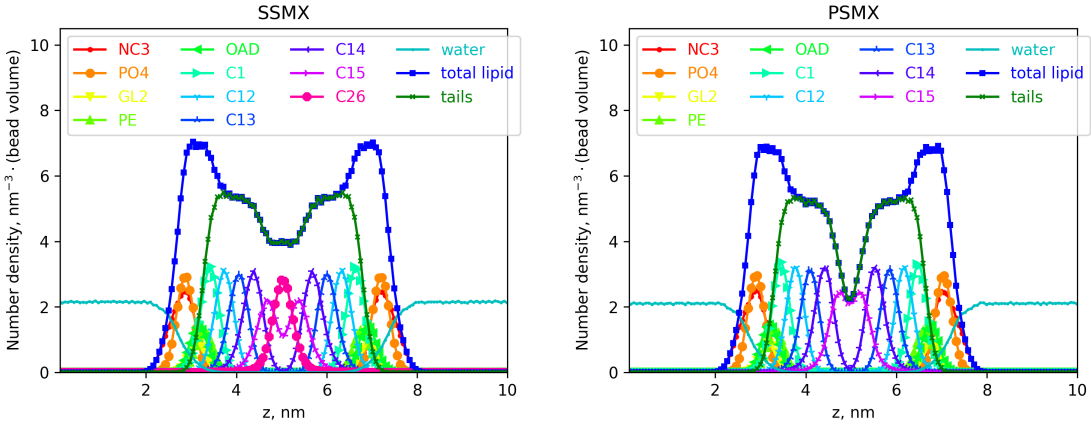


FIGURE C.2: The number density profiles of the model lipids: SSM, PSM, in pure bilayer systems, for all available beads in the molecule, whole lipid and carbon tails.

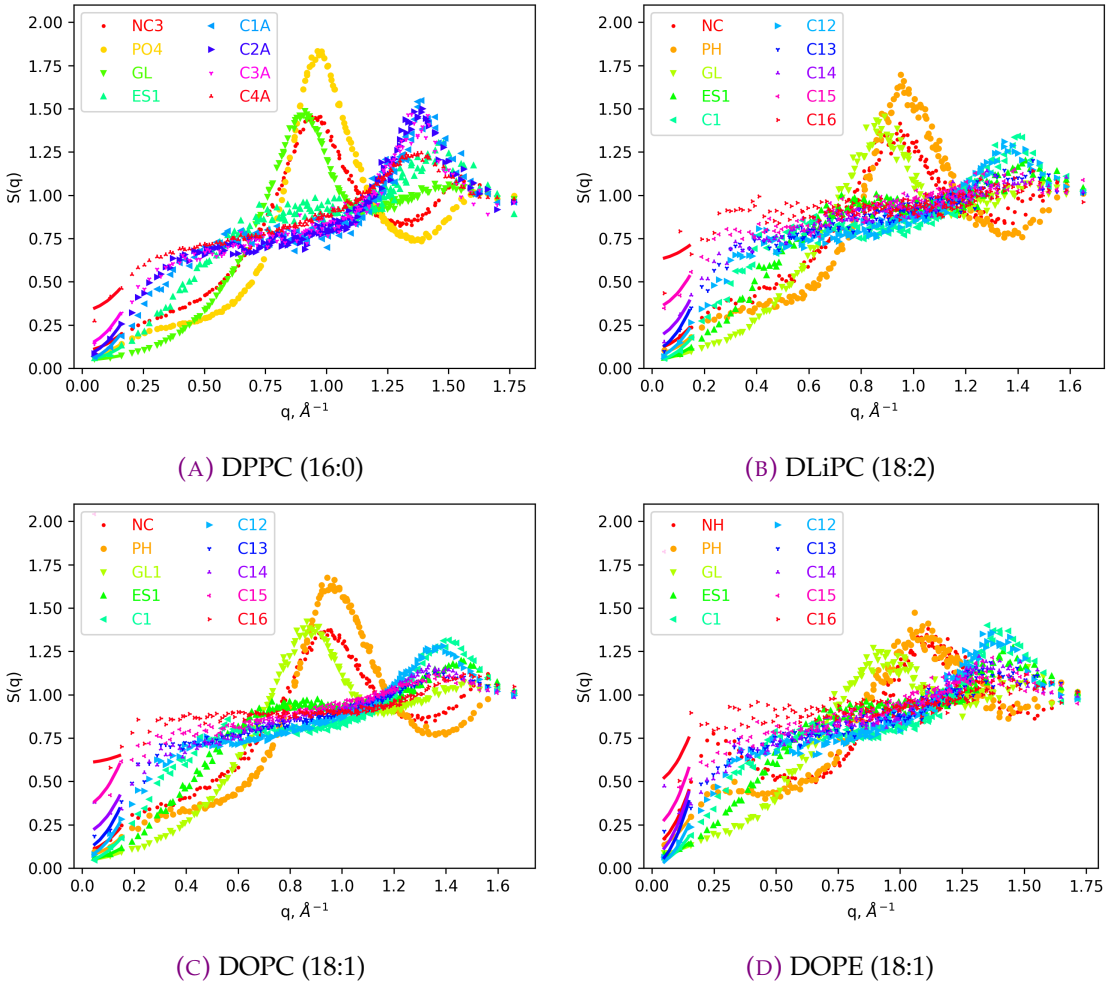
bilayers are shown at 298 K, while sphingomyelin is at 318 K. Data point are shown by points that were cleared for better visualization aspects. By that here we mean that only each 10th data point is shown. Thin lines correspond to the approximation of the structure factor made to obtain zero-values of $S(\mathbf{q})$.

One can see how uncertain are the resulting curves for unsaturated phospholipids. This can be because they are less restricted in the space exploration due to the double bonds in the tails. That means that various kinks can be formed and lipid could form more fluid-like structure. As can be concluded, rigid lipids (DPPC, SSM, PSM) show much better consistency in the zero-limit of $\mathbf{q}$ than unsaturated lipids due to the fact that they experience less fluctuations.

C.3 Structure factor of the phospholipid mixtures

This section features the structure factor of phospholipid mixtures. One can see $S(\mathbf{q})$ for the mixture of DPPC and DLiPC systems of two sizes: 512 and 1024 lipids. Both systems were contained in 8192 SPICA water solvent which gives enough hydration for two cases. Three concentrations were studied: 25% DPPC 75% DLiPC, 50% DPPC 50% DLiPC and 75% DPPC 25% DLiPC. Systems were kept at 298 K and 1 atm for 10 μs run with fixed volume. Figures show partial structure factors as was defined in Chapter 2. One can see S_{11} component that corresponds to the first lipid in pair, S_{22} that corresponds to the second lipid in pair and S_{12} that corresponds to their cross-term.

