UNIVERSITY OF SOUTHAMPTON

Physical properties of mixed membranes explored using atomistic and coarse grained molecular dynamic simulations with enhanced sampling techniques

by

Sophia Wheeler

A thesis submitted in partial fulfilment for the
degree of Doctor of Philosophy

in the
Faculty of Natural and Environmental Sciences
School of Chemistry

October 2018

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES
SCHOOL OF CHEMISTRY

Doctor of Philosophy

by Sophia Wheeler

Studies of three different, but interlinked, aspects of membrane properties are presented here. First an analysis of both the lyotropic and thermotropic phase transitions of a phospholipid bilayer using the ELBA1.0 forcefield. This is believed to be the first time such transitions (including the rippled gel phase) have been observed using a coarse grained (CG) forcefield not specifically parameterised to do so. Further, analysis via enhanced sampling methods of the relative free energies of the phases is presented. These analyses confirm the enormous effects comparatively small changes to a forcefield can have on the aggregate behaviour of lipids modelled with it. Molecular dynamics studies on the bending rigidity of bilayers are also presented, including a comparison of 4 different computational methods for calculating the bending rigidity of bilayers. The system size dependency of such methods is compared, as well as their ability to reproduce trends in bending rigidity across lipid species already measured experimentally. The methods are compared across two different atomistic and one CG method for the first time and, despite quite different theoretical bases, are shown to produce surprisingly consistent results both with each other and with previously published experiment. Finally, two new parameterisations of cholesterol using the ELBA forcefield are explored. Their ability to induce the ordering and structure seen in atomistic simulations is measured and compared with another widely using coarse grained forcefield. ELBA's unique (amongst CG forcefields) direct compatibility with atomistic forcefields also allowed dual-resolution simulations of binary bilayers to be analysed. Results from such dual-resolution simulations are consistent with those resulting from atomistic simulation.

Contents

List of Figures

List of Tables

Lipid abbreviations

DMPC	Dimyristoylglycerophosphocholine
DOPC	Dioleoylglycerophosphocholine
DSPC	Diastearoylglycerophosphocholine
POPC	Palmitoyloleoylglycerophosphocholine
SOPC	Stearoyloleoylglycerophosphocholine

Acknowledgements

First, I must thank my supervisor, Prof. Jon Essex for his invaluable guidance and support, for his scientific rigour and for challenging me to achieve my best. Thanks also go to all members of the Essex group, past and present, for the intellectual debate, the friendship and the pub quizzes. Particular thanks to Dr Mario Orsi for introducing me to the wonderful world of ELBA, BRAHMS and LAMMPS; to Dr Samuel Genheden for all his patient advice on dual resolution simulations; and to Dr Richard Bradshaw for his sympathy and unstinting assistance when the code(s) just wouldn't compile.

Many Faculty staff have provided support during my PhD but special thanks go to Nicki Lewin for her advice and kindness in the dark days and encyclopaedic knowledge of University procedures, but most of all for being the heart of the Complexity DTC, joining which has changed my life.

Finally, my deepest gratitude to Rob, Rose and Freddie, I can't express how much your love and support has meant to me - I couldn't have done it without you.

This work was supported by an EPSRC Doctoral Training Centre grant (EP/G03690X/1)

"Nothing exists except atoms and empty space;
everything else is just opinion" Democritus

Chapter 1

Lipids and membranes

1.1 Lipids and membranes

All living organisms are composed of cells. Membranes are vital to life and crucial components of a cell. The cell membrane provides the cell with its envelope and defines it as an autopoeitic system[1] ie an autonomous and self-maintaining entity containing component-producing processes which generate recursively the same network of processes which in turn produced them. They form the enclosure between the carefully regulated intracellular (or sub cellular organelle) contents and the 'outside world', whilst also being themselves more complex than a simple impermeable obstacle. As such, a biological membrane must be capable of being flexible and selective so as to act as a dynamic regulator of the flow of material in and out as well as taking an active role in cell fusion/ division. All therapeutics which have intracellular action must pass through at least one membrane, whilst if the passage of pathogens through membranes can be interrupted, infection can be prevented. A thorough understanding of membranes' behaviour is therefore vital to fully understanding biological behaviour on a cellular and sub-cellular level.

In nature, biological membranes are very complex systems, consisting of a mixture of many different lipids, sugars and proteins which are not even be the same for both leaflets. These constituting parts are involved in a number of processes making the cell membrane a very dynamic structure. It has also been proposed that protein-lipid interactions regulate the functionality of proteins, making the lipid bilayer more than a passive hydrophobic slab or 2-dimensional solvent for proteins[2–8].

1.2 A history of membrane models

Overton is credited with being the first true 'membranologist'[9]. At the time he was working (the late 19th century) it was believed that membranes were only permeable to water. His work showed that other molecules also crossed membranes and he showed that many different cells had similar permeabilities, implying their membranes had a similar properties. He also showed the membranes had similar solubilities in oil, implying a lipid basis to the membrane[10,11]. In 1925 the lipid bilayer structure for cell membranes was first proposed by Gorter and Grendel[12]. They extracted membrane lipids from erythrocyte membranes with benzene then, having floated the lipids in a Langmuir trough, they reported that the resulting monolayer had twice the area of the calculated surface area of the erythrocyte and therefore the membrane must be a bilayer. In the same year, Leathes and Raper[13] suggested that phospholipids might be essential structural elements of cell membranes.

The first model that included proteins was the Danielli-Davison model (1935) where the proteins were merely spread on the surfaces of the bilayer[14], though later Danielli did add protein lined transmembrane channels[14]. The major problems with this model were the complete lack of dynamics, lack of membrane asymmetry or variety and the fact that a coating of globular protein on the membranes' surface would severely hinder the interactions between the lipid head groups and water.

Almost 40 years later the fluid mosaic model[15] (figure 1.1) was introduced which presented the bilayer as a 2D solvent in which proteins were 'dissolved' and freely able to diffuse in those 2 dimensions. Peripheral and integral proteins were distributed asymmetrically across the membrane, carbohydrates were attached to proteins and lipids on the outside of the membrane. The 'fluid' part of the model's name comes from the recognition (absent in previous models) that the biological membrane is dynamic and fluid. The other part - 'mosaic' - comes from the hypothesis that both proteins and lipids are dispersed inhomogeneously in lateral and transmembrane patches. These domains, which came to be termed 'rafts'[16], are discussed further in section 1.3.4.

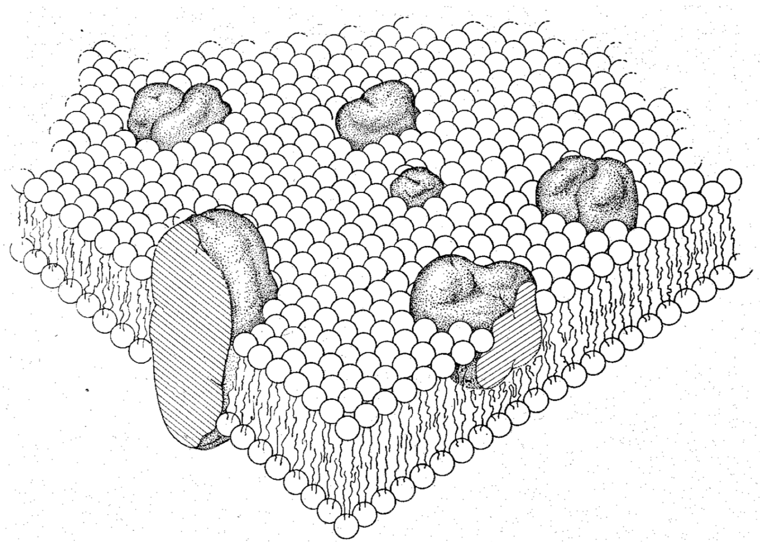


Figure 1.1: The fluid mosaic model. 'The lipid-globular protein model with a lipid matrix...the proteins are randomly distributed at long range...at short range some may form specific aggregates'[15].

1.3 Membrane lipids

There are three common classes of membrane lipids: phospholipids, glycolipids (lipids with carbohydrate attached) and cholesterol. Phospholipids are the major class of membrane lipids[17] and, with cholesterol, are the focus of the work presented in this thesis.

1.3.1 Cholesterol

Cholesterol is a lipid with a unique structure (shown in figure 1.2) compared to the other lipids; it is a sterol, built from four linked hydrocarbon rings, terminated at one end by a hydrocarbon tail and on the other by a hydroxyl group. Cholesterol is universally present in the plasma membranes of all animals, in ratios of 25-50% of the total lipid content; however, it is essentially absent from some intracellular membranes, such as mitochondrial and Golgi[18]. For the reasons further discussed below, cholesterol is important in stabilising membranes, as its presence makes them thicker and less leaky; it is also an essential component of lipid "rafts" or membrane microdomains[16].

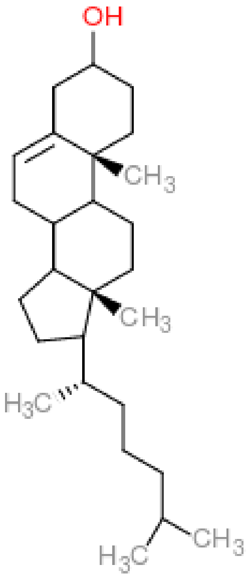


Figure 1.2: Atomistic representation of a cholesterol molecule showing the OH (polar) head group, the stiff, bulky ring region and the single hydrocarbon tail.

1.3.2 Phospholipids

Phospholipids are, like cholesterol, amphiphilic molecules. They are constructed from four components: (1) fatty acids, (2) a “backbone” to which the fatty acids are attached, (3) a phosphate, and (4) an alcohol attached to the phosphate. The fatty acid components constitute the hydrophobic portion, whereas the rest of the molecule has hydrophilic properties. The backbone component may be glycerol - a 3-carbon alcohol - in which case the phospholipid is called a phosphoglyceride, or it may be sphingosine, an amino alcohol, in which case it is referred to as a sphingolipid. For example, ceramide, an important constituent of the skin is a sphingolipid. All the phospholipids considered in this thesis are phosphoglycerides, in particular phosphatidylcholines (PCs).

Phospholipids are named by their fatty acid tails and the the name of the polar alcohol group. For example, the lipid in figure 1.3 is disteroylphosphatidylcholine or DSPC. “di...” because both tails are based on the same 18 carbon fatty acid, stearic acid; “...phosphatidyl...” because of the 1,2-diacylphosphatidyl residue; and “...choline” for the choline head group. Table 1.1 sets out the acyl tail names of the lipids used in this thesis.

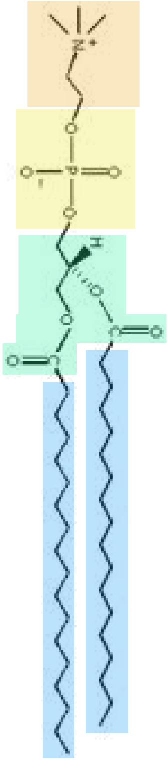


Figure 1.3: Phosphoglyceride molecule, in this case DSPC, showing the charged choline (orange) and phosphate (yellow) head groups, the glycerol and ester groups (green) and the hydrocarbon tails (blue).

Acyl tail name	C:double bonds
lauroyl	12:0
myristoyl	14:0
palmitoyl	16:0
stearoyl	18:0
oleoyl	18:1 cΔ^9

Table 1.1: The naming convention based on lipid tails for PC lipids, listing only those simulated in this thesis. The second column indicates the number of carbon atoms in each lipid tail and the number of unsaturated bonds (and their position) in each tail.

1.3.3 The hydrophobic effect

In water, a ubiquitous presence in biological systems, these amphiphilic lipids aggregate such that their tails are screened as much as possible from the water.

This is due to the hydrophobic effect, a largely entropic effect at physiological temperatures[18]. Hydrophobic molecules (and similarly, the parts of amphiphilic molecules which are hydrophobic), by definition, have no ability to form hydrogen bonds. The addition of hydrophobic molecules to water therefore interferes with the waters' tetrahedral hydrogen bonding network in the vicinity of each hydrophobic molecule, reduces the mobility of the water molecules in the solvation shell, increases their order and thus greatly decreases the entropy of the system as a whole. The interaction enthalpy between the hydrophobic molecule and the water molecules is small and similar in size to the interaction between the hydrophobic molecules themselves. Lipids are therefore driven to self-organise such that the total hydrophobic surface area exposed to water is minimised, therefore the impact on the waters' hydrogen bonding network is minimised, and hence entropy loss minimised. This outweighs loss of configurational entropy of the insoluble molecules due to the constraints typically imposed by the aggregate structure[19].

The lipids' hydrophobic tails are screened from the water whilst their hydrophilic heads (which can participate in the hydrogen bonding network) can be embedded in it. As can be seen from figure 1.4, a multiplicity of supramolecular aggregates can thereby be formed spontaneously; which structure is formed is partially dependent on the physical parameters of the constituent lipids.

One of the factors is the lipid's molecular shape[20–22]. If a packing parameter, P, is defined as:

$$P = \frac{v}{aL} \tag{1.1}$$

where a is the cross-sectional area of the lipid head group (the hydrophilic region), and v and L are the volume and length of the tail (the hydrophobic region), the stable structure of a lipid aggregate can be inferred from the magnitude of P. As a rule of thumb, though in practice the situation can more complex[23], where $P \approx 1$, each individual lipid molecule is approximately cylindrical and, en masse, such lipids would be expected to aggregate into a cubic or bilayer structure (figures 1.4(e), 1.4(g) or 1.4(i)). However where $P \ll 1$ each lipid is has a very large head group, when aggregated they form micelles (figure 1.4(b). In between, $\frac{1}{3} < P < \frac{1}{2}$, hexagonal packing becomes more likely (figure 1.4(d)). At the other extreme, for small head groups and bulky tails, $P > 1$, lipid packing is difficult and phases are complex or reverse micelles or hexagonal structures are formed (figure 1.4(a))[18].

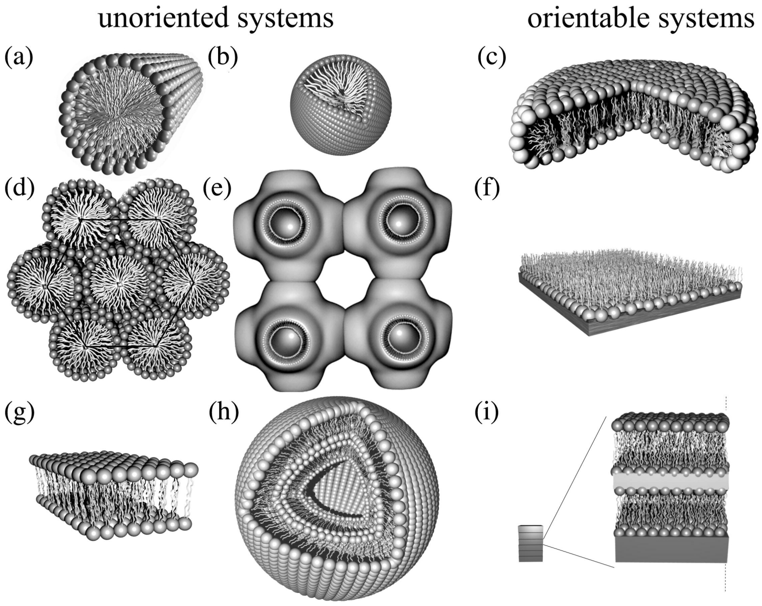


Figure 1.4: Schematic representation of hydrated phases in which lipids may self-assemble. (a) Micellar tubes, (b) spherical micelles, (c) disc-shaped bicelles, (d) hexagonal phase, (e) cubic phase, (f) langmuir monolayer at the air-water interface, (g) lamellar phase, (h) multilamellar liposomes, and (i) stacked bilayers on a solid support[24].