Comparison of two systems sizes gives us a better estimate of the desired property and also can point out to some size-dependent effects. In these two specific systems we did not observe any significant finite size effects. In Figures here one can see how the structure



(A) DPPC (16:0)

(B) DLiPC (18:2)

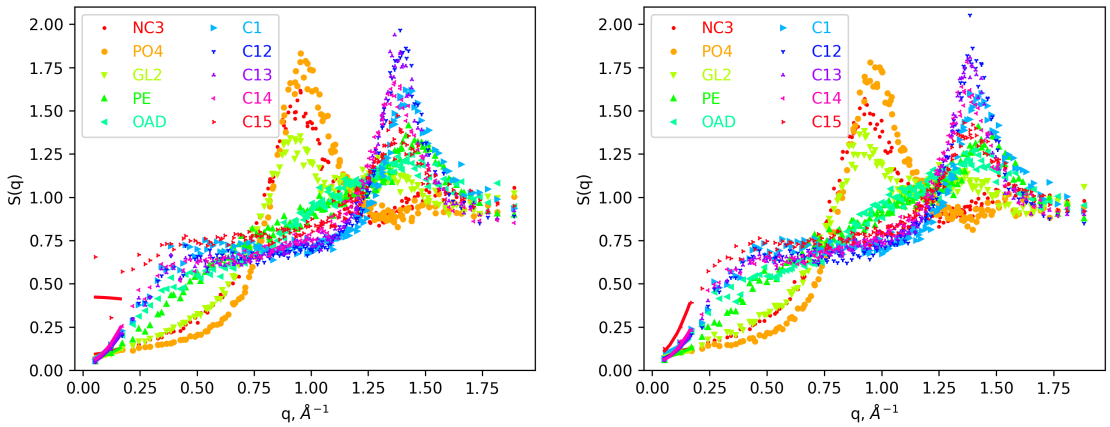
(C) DOPC (18:1)

(D) DOPE (18:1)

FIGURE C.3: A full structure factor dependence on the bead type for available model lipids: DPPC, DLiPC, DOPC, DOPE. Systems of 512 lipids and 8192 water beads at T=298 K and P=1 atm.

factor is dependent on the box size. It is clear that lower $\mathbf{q}$ values can be achieved but on the other hand the short-range information is lost. From Figures in this section one can see that the peaks remain at the same wave-length position for all studied cases. At the same time, the zero limit value of the $S(\mathbf{q})$ can slightly change due to the different number of approximation points and the fact that the bigger system should a priori be closer to the desire zero $\mathbf{q}$. One can see side by side comparison of the obtained structure factors for studied systems of different size in Figs. C.5 (25% DPPC), C.6 (50% DPPC) and C.7 (75% DPPC).

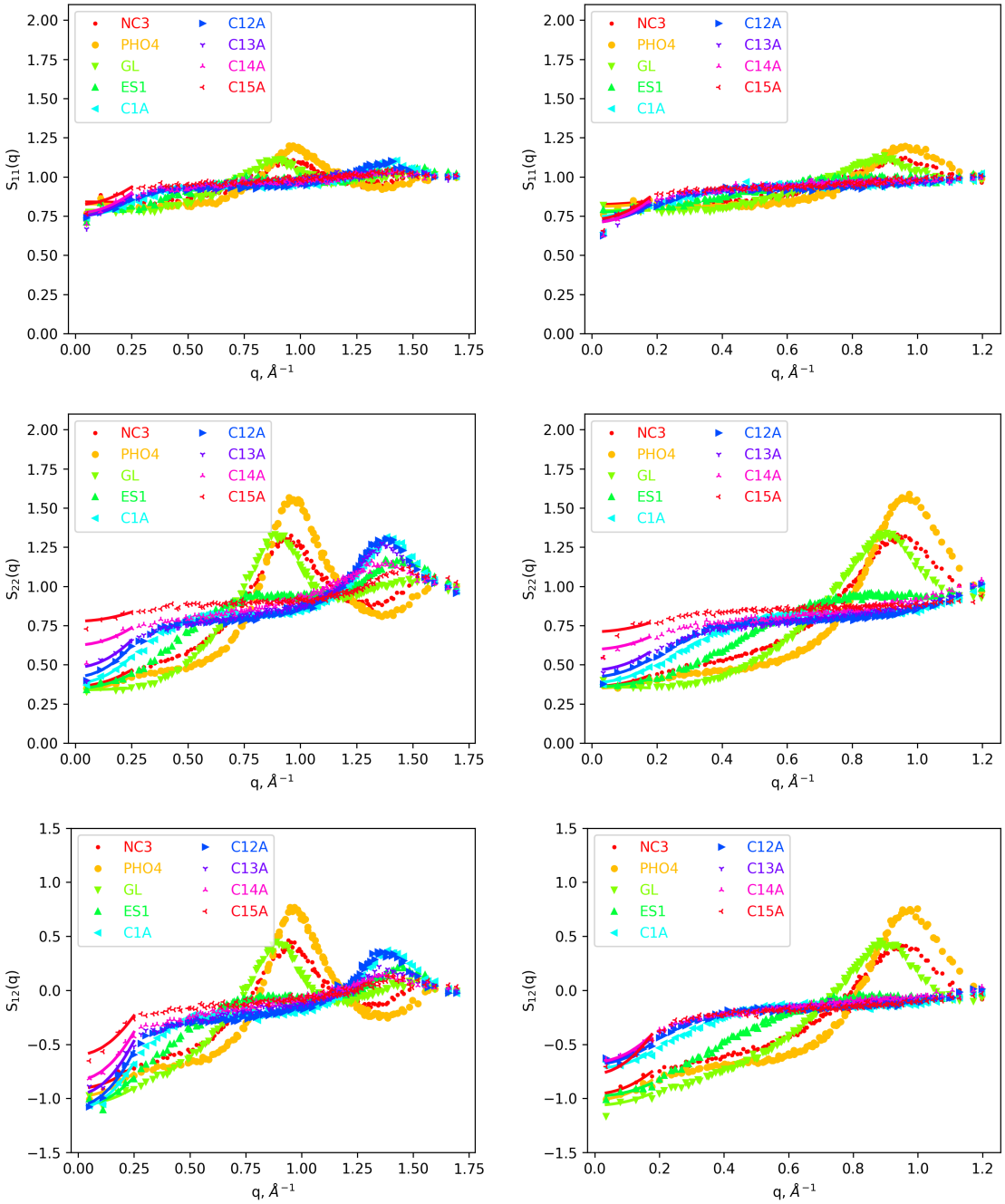
It is also interesting to see how $S(\mathbf{q})$ changes depending on the concentration of the compounds. For example, if one compares Fig. C.6 and Fig. C.7, it can be seen that the dominating lipid species have higher peaks, while the 1:1 system show almost equal behavior of the structure factor for both components. Zero values of $S(\mathbf{q})$ seem to coincide well for the S_{12} component for all cases, while for S_{11} and S_{22} components their sum remains



(A) PSM (B) SSM

FIGURE C.4: A full structure factor dependence on the bead type for available model lipids: PSM and SSM. Systems of 512 lipids and 8192 water beads at T=318 K and P=1 atm.

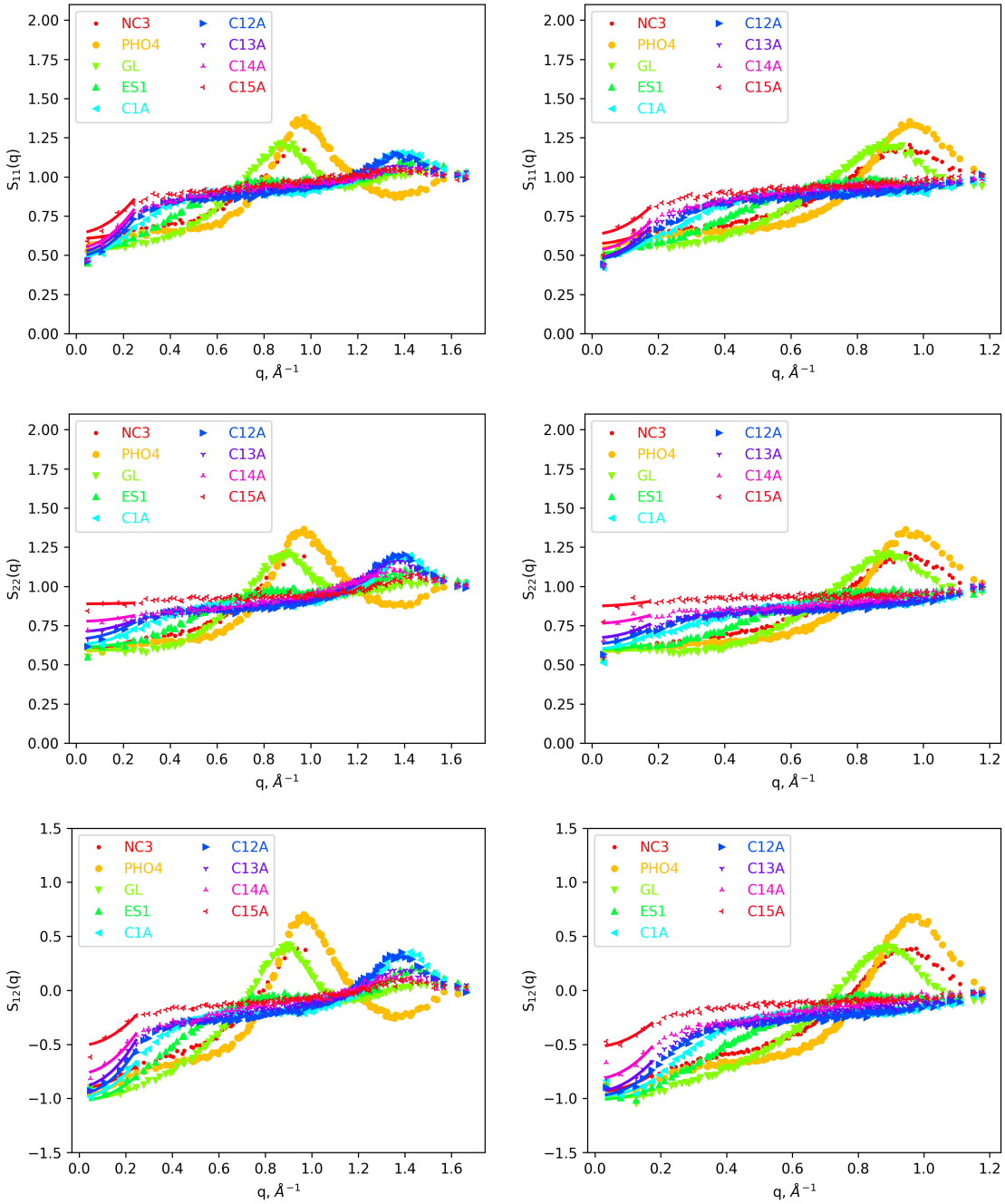
constant. Considering the equation for the compressibility based on the Kirkwood-Buff relations, it is expected that compound with higher $S(q \rightarrow 0)$ has more impact on the overall system compressibility. On the other hand, the mixing parameter should have no or slight dependence on the compound concentration as this quantity is dependent on the cross-relations between zero values of the structure factor and not on the values themselves.



(A) 128 DLiPC 384 DPPC

(B) 256 DLiPC 768 DPPC

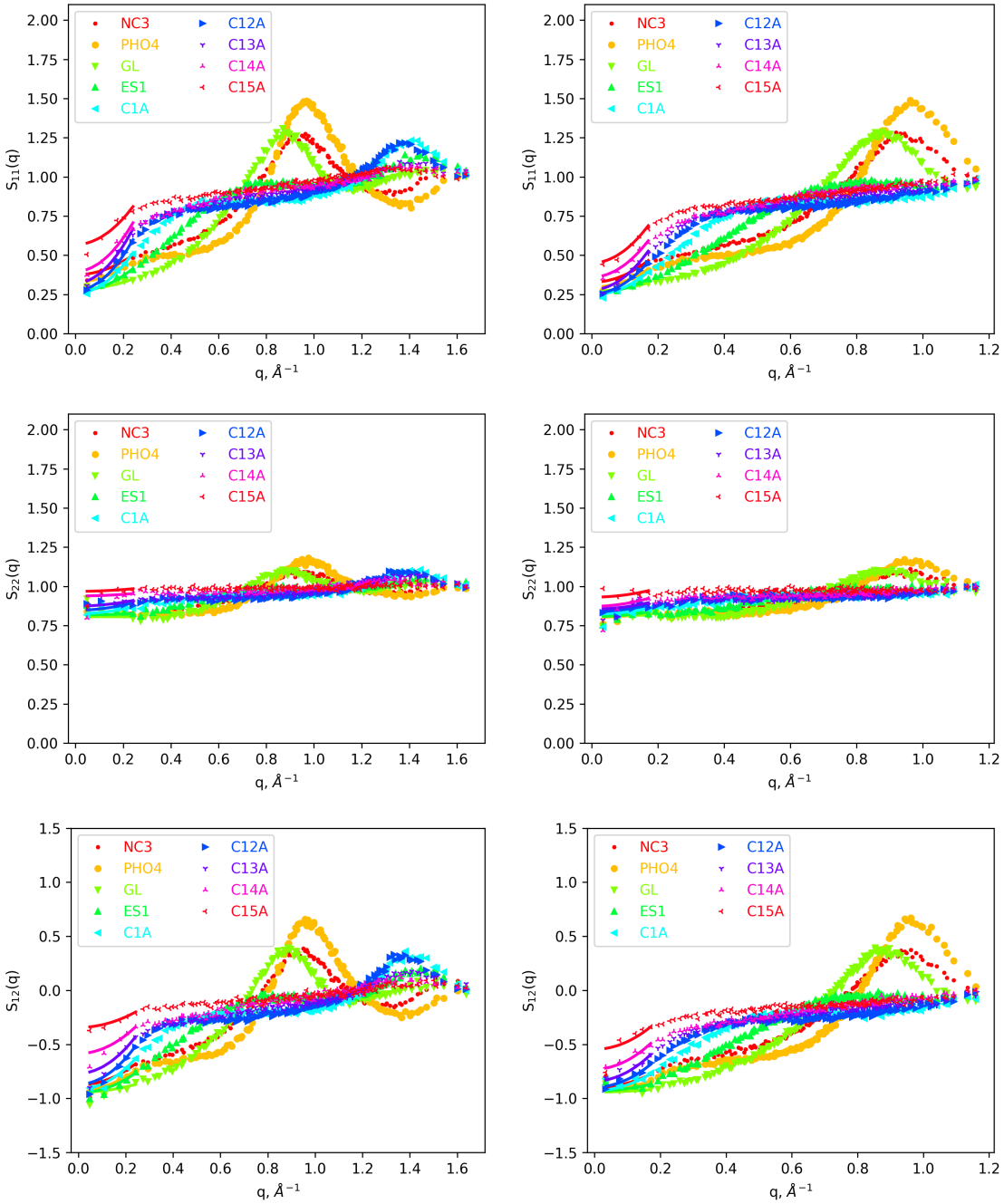
FIGURE C.5: The structure factor of 25% DLiPC 75% DPPC mixture for two system sizes: 512 and 1024 lipids in total.