As well as undergoing transitions between the different aggregate morphologies described above and in figure 1.4, lipid aggregates can also undergo a different kind of phase transition within a particular structure in which they can change from a solid crystalline form to a gel to a liquid disordered form and finally a liquid. This latter set of phase changes are further described and investigated in chapter 3.

The structure of the stable phase of a lipid aggregate under a particular set of physical conditions is not only defined by the physical characteristics of the lipid molecules but also by the temperature and hydration level of the system. The structure can undergo both thermotropic - temperature dependent - and lyotropic - hydration dependant - transitions between phases. Further details of the bilayer phase transitions explored in this thesis is contained in chapter 3 and tabulated transition temperatures for the phase transition $L_{\beta'}$ to L_α of all of the phospholipids considered in this thesis can be found in table 3.1.

The presence of cholesterol in a lipid bilayer also has an effect on the bilayer structure. Ipsen et al.[25] described another phase exhibited by bilayers - liquid ordered (L_o) - induced by the presence of cholesterol. As can be seen in figure 1.2, cholesterol is also an amphiphilic molecule but with a much stiffer and bulkier tail than the other types of membrane lipids. Its stiffness leads to a preference for conformationally ordered lipid chains adjacent to it therefore implying a preference for the gel phase; on the other hand, its bulkiness leads to too much disturbance of the dense packing in that phase. Cholesterol therefore instead stabilises a more ordered phase over a broad range of temperatures and tends to co-localise with saturated lipids. Further details on the effect of cholesterol on the bilayer and the results of our simulations including cholesterol are contained in chapter 5 of this thesis. The L_o phase is like the L_α or L_d phase in that the membrane molecules exhibit a high degree of positional disorder and mobility but also like the gel phase in that the lipid chains have a high degree of conformational order. The addition of cholesterol makes the bilayer thicker and stiffer, ie more resistant to the deformation required to permit small molecules to cross the membrane, whilst retaining membrane fluidity ie having minimal effect on the lateral diffusion of proteins and lipids within the bilayer plane[26]. The effect of cholesterol on aggregate lipid behaviour is considered to be biologically very important - 30-40% by mol of the lipid content of plasma membranes is cholesterol[27].

1.3.4 Lipid Rafts

The term 'lipid rafts' was coined by Simons and Ikonen[16] and refers to cholesterol and sphingolipid enriched regions of biological membranes in the L_o phase that 'float' in the rest of the L_d membrane, as in figure 1.5. The original indication of lipid rafts was their ability to be extracted from the rest of the cell membrane with cold detergents such as Triton X-100, some being dissolved and some forming 'Detergent Resistant Membranes' (DRMs)[28]. It was suggested that the membranes contained tightly packed microdomains which were not accessible to the detergent and that it was these portions which were then found to be insoluble. The DRMs are found to consist of higher concentrations of sphingolipid-cholesterol assemblies containing a particular subset of membrane proteins[16]. In addition, it was found that GPI anchored proteins (particular proteins attached to the membrane with a particular glycolipid) and the glycosphingolipids become Triton resistant during their delivery to the apical membrane supporting the theory that their DRM and/or rafting properties form part of the sorting mechanism. However, controversy arose as to whether the assemblies found by treatment with cold detergent had, as assumed, elucidated pre-existing structures or actually arisen due to the treatment[29,30].

Experiments on model Giant Unilamellar Vesicles (GUVs) containing only saturated and unsaturated phospholipids and cholesterol, showed the formation of domains <500 nm in diameter over a wide range of cholesterol concentrations and temperatures[31–33]. Analysis confirmed the existence of 2 liquid domains; one containing a higher concentration of unsaturated lipids and the other containing a higher concentration of saturated lipids together with cholesterol. Proteins were not introduced into any of the GUV model membranes; in these model systems, lipids self organise into separated L_o-like and L_d-like domains. The drawback to these GUV models is that they are almost always lipid only systems; even when proteins are included, they are rarely (if ever) include proteins at the high ratios which are seen in biological membranes[34]. Furthermore distinction between domains in these model GUV systems is very great - the ordered domains are very ordered, whilst the disordered domains are very disordered. The difference is less distinct in systems such as Giant Plasma Membrane Vesicles (GPMVs). GPMVs are derived from living cells and therefore are made up of a similar diversity of lipids and proteins as the cells from which they are derived but they do not have a cytoskeleton nor is the asymmetry of lipid diversity across the 2 leaflets of the membrane accurately reproduced[34].

Atomistic and coarse grained simulations have been conducted[35,36] which have also shown spontaneous domain formation. The advantage of such simulations for the exploration of domains is that the lateral pressure profile of the membrane can be calculated in a simulation and it is found to be distinctly different within a raft than without.

Some experiments on real cell membranes, including FRET, single particle tracking, GPI-protein cross-linking studies and stimulated fluorescence correlation spectroscopy support evidence for ‘small, dynamic and cholesterol related nanoscale heterogeneity’[37], but at a scale orders of magnitude smaller that that seen in the model GUVs.
On the other hand, more recent experiments involving rearranging GPI anchored proteins in a live cell membrane and measuring the effect on the local membrane environment, show absolutely no effect on the membrane environment beyond that explained by the proteins’ physical size[38]. An excellent recent review[34]

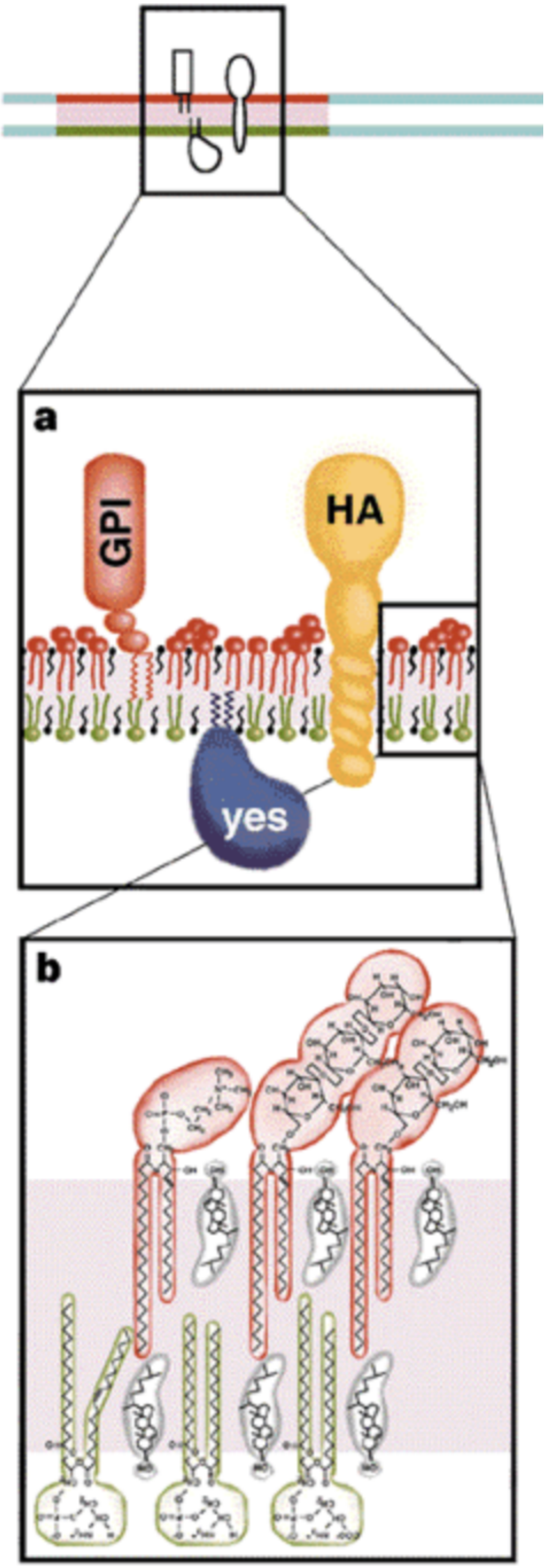


Figure 1.5: A representation of the lipid raft[16]. ‘The rafts (red) segregate from the other regions (blue)...rafts contain proteins attached via GPI anchors... or transmembrane domains’.

describes the history of the lipid raft hypothesis, the measurement techniques employed to elucidate them and notes that 'lipid rafts [*in vivo*] continue to elude direct microscopic detection; thus, the presence and exact nature of rafts in live cells remain the subject of debate, particularly as different methodologies can often yield seemingly contradictory results'.

This continuing debate, as well as the development of experimental techniques driven by it, has opened opportunities for simulation to contribute. Lyman et al[39] have proposed a 'modelling manifesto' describing the future challenges and opportunities open to biophysical simulation; particularly focussing on the requirement to reproduce lipid and protein mobility from simple to complex systems, the need for sufficient system scale so that finite scale effects are minimised and recommending atomistic forcefields (especially CHARMM36) 'incorporating sufficient chemical accuracy' to 'faithfully represent lipid protein interactions'.

1.4 Conclusions

Membrane models and knowledge of the lipid constituents of them have come a long way since the late 19th century but much remains to be discovered. Improved imaging techniques for *in vivo* cell measurement, GUV models and simulation techniques will continue to refine these models into the future. The last of these forms the motivation for this work; we explore the physical behaviour of simulated membranes to better understand their structure and function and also to examine the strengths and limitations of the simulation (and analysis) techniques themselves.

A thorough understanding of membranes' behaviour remains elusive but remains vital to fully understanding biological behaviour on a cellular and sub-cellular level.

Chapter 2

Molecular Dynamics and forcefields

One way to better understand the complex behaviour of lipids and membranes, is to model their complex processes as accurately as possible. A good model has to provide a precise description of the process and to allow the prediction of the future behaviour of the system or predict the results of similar processes. Models can be physical or theoretical but can also be computer simulation based. Those employed in this thesis to investigate lipid bilayers' physical behaviour are exclusively computational and are specifically, molecular dynamics simulations.

2.1 Forcefields and molecular dynamics

Molecular dynamics (MD) simulations involve using Newton's classical equations of motion to establish how the positions and energy of systems of atoms and molecules evolve over time ie the 'trajectories' of the system. Analysis of trajectories allows time dependent properties of the system such as diffusion coefficients to be calculated as well as ensemble dependent properties (assuming the ergodic principle holds). However, because the force on one particle depends on the positions of all the other particle in the system, numerical solution methods must be used; when applied to many-particle systems, an analytical solution of Newton's equations is not possible. This chapter describes such numerical methods as well as the potentials employed in the simulations reported elsewhere in this thesis.

2.1.1 Potentials

The theoretical underpinnings of molecular dynamics are no more than Newton's equations of motion. Molecular dynamics simulations use the Born-Oppenheimer approximation that the fast and slow degrees of freedom (ie the motions of the electrons and nucleii respectively) are separable and then make the assumption that nuclei are sufficiently heavy that the mass and separation of the electrons can be neglected.

Equation 2.1 shows that 'all' that is required to predict an atoms's (or other particle's) motion, whether on its own or as part of a molecule, is to establish the force on the atom at any particular time ($\boldsymbol{F}$), and then integrate to find its velocity due to that force ($\boldsymbol{a}$). But in order to establish the force that the particle experiences, we need the gradient of the potential energy surface at the point where it sits, ie the negative spatial derivative of the potential energy function (U).

$$-\frac{dU}{d\boldsymbol{r}} = \boldsymbol{F} = m\boldsymbol{a} = m\frac{d^2\boldsymbol{r}}{dt^2} \tag{2.1}$$

The potential energy function (U) describes the interactions between the particles in the system. Therefore before MD can be employed, a functional form and associated set of interaction parameters must be defined - known as a forcefield - which is then used to calculate the potential energy surface.

There are many different FF available. They have usually been prepared by fitting energy functions and parameters to experimental data, "empirical forcefields", or, in the case of coarse grained forcefields, fitting to atomistic simulation data, which is the approach taken in the attempt to paramterise cholesterol in the ELBA forcefield in chapter 5. Alternatively, some have been prepared using *ab initio* quantum calculations, as is the case for the SLipid forcefield[40]. A brief discussion of the parameterisation and specific details of each different forcefield described in this thesis is set out below, but generally, a forcefield's functional form involves separating the energy contributions into 2 main categories; those resulting from interactions with other particles bonded to the particle of interest (bond stretching, bending and torsion) and interactions with all the particles to which it is NOT bonded (van der Waals and electrostatic) as set out in equation 2.2.

$$U_{tot} = U_{bonded} + U_{non-bonded} = (U_{bond} + U_{angle} + U_{dihedral}) + (U_{vanderWaals} + U_{electrostatic}) \tag{2.2}$$

2.1.1.1 Bonded interactions

For most forcefields, the interactions between bonded particles can be modelled as if the particles were connected by flexible springs: bond stretching, angle bending and torsion potentials are therefore considered.

Bond stretching The energy contribution from bond stretching (figure 2.1a) is usually modelled by a harmonic (Hooke's law) potential as seen in equation 2.3, where r_0 is the equilibrium bond length, r_{ij} the instantaneous bond length and k_r the force constant. Alternatively, the Morse potential (equation 2.4) is used in some forcefields.

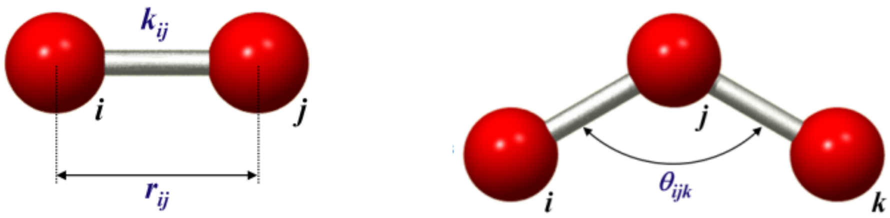
$$U_{stretch} = k_r(r_{ij} - r_0)^2 \tag{2.3}$$

$$U_{stretch} = D_e\{1 - exp\left(-a\left(r_{ij} - r_0\right)\right)\}^2 \tag{2.4}$$

where D_e is the energy well depth defined in relation to dissociated atoms, a is a constant which defines the width of the energy well and r_0 and r_{ij} are, as in equation 2.3 the equilibrium bond length and instantaneous bond length respectively.