(A) 256 DLiPC 256 DPPC

(B) 512 DLiPC 512 DPPC

FIGURE C.6: The structure factor of 50% DLiPC 50% DPPC mixture for two system sizes: 512 and 1024 lipids in total.



(A) 384 DLiPC 128 DPPC

(B) 768 DLiPC 256 DPPC

FIGURE C.7: The structure factor of 75% DLiPC 25% DPPC mixture for two system sizes: 512 and 1024 lipids in total.

APPENDIX D

Example LAMMPS input scripts

This appendix provides three examples of LAMMPS scripts for external force application using LAMMPS native environment: force applied to a specific bead (here, GL) D.2, to a real center of mass of each lipid D.4 and to a pseudo center of mass of each lipid D.3.

```
# lammps input script
# Initialisation of the system
variable        t equal $T # Temperature
variable        N equal ${Wat} # Number of water molecules
variable        k equal 123 # Marker for number of modes
variable        c equal 5 # Bias amplitude x10

log             log.fx.txt # Log output
units           real
dimension       3
atom_style      full
newton          on off

read_restart    system.cg.nvt.rest

pair_style      lj/sdk/coul/long                    15.0
bond_style      harmonic
angle_style     hybrid sdk harmonic
dihedral_style  charmm
special_bonds   lj/coul 0.0 0.0 1.0

kspace_style    pppm/cg 0.00001
```

```
kspace_modify   mesh 20 20 20
kspace_modify   order 3

include         PARM.FILE # Spica parameter file

neighbor        2.0 bin
neigh_modify    delay 4

timestep        10.0 # fs

########################################################################

# External potential procedure

########################################################################
reset_timestep 0

# Output
dump            1 all xtc 1000 dppc${Nlip}w$NT$t.fx0$ck$k.bead.xtc
dump_modify     1 unwrap yes precision 1000

thermo_style    custom step temp pe lx ly lz press
thermo          1000

# MD run of 10000000x1e-15x10 = 100 ns
restart         50000   system.cg.rest system.cg.rest
run             10000000
```

LISTING D.1: General LAMMPS input for the biased simulation.

```
# External potential procedure
# Bead choice
group           GL type 3

# External potential amplitude U0
variable        C equal 0.$c
# External force on each atom GL
variable        Fx atom $C*2*PI*(sin(2*PI*x/lx)+2*sin(4*PI*x/lx)+3*sin(6*
    PI*x/lx))/lx

# Thermostating all particles
fix             1 all nvt temp $t.15 $t.15 500.0
# Force is added only in x axis to GL beads
fix             2 GL addforce v_Fx 0.0 0.0
```

LISTING D.2: An external potential applied to GL bead on qx = 1, 2, 3 modes.

```
# External potential procedure
# Group of all particles in all lipids
group           alllip molecule 0:${Nlip}
# Number of atoms per molecule
variable        nat equal 15 # 15 for DPPC

# External potential amplitude U0
variable        C equal 0.$c
# External force on each atom in lipid molecule
variable        Fx atom $C*2*PI*(sin(2*PI*x/lx))/lx/${nat}

# Thermostating all particles
fix             1 all nvt temp $t.15 $t.15 500.0
# Force is added only in x axis to all beads of all lipid molecules
fix             2 alllip addforce v_Fx 0.0 0.0
```

LISTING D.3: An external potential applied to a pseudo center of mass.

```
# External potential procedure
# Group of all particles in all lipids
group           alllip molecule 0:${Nlip}
# Groups of beads
group           NC type 1
group           PH type 2
group           GL type 3
group           EST type 4 7 # EST1 and EST2 for two tails
group           CM type 5
group           CT type 6 # last tail beads

# Total mass of lipids
variable        Mtot equal mass(lipids)
# Effective mass of beads
variable        mNC equal mass(NC)/${Mtot}
variable        mPH equal mass(PH)/${Mtot}
variable        mGL equal mass(GL)/${Mtot}
variable        mEST equal mass(EST)/${Mtot}
variable        mCM equal mass(CM)/${Mtot}
variable        mCT equal mass(CT)/${Mtot}

# External potential amplitude U0
variable        C equal 0.$c
# External force on each bead
variable        Fx1 atom ${mNC}*$C*2*sin(2*PI*x/lx)*PI/lx
variable        Fx2 atom ${mPH}*$C*2*sin(2*PI*x/lx)*PI/lx
variable        Fx3 atom ${mGL}*$C*2*sin(2*PI*x/lx)*PI/lx
variable        Fx4 atom ${mEST}*$C*2*sin(2*PI*x/lx)*PI/lx
variable        Fx5 atom ${mCM}*$C*2*sin(2*PI*x/lx)*PI/lx
```

```
30 variable        Fx6 atom ${mCT}*$C*2*sin(2*PI*x/lx)*PI/lx
31
32 # Thermostating all particles
33 fix             1 all nvt temp $t.15 $t.15 500.0
34
35 # Force is added only in x axis to corresponding beads of all lipid
      molecules
36 fix             2 NC addforce v_Fx1 0.0 0.0
37 fix             3 PH addforce v_Fx2 0.0 0.0
38 fix             4 GL addforce v_Fx3 0.0 0.0
39 fix             5 EST addforce v_Fx4 0.0 0.0
40 fix             6 CM addforce v_Fx5 0.0 0.0
41 fix             7 CT addforce v_Fx6 0.0 0.0
```

LISTING D.4: An external potential applied to a real center of mass.

Bibliography

[1] B. J. Alder and T. E. Wainwright. "Phase Transition for a Hard Sphere System". In: *The Journal of Chemical Physics* 27.5 (Nov. 1957), pp. 1208–1209. DOI: 10.1063/1.1743957.

[2] M.P. Allen et al. *Computer Simulation of Liquids*. Oxford Science Publ. Clarendon Press, 1989. ISBN: 9780198556459. URL: https://books.google.fr/books?id=O32VXB9e5P4C.