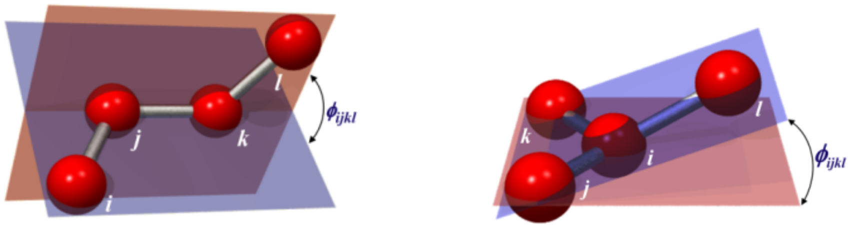
Bond bending The energy contribution from bond bending (figure 2.1b) is also usually modelled by a harmonic (Hooke's law) potential (equation 2.5), although, because the energy required to bend a bond is usually much lower than for bond stretching, with a lower force constant

$$U_{bend} = k_\theta \left(\theta - \theta_0)\right)^2 \tag{2.5}$$



(a) Bond stretching[41] (b) Bond bending[41]

Figure 2.1: Schematic diagrams for bond bending and bond stretching



(a) Bond torsion - rotation about chemical bonds, in this diagram, the rotation about the bond between atoms j and k is considered[41].

(b) Bond torsion - rotation out of the plane. This term is used, for example to keep planar molecules, such as benzene, planar. A harmonic potential can also be used for this term[41]

Figure 2.2: Schematic diagrams for bond torsion

Bond torsion There are 2 types of bond torsion (dihedrals) commonly used, see figure 2.2; bond torsion ie rotation about a chemical bond, and rotation out of the plane.

The exact function form of the bonded interactions is different for some of the forcefields employed in this thesis and is therefore described in more detail for each separately later in this chapter.

2.1.1.2 Nonbonded interactions

Non-bonded interactions can be divided into van der Waals type interactions and electrostatics. Calculation of these non-bonded interactions is the most computationally demanding part of an MD simulation. This is because each atom is bonded, in the case of carbon, to a maximum of 4 others, thus bonded interactions are inherently limited in number. However, for the non-bonded

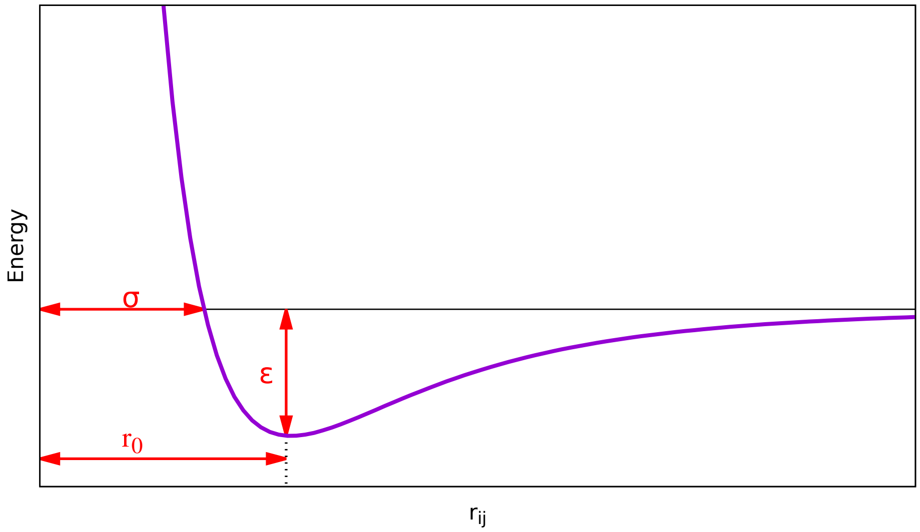


Figure 2.3: The shape of the Lennard Jones potential.

interactions, the force on an individual particle is the sum of the forces on it due to its interaction with all the other particles. We can reduce the number of calculations by only considering pairwise interactions and introducing a cut-off distance beyond which interactions are simply not computed.

van der Waals potential The way particles experience non-electrostatic interactions is made up of two principal features; resistance to compression and therefore a repulsive force at close range, together with an attractive force which causes the particles to aggregate and reproduces for example, the aggregation of atoms in liquids and solids. Potential functions have been developed which exhibit both these characteristics and approximate real data. The most commonly used of these is the Lennard Jones (LJ) potential[42,43] as plotted in figure 2.3 and set out in equation 2.6, which describes the potential energy for a non-bonded pair of particles i and j located at $\mathbf{r}_i$ and $\mathbf{r}_j$ respectively, with $\mathbf{r}_{ij}$ being the distance between them. The strength of the interaction is governed by the parameter ϵ, the 'well depth', and σ determines the interaction length scale ie the point at which the interaction ceases to be repulsive and becomes attractive.

$$U_{ij} = 4\epsilon\left(\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right) \tag{2.6}$$

For particles which have charges and/or dipoles, there are additional elements of the non-bonded potentials:

Charge-charge potential The interaction potential energy $U(r)$ between two point charges Q_i and Q_j located with a distance r between them is defined by Coulomb's law:

$$U(r) = \frac{Q_i Q_j}{4\pi\varepsilon_0 r} \tag{2.7}$$

with ε_0 the relative permittivity of the vacuum.

Charge-dipole potential. For a charge Q interacting with a dipole of magnitude μ, the electrostatic potential energy is:

$$U(\boldsymbol{\mu}, \mathbf{r}) = \frac{Q\mu}{4\pi\epsilon_0 r^2} \left(\boldsymbol{\mu}_i \bullet \mathbf{r}_i\right) \tag{2.8}$$

with $\boldsymbol{\mu}_i$ and $\mathbf{r}_i$ being the unit vector along the dipole, and the charge-dipole distance vector, respectively.

Dipole-dipole potential. For dipole-dipole interactions, the interaction energy is:

$$U(\boldsymbol{\mu_i}, \boldsymbol{\mu_j}, \mathbf{r}) = \frac{\mu^2}{4\pi\epsilon_0 r^3} \left(\boldsymbol{\mu_i} \cdot \boldsymbol{\mu_j} - 3\,\frac{\boldsymbol{\mu_i} \bullet \mathbf{r}}{r|\boldsymbol{\mu_i}|}\,\frac{\boldsymbol{\mu_j} \bullet \mathbf{r}}{r|\boldsymbol{\mu_j}|} \right) \tag{2.9}$$

with $\boldsymbol{\mu}_i$ and $\boldsymbol{\mu_j}$ are the unit vectors along the directions of the two dipoles.

2.1.1.3 Cut offs, neighbour lists and long range interactions

As described above, calculation of nonbonded interactions is the most computationally expensive part of an MD simulation. One simplification that reduces the load is to only consider pairs of particles (pairwise interactions). Another approximation is made by disregarding all such pair-wise interactions beyond a certain distance. Efficiency is then further increased by constructing a "neighbour list" for each particle. All particles within a specific radius (the neighbour list cut-off) are found and then those which have a bonded interaction are discarded, leaving a list of pairwise exclusively non-bonded interactions to be considered. Because the neighbour list cut-off is typically larger than the cut-offs for the non-bonded interactions, this neighbour list need not be updated every timestep and both the MD engines, GROMACS and LAMMPS, used in this thesis allow the user to specify how often the neighbour list is updated - every 20, 30, etc timesteps depending on the system.

As can be seen from figure 2.3, the use of a cut-off for the van der Waals interactions represented by the LJ potential is reasonable because the potential decays rapidly as the distance between particles increases. The effect interactions beyond the cut-off can be added back into the total energy by including an analytically solvable 'tail correction'. This correction assumes that $g(r > r_c) \approx 1$ and therefore the tail correction term to be added to the pairwise energy calculation for $r < r_c$ is:

$$U_{tailcorrection} = \int_{r_c}^{\infty} g(r) 4\pi r^2 dr \tag{2.10}$$

The most computationally efficient way to deal with the electrostatic terms is also to use a cutoff radius similar to that used for the van der Waals terms. However, the use of cut-offs with respect to the electrostatic terms is more problematic because they do not fall off rapidly with distance and long-range electrostatic interactions are often important features of the system. A 'digital' cutoff in an energy term introduces a sharp discontinuity for particles at or approaching the cut off distance which can result in artificial energy fluctuations and differential heating between components of the system depending on their environment (which is why MD simulations tend to thermostat the solvent separately from other components). These fluctuations are particularly problematic when the energy terms are not close to zero at the cut-off point. Instead of applying a digital cut-off, shifted or scaling functions can be used to modulate the potential by multiplying the calculated energy by a smoothly varying scaling factor from 0 to 1 over a cutoff radius. A switched form uses an additional sigmoidal function within the cutoff radius to switch the potential energy to zero. A shifted form uses a different functional form within the cut-off radius in order that it decays to zero.

Alternatively, the electrostatic terms can be divided into short- and long-range components. For the short-range, a digital cut-off is applied to the potential but beyond that cut-off the potential is not set to zero; a more accurate but also more computationally intensive method is used to model the potential, the Ewald summation, usually using a particle mesh Ewald (PME) algorithm[44].

Using this technique, the long-range electrostatic potential is represented by a finite Fourier series which is solved in reciprocal space using fast Fourier transforms and a grid-based distribution of partial-charges.

Forcefields are a combination of the functional forms discussed above and the associated parameters for each atom/bead type or combination of types. The parameters associated with bonded functional forms are specifically designed for the atoms/beads involved in that bond type and therefore the number of possible bonded interactions between atom types is limited to those pre-defined in the forcefield. For non-bonded interactions, for example LJ interactions, an atomistic forcefield usually defines the σ and ϵ parameters for 2 'type A' atoms interacting with each other and 2 'type B' atoms interacting with each other and the combining rules are then used to calculate the values for the interaction between type A and type B. The most common combining rules, and those used by most of the forcefields considered in this thesis, are the Lorentz Berthelot (LB) rules. The σ to be used for the interaction of an A particle with a B particle is found by taking the arithmetic mean of the σs for an interaction between a pair of A particles and the σ defined for the interaction between a pair of B particles (equation 2.11), whilst the ϵ for the A-B interaction is the geometric mean of the 2 homoatomic ϵs (equation 2.12).

$$\sigma_{AB} = \frac{1}{2}(\sigma_{AA} + \sigma_{BB}) \tag{2.11}$$

$$\epsilon_{AB} = \sqrt{\epsilon_{AA}\epsilon_{BB}} \tag{2.12}$$

With the exception of ELBA, the majority of CG FFs do not employ mixing rules and therefore the parameters of every pairwise LJ interaction must be pre-defined as part of the forcefield[45].

2.2 Specific Forcefields

There are 2 different classes of FF used in this thesis - atomistic and CG. As the names suggest, atomistic FF calculate each atom's position in the simulation as a separate particle whilst CG forcefields group together several heavy (ie non-hydrogen) atoms and treat each group as a single particle or 'bead'. There are further classes of forcefield (not employed in this thesis) for example, those that sit between these 2 levels of abstraction: the united atom forcefields. In a united atom forcefield, each heavy atom is represented by a separate bead but the hydrogen atoms are not separately represented. In addition, at a greater level of abstraction, continuum models have been developed which treat materials as having no internal structure but instead use a representative volume element, ie the smallest amount of material over which a property can be modelled yielding a result representative of the whole[46].

2.2.1 Atomistic Forcefields

2.2.1.1 CHARMM

The forcefield used for most of the atomistic simulations in this thesis is "Chemistry at HARvard Molecular Mechanics" (CHARMM36)[47,48].

The potential energy functional forms employed by the CHARMM FF are listed in table 2.1. They include the bonded terms listed in equation 2.2 above, with the addition of extra terms representing angle-bending due to non-bonded forces between atoms 1 and 3 ($U_{Urey-Bradley}$), out-of-plane bending ($U_{improper}$) and protein backbone dihedral angle corrections (U_{CMAP}) the last of these are not used in the simulations in this thesis. All the CHARMM simulations in this thesis used a LJ switching function over 8 - 12 Å and PME was employed for long-range (beyond 12 Å) for electrostatic interactions.

2.2.1.2 SLipids

The remaining atomistic simulations in chapter 4 were conducted using an atomistic forcefield parameterised specifically for lipids: Stockholm Lipids ('SLipids')[40,49]. SLipids was developed because it was noted that reproducing lipid behaviour in simulations presents a particular challenge - for example CHARMM is required to use specialized parameters for lipids that differ from the parameters used for proteins to describe the same type of atoms[40].

The functional forms employed by the SLipids forcefield are similar to those employed by the CHARMM forcefield and are listed in table 2.2. The associated

Interaction Type	Description	Functional form
Bonded	Bond-stretching	$k_r(r - r_0)^2$
	Angle bending	$k_\theta(\theta - \theta_0)^2$
	Urey-Bradley	$k_{UB}(b^{1-3} - b^{1-3,0})^2$
	Dihedral	$k_\phi(1 + cos(n\phi - \delta))$
	Improper	$k_\omega(\omega + \omega^0)^2$
NonBonded	van der Waals	LJ modified by a force switch
	Electrostatic	Short range - $\frac{(q_i q_j)}{(4\pi\epsilon_r\varepsilon_0 r_{ij})}$ Long range - PME

Table 2.1: The functional forms employed in the CHARMM36 forcefield[47]. Where the k indicates the respective force constants differentiated by sub-scripts, ϕ the dihedral angle, with n its multiplicity and δ its phase shift; ω the dihedral angle, q_i and q_j are the partial charges on the i^{th} and j^{th} particle respectively; and ε_r is the relative dielectric constant. The superscript '0' denotes the equilibrium value for the relevant variable.

parameters are mainly based upon *ab initio* calculations with fitting to empirical data from bulk alkane lipids to improve the parameterisation of the lipid tails[40,49]. All the SLipid simulations in this thesis used a LJ switching function over 14 - 15 Å and PME was employed for long-range (beyond 15 Å) for electrostatic interactions.

2.2.2 Coarse grained (CG) Forcefields

CG modelling has become an increasingly popular approach to the simulation of biological systems. In a CG forcefield, the particles in the simulation are not individual atoms but 'beads' which each represent several atoms. The computational

Interaction Type	Description	Functional form
Bonded	Bond-stretching	$k_r(r - r_0)^2$
	Angle bending	$k_\theta(\theta - \theta_0)^2$
	Urey-Bradley	$k_b(b^{1-3} - b_0^{1-3,0})^2$
	Torsions	$k_\phi(1 + cos(n\phi - \delta))$
NonBonded	van der Waals	LJ modified with a potential switch
	Electrostatic	Short range - $\frac{(q_i q_j)}{(4\pi\epsilon_r\varepsilon_0 r_{ij})}$ Long range - PME

Table 2.2: The functional forms employed in the SLipid forcefield[40].

advantages of CG forcefields over atomistic forcefields are 3-fold: first, the number of beads in a CG system are much fewer than the number of atoms they represent and therefore the number of interactions and integrations to be calculated are reduced; second, fewer particles means that the degrees of freedom of the system are reduced and therefore the energy landscape is smoother and can be explored more rapidly; and finally, a CG system tends to have no very fast degrees of freedom (in an atomistic simulation the fastest motions are the C-H bond vibrations) which means a longer timestep can be employed for the simulation and hence a longer total simulation time can be explored with the same number of steps.