[3] Paulo F.F. Almeida. "Thermodynamics of lipid interactions in complex bilayers". In: *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1788.1 (2009). Lipid Interactions, Domain Formation, and Lateral Structure of Membranes, pp. 72–85. ISSN: 0005-2736. DOI: https://doi.org/10.1016/j.bbamem.2008.08.007. URL: https://www.sciencedirect.com/science/article/pii/S0005273608002459.

[4] Peter Atkins and Julio de Paula. *Atkins' Physical Chemistry*. Oxford University Press, 2006.

[5] Arieh Ben-Naim. *Molecular Theory of Solutions*. Oxford University Press, 2006, pp. 112–127.

[6] Sandra V. Bennun et al. "Coarse-grained modeling of lipids". In: *Chemistry and Physics of Lipids* 159.2 (2009), pp. 59–66. ISSN: 0009-3084. DOI: https://doi.org/10.1016/j.chemphyslip.2009.03.003. URL: https://www.sciencedirect.com/science/article/pii/S0009308409000607.

[7] J. Javier Brey and M. J. Ruiz-Montero. "Volume fluctuations and compressibility of a vibrated granular gas". In: *Phys. Rev. E* 81 (2 Feb. 2010), p. 021304. DOI: 10.1103/PhysRevE.81.021304. URL: https://link.aps.org/doi/10.1103/PhysRevE.81.021304.

[8] Monica Bulacu, Xavier Périole, and Siewert J. Marrink. "In Silico Design of Robust Bolalipid Membranes". In: *Biomacromolecules* 13.1 (2012). PMID: 22103705, pp. 196–205. DOI: 10.1021/bm201454j. eprint: https://doi.org/10.1021/bm201454j. URL: https://doi.org/10.1021/bm201454j.

[9] Saptarshi Chakraborty et al. "How cholesterol stiffens unsaturated lipid membranes". In: *Proceedings of the National Academy of Sciences* 117.36 (2020), pp. 21896–21905. DOI: 10.1073/pnas.2004807117. eprint: https://www.pnas.org/doi/pdf/10.1073/pnas.2004807117. URL: https://www.pnas.org/doi/abs/10.1073/pnas.2004807117.

[10] P.R. Cullis and B. De Kruijff. "Lipid polymorphism and the functional roles of lipids in biological membranes". In: *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes* 559.4 (1979), pp. 399–420. ISSN: 0304-4157. DOI: https://doi.org/10.1016/0304-4157(79)90012-1. URL: https://www.sciencedirect.com/science/article/pii/0304415779900121.

[11] James Frederic Danielli and Hugh Davson. "A contribution to the theory of permeability of thin films". In: *Journal of Cellular Physiology* 5.4 (1935), pp. 495–508. DOI: 10.1002/jcp.1030050409. URL: https://app.dimensions.ai/details/publication/pub.1053569672.

[12] Haim Diamant. "Model-free thermodynamics of fluid vesicles". In: *Phys. Rev. E* 84 (6 Dec. 2011), p. 061123. DOI: 10.1103/PhysRevE.84.061123.

[13] JM Drouffe, AC Maggs, and S Leibler. "Computer simulations of self-assembled membranes". In: *Science* 254.5036 (1991), pp. 1353–1356. ISSN: 0036-8075. DOI: 10.1126/science.1962193. eprint: http://science.sciencemag.org/content/254/5036/1353.full.pdf.

[14] B. Efron. "Bootstrap Methods: Another Look at the Jackknife". In: *The Annals of Statistics* 7.1 (1979), pp. 1 –26. DOI: 10.1214/aos/1176344552. URL: https://doi.org/10.1214/aos/1176344552.

[15] S.-P. Fu et al. "Lennard-Jones type pair-potential method for coarse-grained lipid bilayer membrane simulations in LAMMPS". In: *Computer Physics Communications* 210 (2017), pp. 193 –203. ISSN: 0010-4655. DOI: https://doi.org/10.1016/j.cpc.2016.09.018.

[16] Josep Gelpí et al. "Molecular dynamics simulations: Advances and applications". In: *Advances and Applications in Bioinformatics and Chemistry* 10 (Nov. 2015), p. 37. DOI: 10.2147/aabc.s70333.

[17] T. Gil et al. "Wetting and capillary condensation as means of protein organization in membranes". In: *Biophysical Journal* 73.4 (1997), pp. 1728 –1741.

[18] E. Gorter and F. Grendel. "On bimolecular layers of lipoids on the chromocytes of the blood". In: *Journal of Experimental Medicine* 41.4 (1925), 439–443.

[19] R. J. Gowers et al. In: *MDAnalysis: A Python package for the rapid analysis of molecular dynamics simulations*. Ed. by S. Benthall and S. Rostrup. Austin, TX, SciPy. 2016, pp. 102–109.

[20] Yachong Guo, Vladimir A. Baulin, and Fabrice Thalmann. "Peroxidised phospholipid bilayers: Insight from coarse-grained molecular dynamics simulations". In: *Soft Matter* 12 (1 2015), pp. 263–271. ISSN: 17446848. DOI: 10.1039/c5sm01350j. URL: https://pubmed.ncbi.nlm.nih.gov/26462464/.

[21] Jean Pierre Hansen and I. R. McDonald. *Theory of Simple Liquids: With Applications to Soft Matter: Fourth Edition*. Elsevier Ltd, 2013, pp. 1–619. ISBN: 9780123870322. DOI: 10.1016/c2010-0-66723-x.

[22] Thomas Heimburg. *Thermal Biophysics of Membranes*. Wiley-VCH, 2007.

[23] W. Helfrich and Servuss R.-M. "Undulations, Steric Interaction and Cohesion of Fluid Membranes". In: *Il Nuovo Cimento* 3d (1 1984), pp. 137–151.

[24] Aurelia R. Honerkamp-Smith et al. "Line Tensions, Correlation Lengths, and Critical Exponents in Lipid Membranes Near Critical Points". In: *Biophysical Journal* 95.1 (July 2008), pp. 236–246. DOI: 10.1529/biophysj.107.128421.

[25] Juyang Huang and Gerald W. Feigenson. "A Microscopic Interaction Model of Maximum Solubility of Cholesterol in Lipid Bilayers". In: *Biophysical Journal* 76.4 (1999), pp. 2142 –2157. ISSN: 0006-3495. DOI: http://dx.doi.org/10.1016/S0006-3495(99)77369-8.

[26] William Humphrey, Andrew Dalke, and Klaus Schulten. "VMD – Visual Molecular Dynamics". In: *Journal of Molecular Graphics* 14 (1996), pp. 33–38.

[27] John Hjort Ipsen and Ole G. Mouritsen. "Modelling the phase equilibria in two-components membranes of phospholipids with different acyl-chain lengths". In: *Biochim. Biophys. Acta* 944 (1988), pp. 121–134.

[28] M. K. Jain and Harold B. White. "Long-range order in biomembranes." In: *Advances in lipid research* 15 (1977), pp. 1–60. URL: https://www.researchgate.net/profile/Harold-White/publication/22802371_Long-Range_Order_in_Biomembranes/.