To further increase simulation efficiency, the representations of water and electrostatics are usually highly simplified in CG FF. It is common to not explicitly include either the electrostatics of the water molecules themselves nor the dipoles in the glycerol-ester region of the lipids, though the charges in the lipid head groups are represented using Coulombic potentials (equation 2.7). The reduction in explicit electrostatics means that properties which are dependent on electrostatic interactions cannot be represented physically accurately. In particular, the representation of water as an apolar solvent means that effects which are dependent on the collective alignment of the water dipoles or upon the dielectric screening provided by water can not be captured save by (over) parameterisation of the Lennard Jones potential and use of an explicit dielectric constant.

2.2.2.1 MARTINI

One of the most popular CG forcefields is MARTINI[50,51]. The MARTINI force field was developed using a ‘top-down’ approach by iterative calibration of the non-bonded interactions of the various particle types against experimental data, in particular thermodynamic data such as oil/water partition coefficients. It represents 3-4 heavy atoms or 4 water molecules as a single bead. Electrostatic interactions are only present between lipid headgroups, where they are treated through a Coulombic potential. Because no electrostatics are included in the water model in MARTINI, a relative dielectric constant $\varepsilon_r = 20$ must employed for explicit screening[50]. The LB rules for combining LJ parameters cannot be used for MARTINI beads; bead-bead interactions, both bonded and non-bonded must be specifically and individually parameterised. The long-range electrostatics in the MARTINI simulations in this thesis did not employ PME but rather a reaction field formulation whereby a constant dielectric environment is assumed beyond the cut-off, 11 Å, with a dielectric constant of $\epsilon_{rf} = 15$.

The specific functional forms employed by the MARTINI forcefield are similar but simpler to those in either of the atomistic forcefields and are listed in table 2.3.

2.2.2.2 ELBA

The ELBA forcefield was developed specifically for simulating lipid membranes with a view to retaining certain aspects of the underlying physical chemistry of the system which are abstracted by other CG forcefields especially MARTINI. Following a similar approach to MARTINI and other similar CG forcefields, molecules are represented by ‘beads’, each representing a group of 3 heavy atoms. In the ELBA forcefield, the main electrostatic features of the system, the water, the glycerol/ester region, as well as the lipid headgroups are included explicitly via charges and point dipoles. This contrasts with MARTINI and other CG methods, which only include charges in the head group region and compensate for water screening effects by introducing explicit dielectric screening constants.

Interaction Type	Description	Functional form
Bonded	Bond-stretching	$k_r(r - r_0)^2$
	Angle bending	$\frac{1}{2}k_\theta(cos\theta - cos\theta_0)^2$
	Dihedrals	$k_\phi(\phi - \phi_0)^2$
NonBonded	van der Waals	LJ adjusted with a verlet list potential shifted function[52]
	Electrostatic	Short range - $\frac{(q_i q_j)}{(\epsilon r_{ij})}$ Long range - reaction field

Table 2.3: The functional forms employed in the MARTINI forcefield[50].

In the first parameterisation of the ELBA force-field; ELBA0, water was represented by a soft sticky dipole and glycerol, ester and hydrocarbon lipid tails by anisotropic Gay-Berne type potentials ie ellipsoid beads. Whilst providing some promising results[53–55], this first approach had limitations, most notably displaying unphysical interdigitation when simulating lipid bilayers in the gel phase and the inability to reproduce inverse hexagonal phases. The Gay-Berne potential is also an overly complex form, requiring six independent user-defined parameters. Finally, an additional complexity in the form of a scaling factor, was required to reduce the strength of the LJ interactions between hydrophilic and hydrophobic groups from that calculated by the standard L-B mixing rules. In the absence of a such a scaling factor, preassembled bilayers simulated with ELBA0 were unstable[56], and dispersions of lipids and water did not self-assemble into bilayers[56].

These issues were resolved by a complete reparameterisation of ELBA in 2011[56]. The ellipsoid tail beads were replaced with a simpler Lennard-Jones based spherical form. This meant that the hydrophobic effect could be reproduced without deviation from the standard L-B mixing rules, save that, in the case of pairs of particles which are capable of forming hydrogen bonds, the ϵ term is increased to account

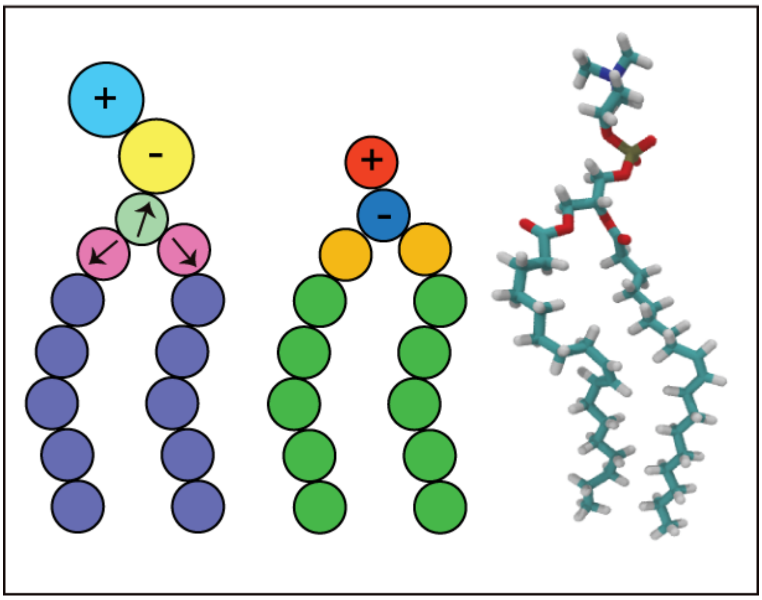


Figure 2.4: Coarse grained and atomistic representations of DSPC. On the left is a DSPC lipid shown in the ELBA model. The electrostatics are shown; positive and a negative point-charges in the headgroups, and point-dipoles (arrows) in the glycerol and ester sites. In the centre is a DSPC molecule in the MARTINI CG forcefield; the only electrostatics are positive and negative point charges in the head group. On the right is an atomistic representation of DSPC.

for such stronger interaction capability (see table 2.4). Each water molecule is represented by a single bead with an embedded point dipole. As shown in figure 2.5, each DSPC molecule is represented by 15 beads: 1 choline bead (with an embedded point positive charge), 1 phosphate bead (with an embedded point negative charge), 1 glycerol bead (with an embedded point dipole), 2 ester beads (each with embedded point dipoles) and a sets of 5 beads representing each of the 2 tails.

The ELBA1.0 water model was parametrized specifically for bulk water in the liquid phase[56]. The dipole moment magnitude was set to 2.3 D, within the range calculated for atomistic water models. The L-J parameters were obtained through incremental changes to match the calculated bulk density from ELBA1.0 simulations and the corresponding experimental measurements. The lipid model was parameterised initially for DOPC, matching dipole magnitude to atomistic simulations and incrementally adjusting σ, ϵ and r_0 to match the experimental density, area per lipid and spontaneous curvature. As noted in chapter 1, DSPC is structurally very similar to DOPC, the only difference being that while DOPC has a double-bond, or unsaturation in the middle of each of the two tails, DSPC tails contain exclusively

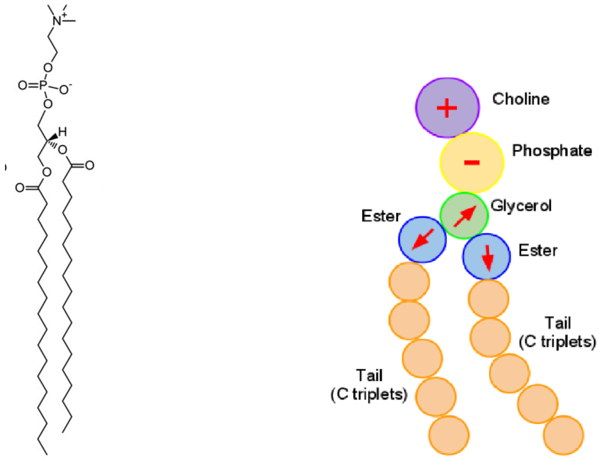


Figure 2.5: DSPC molecule represented atomistically (left) and in the ELBA forcefield (right). The ELBA representation shows the electrostatic features; positive and a negative point-charges in the headgroups, and point-dipoles (arrows) in the glycerol and ester sites.

saturated bonds. The DSPC model was therefore obtained from the DOPC model by simply resetting the angle for the unsaturated bond in the DOPC tails to that used for all the other tail sites corresponding to saturated bonds.

The parameters and functional forms used in ELBA1.0 are set out in tables 2.4 and 2.5.

The form of the Lennard-Jones potential is altered with a shifted function in the ELBA 1.0 forcefield so that it falls smoothly to zero at the cut-off distance (r_c) without introducing discontinuities in the potential and its derivative forces:

$$U_{ij}^{LJ} = 4\epsilon \left\{ \left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right] + \left[6\left(\frac{\sigma}{r_c}\right)^{12} - 3\left(\frac{\sigma}{r_c}\right)^{6} \right] \left(\frac{r}{r_c}\right)^{2} - 7\left(\frac{\sigma}{r_c}\right)^{12} + 4\left(\frac{\sigma}{r_c}\right)^{6} \right\} \tag{2.13}$$

The charge-charge and charge-dipole interaction potentials are also altered by using shifted functions for similar reasons, where:

$$U_{ij}^{QQ} = \frac{Q_i Q_j}{4\pi\varepsilon_0 r_{ij}} \left(1 - \frac{r_{ij}}{r_c}\right)^{2} \tag{2.14}$$

σ_{CC}, σ_{PP}	0.52 nm
σ_{GG}, σ_{EE}	0.46 nm
σ_{TT}	0.45 nm
σ_{WW}	0.30 nm
ϵ_{CC}, ϵ_{PP}	6.0 kJ/mol
ϵ_{GG}, ϵ_{EE}	4.0 kJ/mol
ϵ_{TT}	3.5 kJ/mol
ϵ_{WW}	1.0 kJ/mol
ϵ_{WW}^{tot}	$1.95\epsilon_{WW}$
ϵ_{WP}	$1.8\sqrt{\epsilon_{WW}\epsilon_{PP}}$
ϵ_{WG}	$1.2\sqrt{\epsilon_{WW}\epsilon_{GG}}$
ϵ_{WE}	$1.6\sqrt{\epsilon_{WW}\epsilon_{EE}}$
Q_C	+0.7 e
Q_P	-0.7 e
μ_G	1.6 D
μ_E	2.0 D
μ_W	2.3 D
k_r	1260 kJ/(mol.nm^2)
k_θ	30 kJ/mol
k_{dip}	10 kJ/mol
$\alpha_{0_{CPG}}$	115°
$\alpha_{0_{PGE}}$	160°
$\alpha_{0_{GET}}$, $\alpha_{0_{ETT}}$, $\alpha_{0_{TTT}}^{saturated}$	180°
$\alpha_{0_{TTT}}^{cis-unsaturated}$	120°
m_C, m_P	90 amu
m_G, m_E	62 amu
m_T	42 amu
m_W	40 amu

Table 2.4: The parameters of the ELBA1.0 forcefield[55], see reference for definition of symbols.

Interaction Type	Description	Functional form
Bonded	Bond-stretching	$k_r(r-r_0)^2$
	Angle bending	$\frac{1}{2}k_\theta(cos\theta - cos\theta_0)^2$
NonBonded	van der Waals	shifted function (see equation 2.13
	Electrostatic	Charge-charge - shifted function (see equation 2.14
		Charge-dipole - shifted function (see equation 2.15)
		Dipole-dipole - switched function (see equation 2.16)

Table 2.5: The functional forms employed in the ELBA1.0 forcefield[55].

$$U_{ij}^{Q\mu} = \frac{Q_i}{4\pi\varepsilon_0 r_{ij}^3}\left[1 - 3\left(\frac{r_{ij}}{r_c}\right)^2 + 2\left(\frac{r_{ij}}{r_c}\right)^3\right]\boldsymbol{\mu}_j \bullet \boldsymbol{r}_{ij} \tag{2.15}$$

where Q_i and Q_j are the charges on the 2 particles, $\boldsymbol{\mu}_i$ and $\boldsymbol{\mu}_j$ the dipole vectors on the particles, r_{ij} he distance between them and ε_0 the permittivity of vacuum.

The dipole-dipole interactions are slightly different. A switching function (equation 2.16 is used which alters the form of the potential when the interaction distance is less than the cut off (r_c) but greater than the switching distance (r_s):

$$U_{ij}^{\mu\mu} = \frac{1}{4\pi\varepsilon_0}\left[\frac{\boldsymbol{\mu}_i \bullet \boldsymbol{\mu}_j}{r_{ij}^3} - 3\frac{(\boldsymbol{\mu}_i \bullet \boldsymbol{r}_{ij})(\boldsymbol{\mu}_j \bullet \boldsymbol{r}_{ij})}{r_{ij}^5}\right]S_{ij} \tag{2.16}$$

$$S_{ij} = \begin{cases} 1 & r_{ij} \leq r_s \\ \frac{(r_c - r_{ij})^2(r_c + 2r_{ij} - 3r_s)}{(r_c - r_s)^3} & r_s < r_{ij} < r_c \end{cases}$$

The point dipoles embedded in the glycerol and ester sites are also subjected to a potential that restrains their orientation relative to the lipid molecule to which they belong. In particular, the glycerol dipole is restrained to lie along the direction of the bond vector going from the glycerol to the phosphate site, whereas the ester dipoles are restrained to the bond vectors going from the ester to the adjacent tail site of the corresponding tail. γ is defined as the angle between the dipole vector and its "reference" bond vector and the following potential is used:

$$U = \frac{k_{dip}}{2}(cos\gamma - 1)^2 \tag{2.17}$$

The forcefield was updated to ELBA1.1 in 2013[57] to include a shifted form of the dipole-dipole interaction in the form in equation 2.18. This affected the water bulk-properties and lipid head group areas[55,57] and therefore necessitated a re-parameterisation of the water model in particular. The functional forms and parameters of ELBA1.1 are set out in tables 2.6 and 2.7.

$$U^{\mu\mu} = \left[1 - 4\left(\frac{r}{r_c}\right)^3 + 3\left(\frac{r}{r_c}\right)^4\right]\left[\frac{1}{r^3}(\boldsymbol{\mu}_i \bullet \boldsymbol{\mu}_j) - \frac{3}{r^5}(\boldsymbol{\mu}_i \bullet \boldsymbol{r})(\boldsymbol{\mu}_j \bullet \boldsymbol{r})\right] \tag{2.18}$$

All the simulations using ELBA1.1 described in this thesis use a shifted force cut-off of 12 Å which has been used for both LJ and electrostatics. This has been shown to produce trajectories comparable to those using PME methods[58].