[29] M J Karnovsky et al. "The concept of lipid domains in membranes". In: *Journal of Cell Biology* 94.1 (July 1982), pp. 1–6. ISSN: 0021-9525. DOI: 10.1083/jcb.94.1.1. eprint: https://rupress.org/jcb/article-pdf/94/1/1/1076005/1.pdf. URL: https://doi.org/10.1083/jcb.94.1.1.

[30] John G. Kirkwood and Frank P. Buff. "The Statistical Mechanical Theory of Solutions. I". In: *The Journal of Chemical Physics* 19.6 (June 1951), pp. 774–777. DOI: 10.1063/1.1748352.

[31] John G. Kirkwood and Frank P. Buff. "The statistical mechanical theory of solutions. i". In: *The Journal of Chemical Physics* 19 (6 June 1951), pp. 774–777. ISSN: 00219606.

DOI: 10.1063/1.1748352. URL: http://aip.scitation.org/doi/10.1063/1.1748352.

[32] Jeffery B. Klauda et al. "Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types". In: *Journal of Physical Chemistry B* 114 (23 June 2010), pp. 7830–7843. ISSN: 15205207. DOI: 10.1021/jp101759q. URL: https://pubs.acs.org/doi/abs/10.1021/jp101759q.

[33] K. Koga and B. Widom. "Thermodynamic functions as correlation-function integrals". In: *The Journal of Chemical Physics* 138.11 (Mar. 2013), p. 114504. DOI: 10.1063/1.4795498.

[34] Axel Kohlmeyer. *TopoTools*. 2017.

[35] Shigeyuki Komura and David Andelman. "Physical aspects of heterogeneities in multi component lipid membranes". In: *Advances in Colloid and Interface Science* 208 (2014), pp. 34–46. ISSN: 00018686. DOI: 10.1016/j.cis.2013.12.003. URL: https://pubmed.ncbi.nlm.nih.gov/24439258/.

[36] Martin R. Krause and Steven L. Regen. "The Structural Role of Cholesterol in Cell Membranes: From Condensed Bilayers to Lipid Rafts". In: *Acc. Chem. Res.* 47 (Dec. 2014), pp. 3512–3521. DOI: 10.1021/ar500260t. URL: https://doi.org/10.1021/ar500260t.

[37] Jennifer Loschwitz et al. *Computer simulations of protein membrane systems*. 2020. DOI: 10.1016/bs.pmbts.2020.01.001.

[38] Siewert J. Marrink et al. "Computational Modeling of Realistic Cell Membranes". In: *Chemical Reviews* 119 (9 May 2019), pp. 6184–6226. ISSN: 15206890. DOI: 10.1021/acs.chemrev.8b00460.

[39] Siewert J. Marrink et al. "The MARTINI force field: Coarse grained model for biomolecular simulations". In: *Journal of Physical Chemistry B* 111.27 (27 July 2007). Pmid: 17569554, pp. 7812–7824. ISSN: 15206106. DOI: 10.1021/jp071097f. eprint: https://doi.org/10.1021/jp071097f. URL: https://pubs.acs.org/doi/abs/10.1021/jp071097f.

[40] D. Marsh. *Handbook of Lipid Bilayers*. CRC Press, 2013. ISBN: 9781420088335. URL: https://books.google.fr/books?id=JgnLBQAAQBAJ.

[41] José Mario Marti and Leandro Marti. "Packing Optimization for Automated Generation of Complex System's Initial Configurations for Molecular Dynamics and Docking". In: *J Comput Chem* 24 (2003), pp. 819–825.

[42] Glenn J. Martyna, Douglas J. Tobias, and Michael L. Klein. "Constant pressure molecular dynamics algorithms". In: *The Journal of Chemical Physics* 101 (5 Sept. 1994), pp. 4177–4189. ISSN: 00219606. DOI: 10.1063/1.467468. URL: http://aip.scitation.org/doi/10.1063/1.467468.

[43] Glenn J. Martyna, Douglas J. Tobias, and Michael L. Klein. "Constant pressure molecular dynamics algorithms". In: *The Journal of Chemical Physics* 101.5 (1994), pp. 4177–4189. DOI: http://dx.doi.org/10.1063/1.467468.

[44] L Martínez et al. "Packmol: A Package for Building Initial Configurations for Molecular Dynamics Simulations". In: *Software News and Update* (Feb. 2009). DOI: 10.1002/jcc.21224.

[45] Robert T. McGibbon et al. "MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories". In: *Biophysical Journal* 109.8 (2015), pp. 1528 – 1532. DOI: 10.1016/j.bpj.2015.08.015.

[46] G. W. Feigenson G. van Meer D. R. Voelker. "Membrane lipids: where they are and how they behave". In: *Natural Reviews Molecular Cell Biology* 9 (2 2008), pp. 112–124.

[47] Gautam I. Menon and Chandan Dasgupta. "Effects of pinning disorder on the correlations and freezing of the flux liquid in layered superconductors". In: *Physical Review Letters* 73.7 (Aug. 1994), pp. 1023–1026. DOI: 10.1103/physrevlett.73.1023.

[48] Nicholas Metropolis et al. "Equation of State Calculations by Fast Computing Machines". In: *The Journal of Chemical Physics* 21.6 (1953), pp. 1087–1092. DOI: 10.1063/1.1699114. eprint: https://doi.org/10.1063/1.1699114. URL: https://doi.org/10.1063/1.1699114.

[49] M Mézard and G Parisi. "A tentative replica study of the glass transition". In: *Journal of Physics A* 29 (1996), p. 6515.

[50] Naveen Michaud-Agrawal et al. "MDAnalysis: A toolkit for the analysis of molecular dynamics simulations". In: *Journal of Computational Chemistry* 32.10 (Apr. 2011), pp. 2319–2327. DOI: 10.1002/jcc.21787. URL: https://doi.org/10.10022Fjcc.21787.

[51] Yusuke Miyazaki, Susumu Okazaki, and Wataru Shinoda. "PSPICA: A Coarse-Grained Force Field for Lipid Membranes Based on a Polar Water Model". In: *Journal of Chemical Theory and Computation* 16 (1 Jan. 2020), pp. 782–793. ISSN: 15499626. DOI: 10.1021/acs.jctc.9b00946.

[52] Ole G. Mouritsen and M. B. D. Bloom. "Mattress model of lipid-protein interactions in membranes." In: *Biophysical journal* 46 2 (1984), pp. 141–53. URL: https://pubmed.ncbi.nlm.nih.gov/6478029/.

[53] Luis A. Bagatolli Ole G. Mouritsen. *Life - As a Matter Of Fat*. Springer-Verlag GmbH, Oct. 16, 2015. ISBN: 3319226134.

[54] Tiago Espinosa de Oliveira et al. "Fluid bilayer phase in aqueous mixtures of fatty alcohol and cationic surfactant". In: *Phys. Rev. Research* 2 (1 Jan. 2020), p. 013075. DOI: 10.1103/PhysRevResearch.2.013075.