ELBA is unique amongst CG force-fields: firstly, ELBA is able to reproduce realistic dynamic behaviour, particularly in relation to lateral diffusion. Simulations of bilayers composed of single species of the phospholipids; DOPC, DOPE and DSPC reproduced realistic physical properties enabling exploration of features including the internal pressure distribution, dipole potential and lipid diffusion[45,55–57]. Notably, the lipid diffusion rate calculated with ELBA matched experimentally determined diffusion considerably better than many other commonly used CG lipid models. Yamashita et al. recently employed the ELBA force-field in equilibrium and non-equilibrium MD and Grand Canonical Monte Carlo (GCMC) simulations to study the thermodynamics and kinetics of capillary evaporation of water within hydrophilic, cylindrical mesopores, and in doing so also demonstrated

Interaction Type	Description	Functional form
Bonded	Bond-stretching	$k_r(r - r_0)^2$
	Angle bending	$\frac{1}{2}k_\theta(cos\theta - cos\theta_0)^2$
NonBonded	van der Waals	shifted function (see equation 2.13
	Electrostatic	Charge-charge - shifted function (see equation 2.14
		Charge-dipole - shifted function (see equation 2.15
		Dipole-dipole - shifted function (see equation 2.18)

Table 2.6: The functional forms employed in the ELBA1.1 forcefield[55]. The highlighted cell indicates a change from ELBA1.0.

the versatility and physical accuracy afforded by the ELBA force field[59]. In addition, Siani et al. have extended the ELBA1.1 parameters to enable simulation of hydroperoxidised phospholipids, important products of environmental damage to membranes implicated in diseases such as Parkinsons and Alzheimers[60]. They reported good agreement with all-atom simulations for electron density profiles and water permeation.

The other major advantage of the ELBA forcefield is that, as is common in atomistic force-fields, the Lennard-Jones interactions between different bead types are treated simply using the standard LB mixing rules (with some scaling adjustment for hydrogen bonding moieties), in contrast to being specifically and individually parameterised, as is necessary for the majority of other commonly used CG models. The important consequence of the direct use of the L-B mixing rules is that ELBA is inherently compatible with standard atomistic force-fields. This means that dual-resolution simulations are straightforward. A bulk of the

σ_{CC}	0.54 nm
σ_{PP}	0.52 nm
σ_{GG}, σ_{EE}	0.46 nm
σ_{TT}	0.45 nm
σ_{WW}	0.305 nm
ϵ_{CC}, ϵ_{PP}	6.0 kJ/mol
ϵ_{GG}, ϵ_{EE}	4.0 kJ/mol
ϵ_{TT}	3.5 kJ/mol
ϵ_{WW}	1.534 kJ/mol
ϵ_{WW}^{tot}	$1.5\epsilon_{WW}$
ϵ_{WP}	$1.3\sqrt{\epsilon_{WW}\epsilon_{PP}}$
ϵ_{WG}	$1.1\sqrt{\epsilon_{WW}\epsilon_{GG}}$
ϵ_{WE}	$1.2\sqrt{\epsilon_{WW}\epsilon_{EE}}$
Q_C	+0.7 e
Q_P	-0.7 e
μ_G	1.6 D
μ_E	2.0 D
μ_W	2.6 D
k_r	1260 kJ/(mol.nm^2)
k_θ	30 kJ/mol
k_{dip}	10 kj/mol
$\alpha_{0_{CPG}}$	115°
$\alpha_{0_{PGE}}$	90°
$\alpha_{0_{GET}}$, $\alpha_{0_{ETT}}$, $\alpha_{0_{TTT}}^{saturated}$	180°
$\alpha_{0_{TTT}}^{cis-unsaturated}$	120°
m_C, m_P	90 amu
m_G, m_E	62 amu
m_T	42 amu
m_W	40 amu

Table 2.7: The forcefield parameters for ELBA1.1. The yellow highlighted parameters are those which are different from ELBA1.0

system is simulated using a CG force-field but a particular portion of it, representing, for example a transmembrane protein or a chemical species crossing the membrane or a membrane constituent itself such as cholesterol, is modelled in detail using an atomistic force-field. These dual-resolution or hybrid models are employed in chapter 5.

2.3 Simulation programs, integrators, thermostats and barostats

Three different software programs for molecular dynamics simulation (MD engines) have been used in this project: BRAHMS (http://code.google.com/p/brahms-md/), LAMMPS (http://lammps.sandia.gov/) and GROMACS. The first was developed specifically for the simulation of membranes modelled with the ELBA forcefield and is a serial code. In contrast LAMMPS is an open source code distributed by Sandia National Laboratory. It has been widely used for simulating soft materials, solid state and CG and continuum systems. The main advantage of LAMMPS over BRAHMS is that the former is capable of running in parallel on multiple CPUs using spatial decomposition and OPENMPI message passing techniques as well as GPU acceleration. All simulations described in this thesis including any CG beads modelled with an ELBA potential were simulated using either BRAHMS or LAMMPS, as specified in the relevant chapter. All other simulations employed GROMACS.

The differences between the 3 MD engines employed in this thesis which are particularly pertinent are the algorithms used for integration of the equation of motion and the available thermostats and barostats.

2.3.1 Integrators

As noted above, an exact analytical solution of Newton's equations of motion is not possible for a multi-body system and therefore numerical methods are employed.

Equation 2.1 is solved numerically by considering the Taylor expansion in a small time increment (the time step):

$$\boldsymbol{r}(t+\delta t) = \boldsymbol{r}(t) + \delta t\boldsymbol{v}(t) + \frac{1}{2}\delta t^2\boldsymbol{a}(t) \tag{2.19}$$

GROMACS uses the leapfrog algorithm[61] to solve the expansion. The leap-frog algorithm uses positions $\boldsymbol{r}$ at time t and velocities $\boldsymbol{v}$ at time $t - \frac{1}{2}\Delta t$. It updates the positions and velocities using the forces $\boldsymbol{F}(t)$ determined at time t using:

$$\boldsymbol{v}(t+\frac{1}{2}\delta t) = \boldsymbol{v}(t-\frac{1}{2}\delta t) + \frac{\delta t}{m}\boldsymbol{F}(t) \tag{2.20}$$

$$\boldsymbol{r}(t+\delta t = \boldsymbol{r}(t) + \delta t\boldsymbol{v}(t+\frac{1}{2}\Delta t) \tag{2.21}$$

In contrast, both BRAHMS and LAMMPS use velocity Verlet[62] for integration of the translational motion. In velocity Verlet, positions $\boldsymbol{r}$ and velocities $\boldsymbol{v}$ at time $\boldsymbol{t}$ are used to integrate the equations of motion; velocities at the previous half timestep are not required:

$$\mathbf{v}\left(t+\frac{\delta t}{2}\right) = \mathbf{v}(t) + \frac{\delta t\mathbf{F}(t)}{2m} \tag{2.22}$$

$$\mathbf{r}\left(t+\delta t\right) = \mathbf{r}(t) + \delta t\mathbf{v}\left(t+\frac{\delta t}{2}\right) \tag{2.23}$$

$$\mathbf{v}\left(t+dt\right) = \mathbf{v}\left(t+\frac{\delta t}{2}\right) + \frac{\delta t\mathbf{f}(t+\delta t)}{2m} \tag{2.24}$$

The velocity Verlet has the advantage that it gives positions and velocities at the same time and does not compromise precision. In contrast the leap frog integrator has the disadvantage that the positions and velocities are not synchronised. Both methods are time reversible and symplectic.

The embedded dipoles central to the ELBA forcefield result in anisotropic orientation-dependent potentials and therefore require that the equations of motion are not only solved in respect of translational motion, but also rotational. The algorithms used for the integration of rotational motion were very different in LAMMPS and BRAHMS. The versions of LAMMPS used at the beginning of this project (up to and including Dec 2013) used the same form of integrator, velocity Verlet, for the rotational motion as they did for the translational motion, save with angular momentum substituted for velocity and torque for force, both evaluated in the lab/space frame rather than the particle's body frame.

Rotational integrators designed for the most general situation ie where the rigid body being rotated does not necessarily have any axis of symmetry, require the rotational motion to be integrated in the body's frame of reference. This is because unless the moment of inertia tensor, I is diagonalised (as it is in the body frame), the relation $L = I\omega$, where L is the angular momentum and ω the angular velocity, does not hold. In addition, the lab frame moment of inertia tensor must itself be integrated each timestep to account for the change due to the body's rotation; the body frame moment of inertia does not change.

Other manipulation/analysis within the MD software, particularly evaluation of forces, requires the use of the lab frame and so a rotational transformation is required between the two frames of reference during the MD timestep. There are various methods used to perform the rotation, for example, Euler angles in rotation matrices, quaternions and the Dullweber, Leimkuhler and McLachlan (DLM) method[63] used in BRAHMS; only the last of these is symplectic and time reversible but the others have also been used[64]. In a system such as in the ELBA forcefield, where it is not necessary for the form of the rotational integrator to be completely generalised because only one unique direction has been inserted into the sphere (the dipole), Allen and Tildesley[64] suggest that it may be possible to avoid a complete transformation into the body frame of reference but nonetheless, the algorithm they suggest for the equivalent system of a rigid homo-diatomic molecule is more complex than the original (over) simplified LAMMPS algorithm.

It was hypothesised that the original LAMMPS algorithm for rotational integration is only becomes equivalent to the DLM algorithm used in BRAHMS in the case where the body to be rotated is spherical and therefore the body frame and the lab frame of reference are equivalent and symmetrical and the moment of inertia tensor is diagonalised in both - this is not the case for ELBA beads with embedded dipoles. The DLM method was therefore included in LAMMPS in versions later than December 2013. The difference between the 2 LAMMPS rotation integrators can easily be seen by considering an NVE simulation of ELBA water - ie a LJ fluid, with each LJ bead containing a point dipole. When using the original integrator, the rotational energy of each bead partially 'leaked' away - energy was being lost from the system in a supposedly NVE ensemble. Only after analysing each bead's rotational energy in each of its 3 body-frame degrees of freedom ie rotation about the dipole vector and the 2 orthogonal vectors, did it become clear that the energy leak was almost exclusively from the rotational degree of freedom about the dipole vector. The issue is not seen in an NVT (or indeed NPT) simulation as the thermostat ensures that each degree of freedom contains the required $\frac{1}{2}k_BT$ of kinetic energy. Once the original integrator was replaced with a more complex DLM integrator, therefore taking account of the translations between frames of reference, the energy leak in the rotational degrees of freedom in the NVE ensemble no longer occurred and energy was conserved.

2.3.1.1 rRespa integrator - dual timestep

A slightly more complicated algorithm, the rRespa multi-timescale integrator[65], is used for the dual resolution simulations described in chapter 5. This means that certain parts or sets of particles in the simulation, for example the atomistic parts of a dual-resolution simulation, are integrated and forces between them calculated at a short timestep, whilst other particles, for example the CG parts, are calculated at a longer timestep. This improves the computational efficiency of a dual-resolution simulation because the majority of the force calculations are computed less often, because the number of CG beads will exceed the number of atoms.

2.3.2 Thermostats

All 3 MD engines used in this thesis use different thermostats for maintaining the simulation system at a constant temperature in the 'production' simulations; GROMACS Nosé-Hoover, BRAHMS Berendsen and LAMMPS Langevin.

For each degree of freedom of each particle, the temperature of the system is related to the time average of the kinetic energy which is given by:

$$\langle \mathcal{H} \rangle_{NVT} = \frac{1}{2} k_B T \tag{2.25}$$

where the kinetic energy Hamiltonian, H, of each degree of freedom of a particle is related to the translational or rotational velocity: $KE_{trans} = 3\left[\frac{1}{2}mv^2\right]$ and $KE_{rot} = 3\left[\frac{1}{2}I\omega^2\right]$.

Using a Berendsen thermostat, the system is coupled to an external heat bath fixed at the desired temperature, T_{bath},[66]. The bath acts as a source of thermal energy, supplying or removing heat from the system as necessary. The particles' velocities are scaled at each timestep such that the rate of change of the temperature is proportional to the difference in temperature between the bath and the system:

$$\frac{dT(t)}{dt} = \frac{1}{\tau}(T_{bath} - T(t)) \tag{2.26}$$

How tightly the bath and system are coupled together is governed by the coupling constant τ. If τ is large, the coupling will be weak, if small, the coupling will be strong.
The scaling factor (λ) for the velocities is thus:

$$\lambda^2 = \frac{dt}{\tau}\left(\frac{T_{bath}}{T(t)} - 1\right) \tag{2.27}$$

If τ is the same as the timestep then the thermostat amounts to a simple velocity rescaling algorithm. A coupling constant such that $\frac{dt}{\tau} \approx 0.075$ has been suggested as an appropriate approach[57]. The disadvantage of the Berendsen thermostat is that does not produce averages in complete accordance with the canonical ensemble - though the average temperature produced is correct, the precise fluctuations of the system temperature are not[67].

In contrast, the Langevin thermostat which is the only thermostat available in LAMMPS for the ELBA forcefield does generate a rigorous canonical ensemble. The motion of the particles is adjusted by modifying the forces (and torques) experienced by the particle. The particles of the system can be considered to be embedded in a continuum of smaller particles. The particles experience a drag force (F_{drag}) due to this fictional continuum proportional to their velocity and to the damping constant *damp*, but also experience random 'kicks' due to the fictional particles bumping into the system particles.

$$\boldsymbol{Force} = \boldsymbol{F_i} + \boldsymbol{F_{drag}} + \boldsymbol{F_{random}} \tag{2.28}$$

$$\boldsymbol{F_{drag}} = -\frac{m}{damp}\boldsymbol{v} \tag{2.29}$$

$$\boldsymbol{F_{random}} \propto \sqrt{\frac{K_B T m}{dt.damp}} \tag{2.30}$$

$\boldsymbol{F}_i$ is the force computed via the usual inter-particle interactions. The $\boldsymbol{F}_{drag}$ and $\boldsymbol{F}_{random}$ terms are added on a per-particle basis. $\boldsymbol{F}_{random}$ is a force due to fictional solvent atoms at a temperature T randomly bumping into the particle. The direction of this force is randomised.

Lastly, the Nosé-Hoover thermostat[68,69] introduces a term representing 'thermal resevoir' into the equations of motion as well as a frictional term ξ[70]:

$$\frac{d^2\boldsymbol{r}_i}{dt^2} = \frac{\boldsymbol{F}_i}{m_i} - \frac{\boldsymbol{p_\xi}}{Q}\frac{d\boldsymbol{r}_i}{dt} \tag{2.31}$$

Where ξ has its own momentum $\boldsymbol{p_\xi}$ and equation of motion. Q is the mass parameter of the thermal reservoir and, with T_0, defines the strength of the coupling. The derivative of $\boldsymbol{p_\xi}$ is calculated from the difference between the current temperature T and the reference temperature T_0:

$$\frac{d\boldsymbol{p_\xi}}{dt} = T - T_0 \tag{2.32}$$

The Nosé-Hoover thermostat does correctly reproduce the canonical ensemble but the relaxation of the temperature takes longer than the other methods described and can result in oscillatory behaviour[71] - hence being reserved for production simulations only in this thesis.

2.3.3 Barostats

Both BRAHMS and LAMMPS employ a Berendsen barostat[66] to regulate the pressure of the simulation system whereas the GROMACS simulations employ a Parinello-Rahman barostat for the production simulations.

The Berendsen barostat maintains the pressure by adjusting the simulation box dimensions. It is analogous to the Berendsen thermostat described above, in that the system, with pressure $P(t)$ is coupled to a pressure ‘bath’ with reference pressure P_{bath}, by a coupling constant τ_P. The volume of the simulation box is scaled by a factor λ:

$$\lambda = \left(1 + \frac{dP}{\tau_P}(P(t) - P_{bath})\right)^{\frac{1}{3}} \tag{2.33}$$

The barostat equivalent of the Nosé-Hoover thermostat is the Parinello-Rahman barostat[72]. The correct NPT ensemble is returned by this barostat but, like the Nosé-Hoover thermostat, it takes longer than other barostats to reach the correct pressure and, if started too far away from the reference pressure, it can lead to oscillatory behaviour. The Parinello-Rahman barostat changes the equation of motion:

$$\frac{d^2\boldsymbol{r}_i}{dt^2} = \frac{\boldsymbol{F}_i}{m_i} - \boldsymbol{M}\frac{d\boldsymbol{r}_i}{dt} \tag{2.34}$$

where:

$$\boldsymbol{M} = \boldsymbol{b}^{-1}\left[\boldsymbol{b}\frac{d\boldsymbol{b}'}{dt} + \frac{d\boldsymbol{b}}{dt}\right]\boldsymbol{b}'^{-1} \tag{2.35}$$

$\boldsymbol{b}$ is a matrix representing the box vectors and is calculated:

$$\frac{d^2\boldsymbol{b}}{dt^2} = V\boldsymbol{W}^{-1}\boldsymbol{b}'^{-1}(\boldsymbol{P} - \boldsymbol{P}_{ref}) \tag{2.36}$$

where V is the box volume, $\boldsymbol{W}$ is a matrix parameter determining coupling strength and $\boldsymbol{P}$ and $\boldsymbol{P}_{ref}$ refer the the current and reference pressure matrices. In GROMACS, the inverse of $\boldsymbol{W}$ is defined as:

$$(\boldsymbol{W}^{-1})_{ij} = \frac{4\pi^2\beta_{ij}}{3\tau_p^2 L} \tag{2.37}$$

where β is the isothermal compressibility, τ_p is the pressure time constant and L is the largest box matrix element.

2.3.4 Boundary conditions

Despite best efforts in coarse graining, increased computational efficiency etc, the number of atoms/particles simulated in a molecular dynamics simulation will always be very small compared to Avogadro's number or to the numbers of atoms/molecules involved in physical experiments. If the simulation were to be enclosed by boundaries, a significant proportion of the particles would be subjected to edge effects which would not represent bulk properties. All the MD simulations in this thesis therefore employ periodic boundary conditions (PBC) whereby an infinite number of replicas of the systems surround the original and as each particle moves out of the domain from one side, its copy moves in from the other. Figure 2.6 shows a 2-D representation of PBC, the original system in the centre. The dotted circle represents a cut off distance and illustrates that a particle can be counted as interacting with a copy; the 'minimum image convention' is employed - each particle interacts with the nearest image of any other particle.

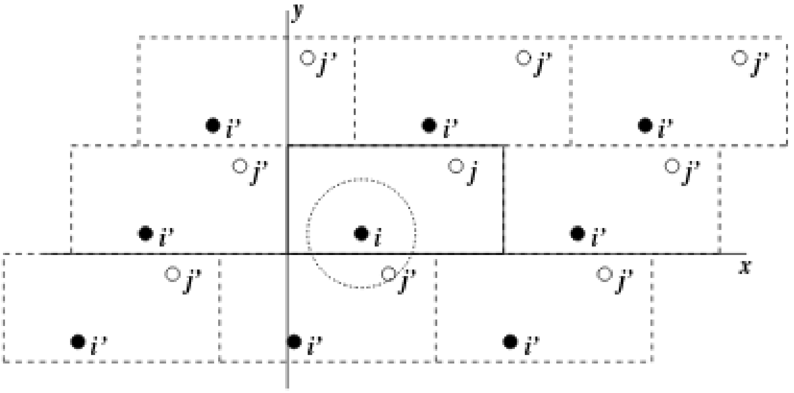


Figure 2.6: Periodic Boundary conditions in 2 dimensions[70]

2.4 Summary

In MD simulations such as those employed in this thesis, atoms (or groups of atoms) are modelled as spheres connected by bond interactions rather than there existing separate entities called 'bonds'. Bonding interactions are modelled as springs and non-bonded interactions via simple Lennard-Jones and electrostatic potentials. The charges on atoms or the magnitude of dipoles do not change during the simulations and bonds (or bonding interactions) are not created or destroyed. The movement of these atoms is calculated from the set of forces on them. These forces are derived from the multi dimensional potential energy landscape and solely depend on the atoms' position in space. This potential energy landscape is defined by the parameters of the forcefield in use which may have been derived from fitting to experimental measurements of macro-properties, *ab initio* quantum calculations or a combination of the 2. Having defined the forces on the atoms, Newton's equations of motion are used to calculate their motions at each timestep. Further, thermostatting and barostatting algorithms are applied in order to match the systems' behaviour to that observed in the NVT or NPT ensembles in the laboratory.

Chapter 3

Phase Transitions

The stable phase of a lipid aggregate is not only defined by the physical characteristics of the lipid molecules but also by the temperature and hydration level of the system. The structure can undergo both thermotropic - temperature dependent - and lyotropic - hydration dependent - transitions between phases. In particular the level of hydration as an important factor in the phase behaviour of lipids is often overlooked, as Rand and Parsegian have noted "...it [is] too easy to forget about the role of water in the polymorphism of lipid assemblies, too easy to think of it simply as providing the right environment for amphiphilic molecules to assemble themselves..."[73]. Phase transitions at low hydration levels are of particular relevance in fields such as cryobiology as much of the damage done to biological material on freezing is not due to the low temperature per se but the reduction in the amount of liquid water present[74].

During lipid bilayer phase transitions the degree of order of the lipids' hydrocarbon chains inside the bilayer changes. For instance, phosphatidylcholines exhibit the following bilayer phases, increasing in order; a fluid, liquid crystal phase L_α, two gel phases $P_{\beta'}$ (ripple phase) and $L_{\beta'}$ and a crystalline phase L_c. The gel phases differ from the fluid or liquid-crystalline phase by a number of key features[75,76]. In the gel phase: (i) the area per lipid is lower; (ii) the lipid tails are almost fully extended; (iii) the lipid head groups are hexagonally ordered in the plane of the bilayer; and (iv) the lipids' lateral mobility is greatly reduced. The difference between the tilted and untilted gel phase is the presence of an average tilt of the lipid tails with respect to the bilayer normal. At even lower temperatures, most phospholipids adopt a crystal phase[75]. In the crystal phase,

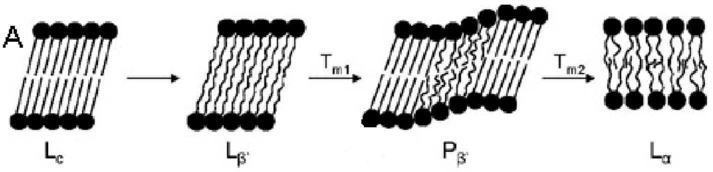


Figure 3.1: Schematic representation of bilayer phase transitions.

the lipids are fully ordered and virtually immobile. Figure 3.1 shows the bilayer transitions of a typical phospholipid with increasing temperature and figure 3.2 shows the phase diagram for DMPC. The different areas of the diagram show that the phases adopted by a DMPC aggregate depend both on hydration level as well as temperature.

The temperature at which the main phase transition occurs depends on the headgroup of the lipid and also on the length of its tails[77]. As a first order transition, the Gibbs free energy is continuous but there are discontinuous changes, ΔH_t and ΔS_t in enthalpy and entropy at transition (transition values indicated by the subscript ‘t’).

$$\Delta G_t = \Delta H_t - T_t \Delta S_t = 0 \tag{3.1}$$

where T_t is the transition temperature. Therefore:

$$T_t = \frac{\Delta H_t}{\Delta S_t} \tag{3.2}$$

For lipids with 2 saturated, symmetric tails ΔH_t and ΔS_t depend linearly on chain length[78,79]:

$$\Delta H_t = \Delta H_{inc}(n - n_H) \tag{3.3}$$

$$\Delta S_t = \Delta S_{inc}(n - n_S) \tag{3.4}$$

where ΔH_{inc} and ΔS_{inc} are the incremental transition enthalpy and the incremental transition entropy per CH_2 group. n_H and n_S are the chain lengths for which the transition enthalpy and transition entropy respectively extrapolate to zero

Lipid name	length of tails	calculated T_m (K)	measured T_m[83] (K)
DMPC	14	297.4	297.1
DPPC	16	315.6	314.6
DSPC	18	329.1	328.45

Table 3.1: The T_m for the saturated lipids considered in this thesis calculated by equation 3.5 and measured experimentally.

and therefore $\Delta H_{inc} n_H$ and $\Delta S_{inc} N_S$ account for all contributions from the head group and the terminal methyl group.

Combining these equations leads to:

$$T_t(n) = T_t^{\infty}(1 - \frac{n_H - n_S}{n - n_S}) \tag{3.5}$$

Where T_t^{∞} is the transition temperature extrapolated to infinite chain length.

For fully hydrated PC lipids with 2 saturated symmetric tails 12-24 carbons long $T_m^{\infty} = 420.8 \pm 4.1$K, $n_H - n_S = 3.39 \pm 0.24$, $n_S = 2.44 \pm 0.55$[80–82] which yields the calculated T_m for the saturated PC lipids discussed in this thesis and tabulated in table 3.1.

The T_m for phospholipids with a single double bond in either/both tails is lower and the position of the double bond also affects T_m. For otherwise identical PC lipids, the T_m falls to a minimum as the position of the double bond approaches the centre of the tail(s) and rises again as it moves past the centre[77]. The measured T_m for DOPC and POPC are 255.85 K[84] and 269.75 K[85] respectively.

Lipid phase behaviour of fully hydrated lipid bilayers has been explored experimentally, and both atomistic and CG simulations have also been used to explore thermotropic phase transitions[36,87–89]. However, there have been very few simulation studies seeking to observe lyotropic phase transitions. MD studies in which the hydration of bilayers is varied have mainly focused on the existence and origin of the hydration forces which act at short range to repel hydrated lipid aggregates from one another[2]. CG models have been used extensively[90,91] in modelling membrane phases and have shown good semi-quantitative agreement with experiment. However such simulations have tended to start with blocks of lipids in the 'correct' phase for the temperature and hydration to investigate domain separation rather than

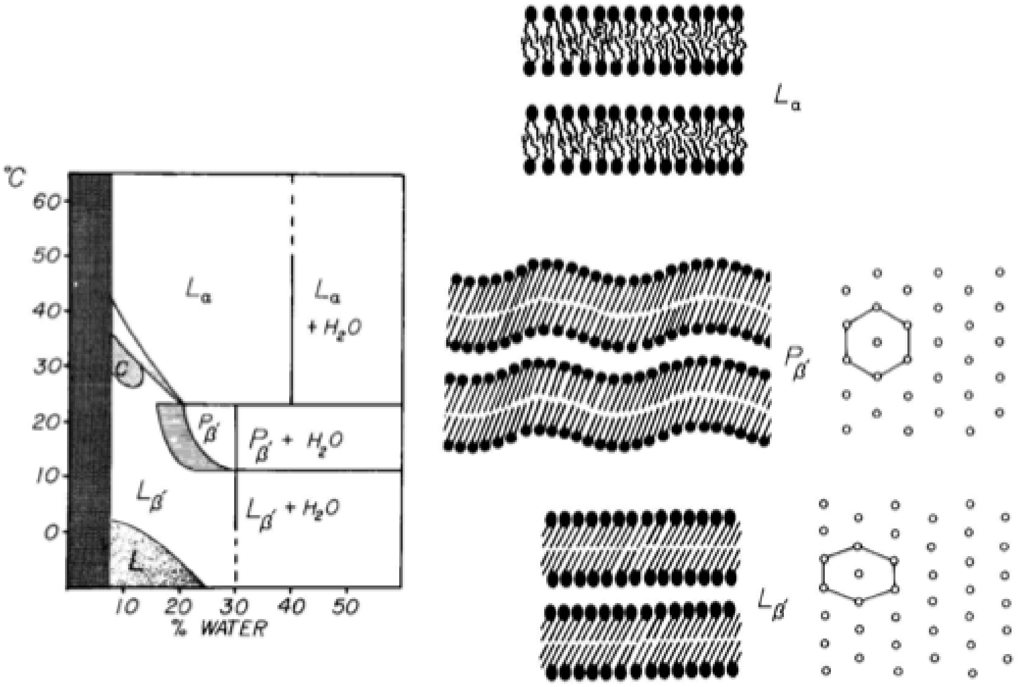


Figure 3.2: Phase diagram of hydrated DMPC bilayers, together with representations of the L_α, $P_{\beta'}$ and $L_{\beta'}$ phases. The hydrocarbon chain packing is a hexagonal array for the P_β phase and a "distorted" hexagonal lattice for the L_β phase. Reproduced from Janiak et al[86].

phase transitions.

This chapter explores the phase diagram of DSPC using the ELBA forcefield, considering both the effects of temperature and the effects of hydration on its phase behaviour.

At the time of initiating this project, DSPC was the only saturated PC lipid parameterised in the ELBA forcefield. It was not possible to find a completely experimentally determined phase diagram for DSPC to compare against our simulation results, though it has been proposed[86] that the phase diagram for DSPC follows a similar pattern as seen in other saturated di-acyl phosphatidylcholines; the phase boundaries are a similar shape but are shifted to higher temperatures as the length of the saturated acyl tails increases. Thus DMPC, 14 carbons per tail, has been found to have a lower main phase transition temperature than DPPC, 16 carbons per tail. As discussed in chapter 1, both DMPC and DPPC have a similar structure to DSPC save with shorter tails. The phase diagram for the saturated di-acyl phosphatidylcholines proposed by Janiak et al.[86] is set out in figure 3.3. The solid horizontal line to the right, representing the phase change at constant temperature and increasing hydration, was experimentally determined for DSPC, the dotted sloping phase boundary to left, representing the increasing phase transition temperature at lower hydration, is hypothesised based upon the experimentally determined equivalent phase boundaries for shorter lipids.