[55] W. K. Den Otter. "Area compressibility and buckling of amphiphilic bilayers in molecular dynamics simulations". In: *Journal of Chemical Physics* 123 (21 Dec. 2005),

p. 214906. ISSN: 00219606. DOI: 10.1063/1.2132287. URL: http://aip.scitation.org/doi/10.1063/1.2132287.

[56] Karel Paukner, Ivana Králová Lesná, and Rudolf Poledne. "Cholesterol in the Cell Membrane; An Emerging Player in Atherogenesis". In: *International Journal of Molecular Sciences* 23.1 (2022). ISSN: 1422-0067. DOI: 10.3390/ijms23010533. URL: https://www.mdpi.com/1422-0067/23/1/533.

[57] Atkins Peter, de Paula Julio, and Keeler James. *Physical Chemistry*. 11th ed. OUP Oxford, 2017. ISBN: 978-0198769866.

[58] E Pitard et al. "Critical behaviour of a fluid in a disordered porous matrix : an Ornstein-Zernike approach." In: *Physical Review Letters* 74 (1995), p. 4361.

[59] Steve Plimpton. "Fast Parallel Algorithms for Short-Range Molecular Dynamics". In: *Journal of Computational Physics* 117 (1995), pp. 1–19.

[60] Irina D. Pogozheva et al. "Comparative Molecular Dynamics Simulation Studies of Realistic Eukaryotic, Prokaryotic, and Archaeal Membranes". In: *Journal of Chemical Information and Modeling* 62.4 (2022). Pmid: 35167752, pp. 1036–1051. DOI: 10.1021/acs.jcim.1c01514. eprint: https://doi.org/10.1021/acs.jcim.1c01514. URL: https://doi.org/10.1021/acs.jcim.1c01514.

[61] N. A. Porter, R. A. Wolf, and J. R. Nixon. "Separation and purification of lecithins by high pressure liquid chromatography". In: *Lipids* 14.1 (1979), pp. 20–24. DOI: https://doi.org/10.1007/BF02533561. eprint: https://aocs.onlinelibrary.wiley.com/doi/pdf/10.1007/BF02533561. URL: https://aocs.onlinelibrary.wiley.com/doi/abs/10.1007/BF02533561.

[62] Arun Radhakrishnan and Harden McConnell. "Condensed Complexes of Cholesterol and Phospholipids:" in: *Biophysical Journal* 77 (1999), pp. 1507–1517.

[63] Smita Raghava et al. "The SV40 Late Protein VP4 Is a Viroporin that Forms Pores to Disrupt Membranes for Viral Release". In: *PLOS Pathogens* 7.6 (June 2011), pp. 1–15. DOI: 10.1371/journal.ppat.1002116. URL: https://doi.org/10.1371/journal.ppat.1002116.

[64] W Rawicz et al. "Effect of Chain Length and Unsaturation on Elasticity of Lipid Bilayers". In: *Biophysical Journal* 79 (2000), pp. 328–339. DOI: 10.1016/s0006-3495(00)76295-3.

[65] E. Reister and U. Seifert. "Lateral diffusion of a protein on a fluctuating membrane". In: *Europhysics Letters* 71.5 (2005), pp. 859–865. DOI: DOI:10.1209/epl/i2005-10139-6.

[66] JD Robertson. "The cell membrane concept". In: *Journal Of Physiology-london*. Vol. 140. 3. Cambridge Univ Press 40 West 20th Street, New York, Ny 10011-4211. 1958, P58–p59.

[67] Tomasz Róg and Ilpo Vattulainen. "Cholesterol, sphingolipids, and glycolipids: What do we know about their role in raft-like membranes?" In: *Chemistry and Physics of Lipids* 184 (2014), pp. 82–104. ISSN: 0009-3084. DOI: https://doi.org/10.1016/j.chemphyslip.2014.10.004. URL: https://www.sciencedirect.com/science/article/pii/S0009308414001248.

[68] Sam Safran. *Statistical Thermodynamics of Surfaces, Interfaces and Membranes*. Reading, MA: Addison-Wesley, 1994.

[69] Sangjae Seo and Wataru Shinoda. "SPICA Force Field for Lipid Membranes: Domain Formation Induced by Cholesterol". In: *Journal of Chemical Theory and Computation* 15.1 (1 Jan. 2019). Pmid: 30514078, pp. 762–774. ISSN: 15499626. DOI: 10.1021/acs.jctc.8b00987. eprint: https://doi.org/10.1021/acs.jctc.8b00987. URL: https://doi.org/10.1021/acs.jctc.8b00987.

[70] W. Shinoda, R. DeVane, and M. L. Klein. "Multi-property fitting and parameterization of a coarse grained model for aqueous surfactants". In: *Molecular Simulation* 33 (1-2 Jan. 2007), pp. 27–36. ISSN: 0892-7022. DOI: 10.1080/08927020601054050. URL: http://www.tandfonline.com/doi/abs/10.1080/08927020601054050.

[71] Wataru Shinoda, Russell Devane, and Michael L. Klein. "Zwitterionic lipid assemblies: Molecular dynamics studies of monolayers, bilayers, and vesicles using a new coarse grain force field". In: *Journal of Physical Chemistry B* 114.20 (20 May 2010). Pmid: 20438090, pp. 6836–6849. ISSN: 15205207. DOI: 10.1021/jp9107206. eprint: https://doi.org/10.1021/jp9107206. URL: https://pubs.acs.org/doi/abs/10.1021/jp9107206.

[72] Kai Simons and Elina Ikonen. "Functional rafts in cell membranes". In: *Nature* 387 (June 1997). ISSN: 6633. DOI: 10.1038/42408. URL: https://doi.org/10.1038/42408.

[73] S. J. Singer and Garth L. Nicolson. "The Fluid Mosaic Model of the Structure of Cell Membranes". In: *Science* 175.4023 (1972), pp. 720–731. DOI: doi:10.1126/science.175.4023.720. URL: https://www.science.org/doi/abs/10.1126/science.175.4023.720.

[74] D. M. Small. "Lateral chain packing in lipids and membranes". In: *Journal of Lipid Research* 25.13 (1984), pp. 1490–1500. ISSN: 0022-2275. DOI: https://doi.org/10.1016/S0022-2275(20)34422-9. URL: https://www.sciencedirect.com/science/article/pii/S0022227520344229.

[75] A. Sokal. "Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms Note to the Reader". In: *Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms Note to the Reader*. 1996.

[76] Mark M. Stevens, Aurelia R. Honerkamp-Smith, and Sarah L. Keller. "Solubility limits of cholesterol, lanosterol, ergosterol, stigmasterol, and [small beta]-sitosterol

in electroformed lipid vesicles". In: *Soft Matter* 6 (23 2010), pp. 5882–5890. DOI: 10.1039/C0SM00373E.

[77] William Stillwell. "Chapter 8 - From Lipid Bilayers to Lipid Rafts". In: *An Introduction to Biological Membranes (Second Edition)*. Ed. by William Stillwell. Second Edition. Elsevier, 2016, pp. 121–134. ISBN: 978-0-444-63772-7. DOI: https://doi.org/10.1016/B978-0-444-63772-7.00008-7. URL: https://www.sciencedirect.com/science/article/pii/B9780444637727000087.