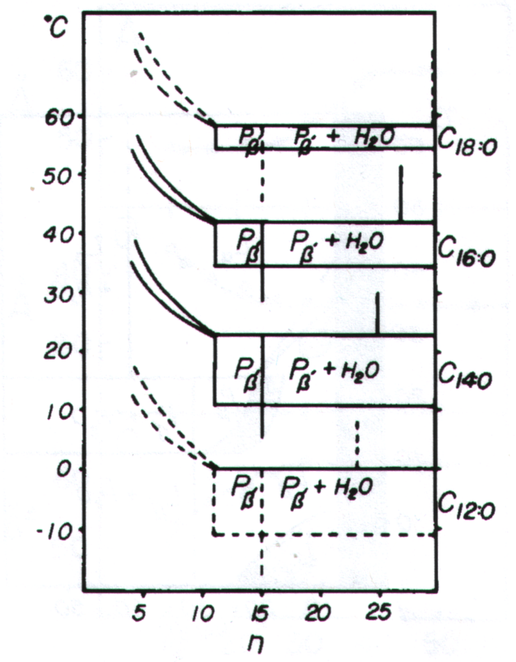


Figure 3.3: This phase diagram shows the increasing transition temperature for PC lipids as acyl chain length is increased and also increasing transition temperature at reduced hydration levels. Reproduced from Fig.12, Janiak et al (1979)[86]

As can be seen from figure 3.3, at low hydration levels, the temperature of the phase transition from gel (L_β) phase to liquid crystalline (L_α) increases from the fully-hydrated T_M of 336 K[92]. Full hydration of a DSPC bilayer is at a level of 33 water molecules per lipid molecule (wpl)[92].

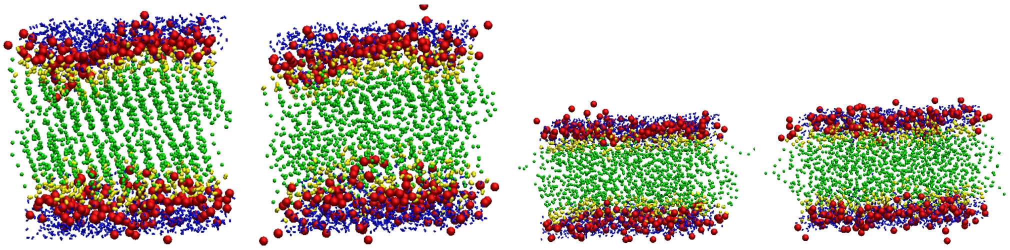
3.1 DSPC phase transition simulated using BRAHMS

Given the success of the ELBA forcefield in reproducing other membrane properties described in chapter 2, we decided to investigate its ability to reproduce phase behaviour, both lyotropically and thermotropically driven, in the fully saturated lipid (DSPC) already parameterised in ELBA.

3.1.1 Method

Forty simulations of 128 DSPC lipids modelled using the ELBA1.0 forcefield were run using the BRAHMS[93] MD engine at various levels of hydration between 3 and 33 waters per lipid and various temperatures between 50°C (323 K) and 78°C (351 K) to investigate whether the gel (L_β) to liquid crystalline (L_α) phase transition was reproduced using the ELBA1.0 forcefield. Each simulation was initialised with the lipids arranged in a regular grid in a bilayer parallel to the xy plane of the simulation box, 64 lipids per leaflet, the x and y dimensions of the simulation box both being 5.50 nm, thereby giving an average area per lipid of 47.3 Å^2 which is that associated with the gel phase of this lipid.[94]. Temperatures and pressures were controlled using a Berendsen thermostat, coupling constant, τ_T, of 200 fs, and a semi anisotropic barostat, coupling constant, τ_P, of 500 fs, respectively. All other forcefield parameters were set in accordance with the standard ELBA1.0 parameters set out in table 2.4. The systems were each simulated for 10000 steps in the NVT ensemble with increasing timesteps (0.015 fs to 1.5 fs) and then the NPT ensemble for 10000 steps with a timestep of 15 fs. After this equilibration phase, a 'production phase' was conducted for each in the NPT ensemble for a minimum of 200 ns.

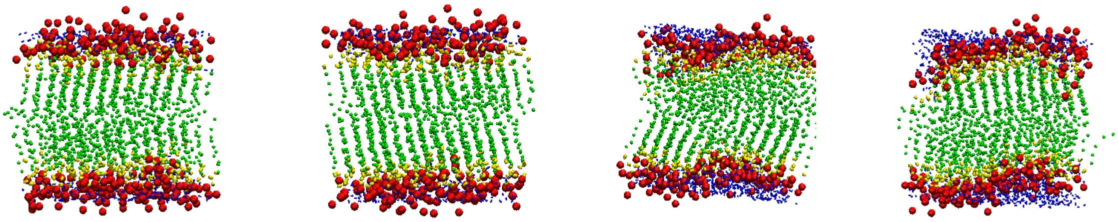
Snapshots of the simulation at 63°C (336 K) with 10.56 waters per lipid are shown in figure 3.4 as a representative example of the phase transition which was observed.



(a) 0 ns (b) 158 ns (c) 170 ns (d) 250 ns

Figure 3.4: A series of snapshots from the final "production" NPT simulation with 10.56 waters per DSPC at 63°C (336 K) showing a phase transition between 158 ns and 170 ns from a gel phase to a liquid disordered phase. The lipid head groups are coloured red, the glycerol and ester beads yellow, tails green and the waters blue. The thickness of the bilayer reduces from and average P-P distance of approximately 48 Åin the gel phase to approximately 38 Åin the liquid disordered and the tails become visibly less aligned.

For comparison, snapshots from a lower hydration at the same temperature are shown in figure 3.5. These show no phase transition.



(a) First frame of 3.0 wpl production run

(b) Last frame of 3.0 wpl production run

(c) First frame of 6.89 wpl production run

(d) Last frame of 6.89 wpl production run

Figure 3.5: Four snapshots from production run of DSPC at 336 K. The bilayer is in the $L_{\beta'}$ phase for the whole of both trajectories, as can be seen from the ordering and tilt of the acyl tails

The images taken from the trajectories also show that, in the gel phase, the acyl chains are not only ordered but also spontaneously assume a tilt with respect to the bilayer normal. In accordance with the theoretical and experimental data, the phase adopted is a $L_{\beta'}$ rather than a L_{β}. This is an improvement over the only other coarse grained study of phase transitions which used the MARTINI force field to investigate thermotropic phase transitions of DPPC and found an untilted gel L_{β} phase rather than the actually experimentally observed tilted gel $L_{\beta'}$[88]. This problem was reported as being improved by reducing the size of the tail sites, but this reparameterisation was reported to have knock-on effects on other properties. In contrast, ELBA1.0 accurately reproduces the tilt of the gel phase without special reparameterisation.

Confirmation that a phase change occurs in the bilayer for the 10.56 waters per lipid system can be seen by considering both the area per lipid and the tail order parameter. As can be seen in figure 3.6 the area per lipid suddenly jumps from 49 Å^2, in good agreement with the experimental value for the area per lipid of the gel phase of ∼47 Å^2[94] to ∼70 Å^2 which again agrees well with the experimentally determined value for the liquid crystal phase of ∼65 Å^2[95].

To quantify tail ordering, the segmental order parameter is calculated[55]:

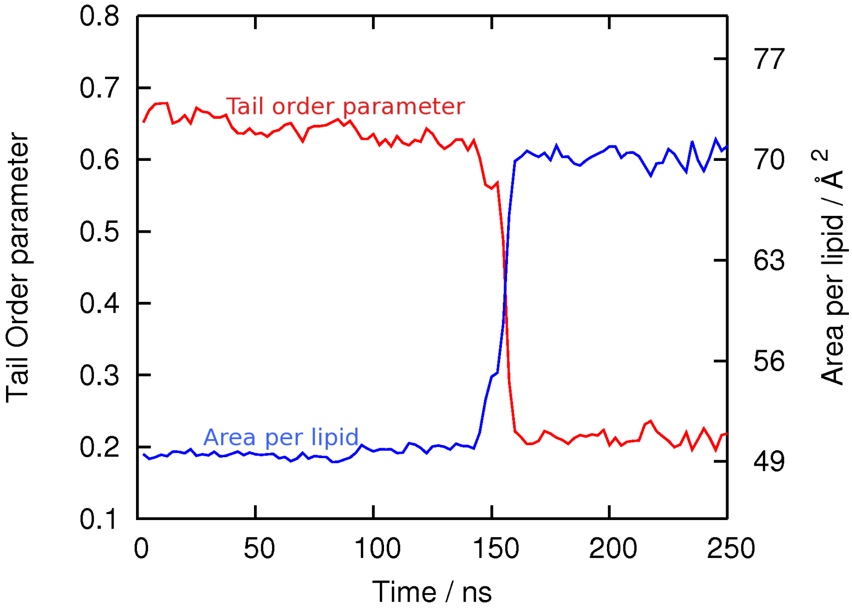


Figure 3.6: A graph showing both the area per DSPC lipid and the tail order parameter changes, during the production run of 10.56 waters per lipid at 63°C (336 K). The area per lipid suddenly increases from that expected for the gel phase to that expected for the liquid crystal phase at the same point that the tail order parameter indicates a sharp increase in the disorder in the acyl tails.

$$S^k_{\mathrm{mol}} = \langle 3\cos^2\eta - 1\rangle/2 \tag{3.6}$$

with η the instantaneous angle between the k-th bond along the tail and the direction normal to the bilayer plane (the z-axis). The angular brackets indicate averaging over the simulation time. By further averaging over the four bonds in each lipid tail, and then over all lipid molecules, a global tail order parameter was obtained which is plotted in Figure 3.6. By definition, $-0.5 \leq S_{\mathrm{mol}} \leq 1$; in particular, $S_{\mathrm{mol}} = -0.5$ indicates alignment parallel to the bilayer plane, $S_{\mathrm{mol}} = 0$ indicates random orientation and $S_{\mathrm{mol}} = 1$ indicates alignment parallel to the normal to the bilayer plane. As can be seen in figure 3.6, the global tail order parameter at the start of the simulation is well over 0.5, indicating a high degree of order in the tails and, at the same time the area per lipid undergoes a dramatic increase, the global tail order parameter falls to approximately 0.2; indicating substantial disorder (though not complete random organisation).

The results of all 40 simulations are summarised in figure 3.7 which shows the phase of the bilayer at the end of the production simulation, blue for gel and red for L_α. The phases were assigned by averaging the area per lipid over the last 10 ns of the production phase of the simulation, if the resulting area was less than 52 Å^2 (being the starting (gel) area per lipid + 10%) the bilayer was labelled as gel phase, if more than 60 Å^2 (being the experimentally determined area per lipid for the liquid disordered phase − 10%) the bilayer was labelled as liquid disordered phase. No simulation resulted in a bilayer with an area per lipid between these two bounds. The phase transitions simulated show remarkable agreement with the experimentally determined/estimated phase boundaries[86] also plotted on figure 3.7.

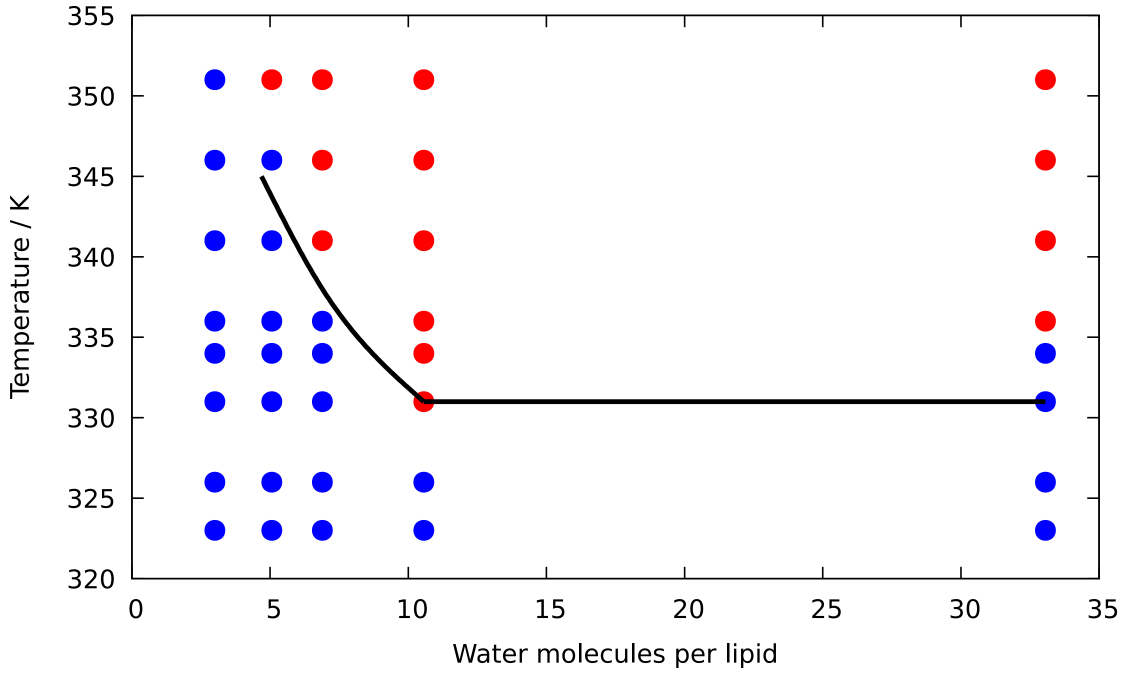


Figure 3.7: A graph summarising the results of all 40 simulations of DSPC run using BRAHMS. A blue label represents a simulation where the bilayer remained in a gel phase, the red labels show those simulations where the bilayer 'melted' to a liquid disordered phase. The black line represents the phase boundaries in the experimentally determined phase diagram[86].

3.2 DSPC phase transition simulated using LAMMPS

A subset of the simulations described above were rerun using the ELBA1.0 forcefield in LAMMPS (Jan 2013). The initial starting structures for these simulations were generated using BRAHMS using the same method described in section 3.1.1 and converted to LAMMPS input format using in-house scripts. The same simulation parameters were used as for the simulations using BRAHMS, save that, as described in chapter 2, the Langevin thermostat was used in the place of the Berendsen thermostat, with the lipids and solvent thermostatted separately with a damping parameter of 1000 fs. The Berendsen barostat was used with a pressure of 1 atmosphere with semi anisotropic coupling. The results are shown in figure 3.8 which was plotted using the same protocol as that used for assigning phases to the points in figure 3.7. The same phase boundaries, both lyotropic and thermotropic were found using LAMMPS.