[78] Mark E. Tuckerman. *Statistical Mechanics: Theory and Molecular Simulations*. 1st. Oxford University Press, Oxford, 2010. ISBN: 978-0-19-852526-4.

[79] Katerina Vafeiadou et al. "Replacement of saturated with unsaturated fats had no impact on vascular function but beneficial effects on lipid biomarkers, E-selectin, and blood pressure: results from the randomized, controlled Dietary Intervention and VAScular function (DIVAS) study". In: *The American Journal of Clinical Nutrition* 102.1 (May 2015), pp. 40–48. ISSN: 0002-9165. DOI: 10.3945/ajcn.114.097089. eprint: https://academic.oup.com/ajcn/article-pdf/102/1/40/23756315/ajcn097089.pdf. URL: https://doi.org/10.3945/ajcn.114.097089.

[80] Stefan Van Der Walt, S Chris Colbert, and Gael Varoquaux. "The NumPy array: a structure for efficient numerical computation". In: *Computing in Science & Engineering* 13.2 (2011), pp. 22–30. DOI: 10.1109/mcse.2011.37.

[81] Oscar Vazquez, J.N. Herrera, and L. Blum. "Static structure factor and thermodynamic properties of a binary Yukawa mixture". In: *Physica A: Statistical Mechanics and its Applications* 325.3 (2003), pp. 319–332. ISSN: 0378-4371. DOI: https://doi.org/10.1016/S0378-4371(03)00146-8.

[82] Loup Verlet. "Computer "Experiments" on Classical Fluids. II. Equilibrium Correlation Functions". In: *Phys. Rev.* 165 (1 Jan. 1968), pp. 201–214. DOI: 10.1103/PhysRev.165.201. URL: https://link.aps.org/doi/10.1103/PhysRev.165.201.

[83] Margus R. Vist and James H. Davis. "Phase Equilibria of Cholesterol/Dipalmitoylphosphatidylcholine Mixtures: H2 Nuclear Magnetic Resonance and Differential Scanning Calorimetry". In: *Biochemistry* 29 (1990), pp. 451–464.

[84] Yi Wang et al. "Enhanced lipid diffusion and mixing in accelerated molecular dynamics". In: *Journal of Chemical Theory and Computation* 7 (10 Oct. 2011), pp. 3199–3207. ISSN: 15499618. DOI: 10.1021/ct200430c. URL: https://pubs.acs.org/sharingguidelines.

[85] Max C. Watson et al. "Determining Biomembrane Bending Rigidities from Simulations of Modest Size". In: *Phys. Rev. Lett.* 109 (2 July 2012), p. 028102. DOI: 10.1103/PhysRevLett.109.028102.

[86] Max C. Watson et al. "Thermal fluctuations in shape, thickness, and molecular orientation in lipid bilayers". In: *The Journal of Chemical Physics* 135.24 (2011), p. 244701. DOI: 10.1063/1.3660673.

[87] Wikipedia. "History of cell membranes". In: (). URL: www.thefullwiki.org.

[88] J. Wolff, C. M. Marques, and F. Thalmann. "Thermodynamic Approach to Phase Coexistence in Ternary Phospholipid-Cholesterol Mixtures". In: *Phys. Rev. Lett.* 106.12 (Mar. 2011), p. 128104. DOI: 10.1103/PhysRevLett.106.128104.

[89] Philip L. Yeagle. "Chapter 2 - The Lipids of Biological Membranes". In: *The Membranes of Cells (Third Edition)*. Ed. by Philip L. Yeagle. Third Edition. Boston: Academic Press, 2016, pp. 27–56. ISBN: 978-0-12-800047-2. DOI: https://doi.org/10.1016/B978-0-12-800047-2.00002-4. URL: https://www.sciencedirect.com/science/article/pii/B9780128000472000024.

[90] Philip L. Yeagle. "Chapter 6 - Laboratory Membrane Systems". In: *The Membranes of Cells (Third Edition)*. Ed. by Philip L. Yeagle. Third Edition. Boston: Academic Press, 2016, pp. 95–114. ISBN: 978-0-12-800047-2. DOI: https://doi.org/10.1016/B978-0-12-800047-2.00006-1. URL: https://www.sciencedirect.com/science/article/pii/B9780128000472000061.

[91] Kai Zhang. "On the Concept of Static Structure Factor". In: *Soft Condensed Matter* (June 2016). URL: http://arxiv.org/abs/1606.03610.

Université de Strasbourg

Yelyzaveta BEREZOVSKA

Mélanges non-idéaux dans les bicouches phospholipidiques: approche par dynamique moléculaire

Résumé

Les phospholipides sont des molécules amphiphiles qui s'auto-assemblent en bicouches dans l'eau et présentent des propriétés physiques et chimiques complexes et diverses en fonction de leur composition. Le rôle des lipides dans les membranes cellulaires est essentiel. La dynamique moléculaire permet d'étudier les structures des bicouches à des échelles nanométriques et picosecondes, qui sont soit difficiles soit impossibles à atteindre expérimentalement.

Nous nous concentrons ici sur les propriétés thermodynamiques des mélanges lipidiques purs et binaires, en utilisant le champ de force lipidique à gros grains SPICA. En adaptant la théorie des solutions de Kirkwood-Buff à notre cas, nous analysons les fluctuations de densité dans l'espace réciproque (facteurs de structure) pour obtenir les paramètres thermodynamiques d'intérêt, tels que la compressibilité ou l'énergie libre de mélange dans une théorie de description régulière des solutions. Nous montrons ensuite comment relier en pratique les propriétés moléculaires des modèles lipidiques à gros grains aux propriétés thermodynamiques collectives du système macroscopique correspondant.

Mots-clés: bicouches lipidiques, membranes, facteur de structure, réponse linéaire, théorie de Kirkwood-Buff, mélanges lipidiques, théorie des solutions régulières, simulations numériques, SPICA, LAMMPS

Résumé en anglais

Phospholipids are amphiphilic molecules that self-assemble into bilayers in water, and display complex and diverse physical and chemical properties depending on the composition. The role of lipids in cell membranes is essential. Molecular dynamics allows to investigate bilayer structures at nanometric and picosecond scales, that are either challenging or impossible to reach experimentally.

We focus here on the thermodynamical properties of pure and binary lipid mixtures, using the SPICA coarse-grained lipid force field. Adapting the Kirkwood-Buff theory of solutions to our case, we analyse the density fluctuations in reciprocal space (structure factors) to obtain the thermodynamic parameters of interest, such as the compressibility or the free-energy of mixing in a theory of regular solution description. We then show how to relate in practice the molecular properties of coarse-grained lipid models to the collective thermodynamical properties of the corresponding macroscopic system.

Keywords: lipid bilayers, membranes, structure factor, linear response, Kirkwood-Buff theory, lipid mixtures, theory of regular solutions, numerical simulations, SPICA, LAMMPS