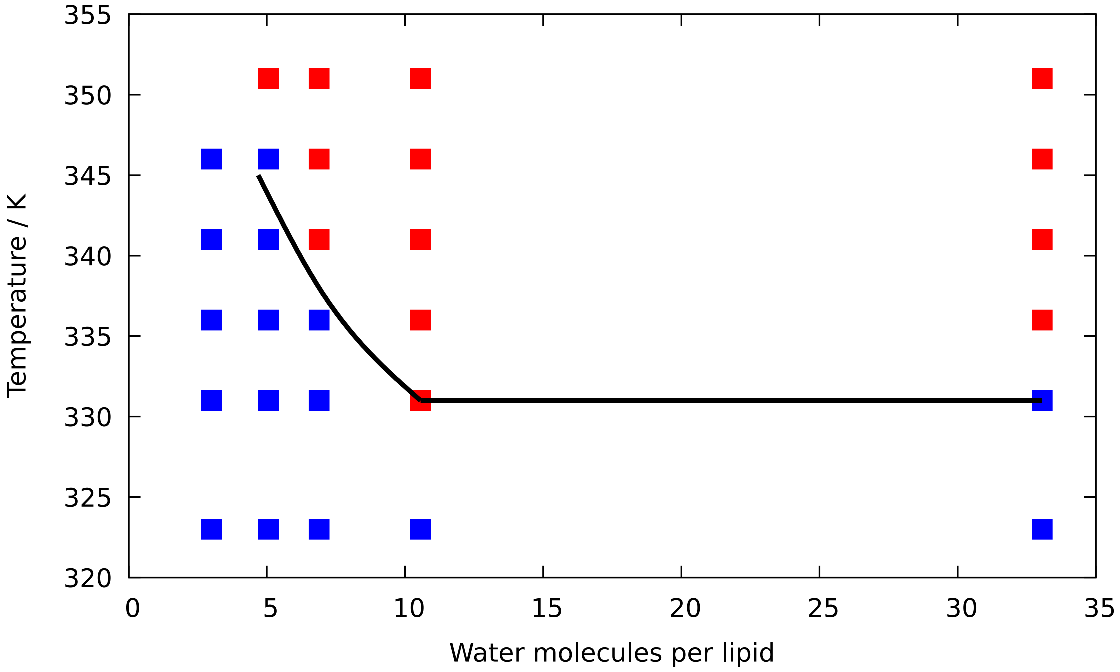


Figure 3.8: A summary of the results of 29 simulations of DSPC run using the ELBA1.0 forcefield in LAMMPS. A blue label represents a simulation where the bilayer remained in a gel phase, the red labels show those simulations where the bilayer 'melted' to a liquid disordered phase. The black line represents the phase boundaries in the experimentally determined phase diagram[86].

LAMMPS simulations using bilayers created in the same way were also conducted using the ELBA1.1 forcefield. Using the same criteria to assign the phases, none of the simulated bilayers using ELBA1.1 remained in the gel phase even during equilibration; even those simulated at the lowest hydration and temperature combination transitioned to L_α before the production phase of simulation had begun. Despite ELBA1.0 and ELBA1.1 being very similar forcefields (as described in further detail in chapter 2) the differences between them obviously have profound effects on the phase behaviour observed in their simulations.

In summary, for those simulations using the ELBA1.0 forcefield and initialised in the gel phase, the simulations provided data in good agreement with the available experimental data. The temperature and hydration of the phase change achieved was also in good agreement with the location of the phase boundary measured and extrapolated from experimental data.

To check for hysteresis in the phase transitions, the reverse phase transitions ie L_α to L_β must be observed. We therefore set up simulations using the output of the high temperature/high hydration simulations above ie those which had transitioned to the L_α phase, removed some of the water beads and reran the same protocols. Despite running a series of ELBA1.0 simulations at hydrations as low as 3 waters per lipid at 323 K for 500 ns each, no phase change was observed in the other direction, ie L_α to $\mathrm{L}_{\beta'}$. Those simulations initialised in the L_α phase remained in that phase even when simulated at reduced hydration and reduced temperature.

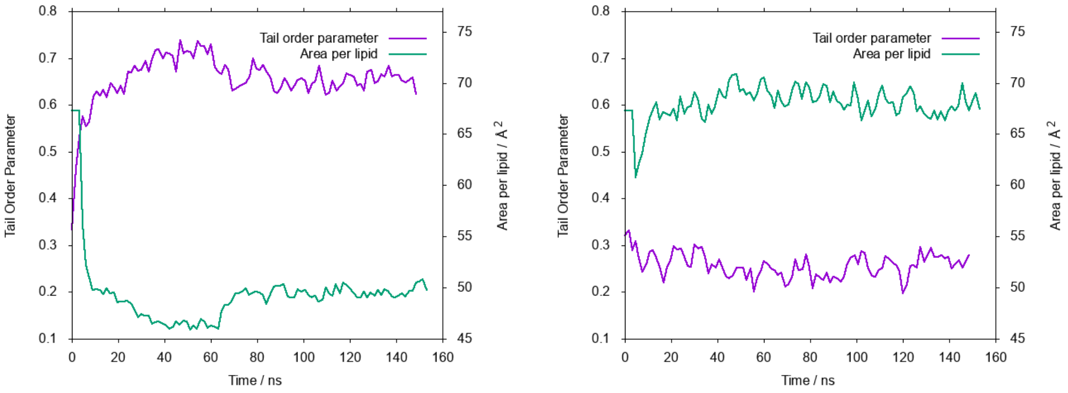
This hysteresis in our simulated phase transitions may be due to insufficient sampling of the energy and conformational landscape in the relatively short simulation time. Molecular simulations of crystallizations in general are considered to be difficult, primarily due to complicated potential energy landscapes and long time scales required to sample them effectively[96–98]. Nucleation of the more ordered phase is a particular problem. However, the inability to simulate a phase transition to the gel phase may also be due to the gel phase conformation not being thermodynamically stable in the ELBA1.0 forcefield at all, and its persistence in low temperature simulations may merely be because transitions out of it are kinetically hindered. The gel phase conformation in the ELBA1.0 forcefield may actually be a metastable state rather than an accurate representation of the thermodynamically stable gel phase of DSPC.

Marrink et al. conducted simulations exploring the gel to L_α and vice versa, phase transitions in DPPC bilayers modelled using the MARTINI CG forcefield[88].

They found that the critical step in the phase transition was the formation of a '20 - 80 gel cluster spanning both monolayers'. To investigate the possible requirement for nucleation to initiate a L_α to gel phase change using the ELBA1.0 forcefield, we constructed 2 'mixed phase' bilayers. They were constructed by simply slicing the bilayer in half. The first was constructed using coordinates of the 64 lipids with the lowest value x-coordinates from the 'gel-annealed' bilayer (prepared in accordance with the protocol described in section 3.5) and the 64 lipids with the highest value of x-coordinates from the 'L_α-annealed' bilayer (prepared in accordance with the protocol described in section 3.5), and concatenating them to prepare a new bilayer, ensuring that there were approximately equal numbers in each phase in each leaflet. The second was similarly prepared using the other half of the lipids from each annealed bilayer. These new mixed bilayers was equilibrated using LAMMPS; firstly, a 10^7 step energy minimisation was conducted to attempt to overcome any bead clashes caused by the construction method, followed by 0.5 ns in total of simulation in the NVE ensemble at timesteps increasing from 0.0015 fs to 15 fs. Followed by 0.5 ns NVT simulation employing separate Langevin thermostats for the lipids and water at a temperature of 323 K and finally, the production portion of the simulation, 150 ns NPT using a semi anisotropic Berendsen barostat to maintain the pressure at 1 atmosphere.

The results of these simulations can be seen in figure 3.9. In the first case (figure 3.9a), the area per lipid and average tail order parameter values at the beginning of the simulation are, not unexpectedly, between those associated with either the gel or the L_α. Within 20 ns of the start both have adjusted to values associated with the gel phase - and area of 49.2 Å^2 per lipid and an average tail order parameter of 0.64. This means that the lipids which were in the L_α phase at the start of the simulation, have ended it in the gel phase and provides evidence that transition to the gel phase requires nucleation to initiate rather than not being observed because the gel phase is only metastable. However, the other simulation (figure 3.9b), the area per lipid and tail order parameter adjusted to values more associated with the L_α phase, 67.6 Å^2 per lipid and 0.35 respectively. This indicates that even with lipids already in the gel phase in close proximity to provide nucleation for the phase transition and the temperature and hydration level set to favour the gel phase, nonetheless the lipids in the gel phase 'melted' to the L_α.

In order to further investigate the energy landscape of the transition and thus gain more evidence that the gel 'phase' found in the ELBA1.0 simulations is a genuinely thermodynamically stable phase rather merely a metastable artefact



(a)

(b)

Figure 3.9: Graphs showing both the area per DSPC lipid and the tail order parameter changes, during two production runs of 10.56 waters per lipid at 50°C (323 K). The initial configurations of the bilayers were different from one another. Each was constructed from 64 DSPC lipids in the gel phase and 64 in the L_α phase.

of the forcefield, we employed the enhanced sampling techniques, umbrella sampling and metadynamics, described below. Enhanced sampling techniques were also employed on the ELBA1.1 simulations to attempt to quantify the effect on the free energy landscape of the seemingly small differences between ELBA1.0 and ELBA1.1.

3.3 Metadynamics theory

Metadynamics has been described as "filling the free energy wells with computational sand"[99]. If a system can be described by some "collective variables" (CVs), the location of the system on the free energy surface defined by those CVs is found and a Gaussian potential is added to the energy surface during the molecular dynamics simulation[100]. This modifies the free energy surface along which the system is evolving. As Gaussians continue to be added, the system is increasingly discouraged from returning to its previous states and the sum of the Gaussians 'fills up' the energy basin, which may be a metastable state in which the system may have become trapped, and the system is then free to enter another energy basin. Eventually, all energy basins in the landscape are explored. Because a record is kept of the added Gaussians, the shape of the original energy landscape can be recovered by considering the negative of the sum of added Gaussians, much as plaster of paris poured into a mould provides a negative relief of that mould once the mould has been gradually filled and set.

More formally, the external history-dependent bias potential, $V(\overrightarrow{s,t})$ is constructed in a space defined by selected degrees of freedom $\vec{s}(q)$ known as CVs.

$$V(\overrightarrow{s,t}) = \sum_{k_B\tau<t} W(k_B\tau)exp\left(-\sum_{i=1}^{d}\frac{(s_i - s_i(q(k\tau)))^2}{2\sigma_i^2}\right) \tag{3.7}$$

where τ is the deposition stride, ie how frequently Gaussians are added during the simulation, σ_i is the width of the Gaussian for the i^{th} CV, d is the total number of CVs and $W(k\tau)$ is the height of the Gaussian.

τ, σ_i and $W(k\tau)$ can all be varied to tune the accuracy of the description of the energy landscape balanced with the computational cost of the simulation; large Gaussians can be used to quickly obtain a rough description of the surface whereas smaller Gaussians will yield more detail but take more computational time to converge[100].

In standard metadynamics, Gaussians of constant height are added for the entire course of a simulation. As a result, the system is eventually pushed to explore high free-energy regions and the estimate of the free energy calculated from the bias potential oscillates around the real value. In well-tempered metadynamics[101], the height of the Gaussian is altered adaptively during the course of the simulation. The Gaussian height is decreased with simulation time according to:

$$W = W_0exp\left(-\frac{V(s(\vec{q}(k\tau))}{k_B\Delta T}\right) \tag{3.8}$$

Where W_0 and W are the initial and current Gaussian heights and ΔT is a user defined input with the dimension of temperature. With this rescaling of the Gaussian height, the bias potential smoothly converges in the long time limit, but it does not fully compensate the underlying free energy because the ensemble sampled is at a temperature higher than the system temperature ie $T+\Delta T$. Setting $\Delta T = 0$ corresponds to standard MD and $\Delta T \rightarrow \infty$ to standard metadynamics. In PLUMED[102] the user input 'bias factor' (γ) is used to define ΔT.

$$\gamma = \frac{T + \Delta T}{T} \tag{3.9}$$

The dimensionality of the energy landscape is equal to the number of CVs which are biased. It is recommended that fewer than 4 CVs are biased in any metadynamics simulation[102]. This is because, assuming the CVs are not correlated with one another, the time required to fill the landscape increases exponentially with the number of biased CVs.

Selecting which aspects of the system to define as CVs and use to bias the simulations is difficult. The whole phase space must be projected onto the CV space with the requirement that none of the fundamental dynamics of interest of the system are lost.

Appropriate CVs for biasing simulations must therefore fulfil the following conditions: firstly that the CV(s) must distinguish between the initial and final states and all intermediates; and secondly, that the CVs must describe all the slow degrees of freedom relevant to the process being studied.

The first requirement is obviously necessary. The necessity of the second can be seen by considering the effect if it is neglected; the selected CV can become trapped in a small region of CV space during the simulation. The system can then ‘jump’ from one energy basin to another in the neglected slow degree of freedom which means that the biased CVs can not access the regions of phase space previously visited. Freely diffusive behaviour in the CV space is a necessary, but insufficient, condition to confirming that the CV selected is appropriate. Trapping in an unbiased slow degree of freedom can also be indicated if the height of the bias applied (W) does not smoothly decay to a small value during the course of the well-tempered metadynamics simulation but instead decays and increases and the system becomes trapped in energy basins defined by unbiased CVs.

Choice of CVs for any particular system relies on chemical intuition and knowledge of the system/process and the change in physical properties. Designing a CV for better sampling commonly requires a certain amount of trial-and-error.

Even having selected one ‘good’ CV, it may be the case that more than one is required to effectively sample the dynamics of the system. This then raises additional consideration that there is likely to be correlation between the two

CVs in so far as the information of some physical aspects of the system each contains may overlap. This overlapping information will therefore be weighted more heavily in the effective bias, having been biased twice (or more depending on the number of CVs).

Whilst a single metadynamics simulation can be as quick and easy to actually run as an unbiased simulation and can potentially provide detailed information on the free energy surface from that one simulation, that does not take account of the fact that selection of the CVs can be difficult and time consuming requiring significant trial and error.

3.4 Umbrella Sampling theory

The method of umbrella sampling[103] involves modifying the potential during the simulation such that conformations along a reaction pathway become more favourable. An additional potential (usually harmonic) is used to restrain the system in a small window close to the chosen value along the path of the reaction co-ordinate. The free energy surface along the chosen coordinate is referred to as the potential of mean force (PMF). A series of simulations are run with different “chosen values”. The output of these separate simulations are then combined to produce a PMF profile along the reaction pathway. The windows must overlap to ensure a smooth sampling distribution. In order to obtain the unbiased PMF and thus the free energy difference, the Weighted Histogram Analysis Method (WHAM) is used to combine the output of the simulations[104,105].

For example, if we were seeking to use umbrella sampling to establish the PMF profile for the red circle in figure 3.10 moving from left to right along the x axis, we could run a series of 11 simulations with the dot restrained in each of the blue positions with a harmonic potential. The restraining constant for the harmonic potential is selected such that the dot is not fixed rigidly in place during each simulation; it is free to move in a window (represented by the area enclosed by the black vertical lines in the figure). Each curve at the bottom of the figure represents the distribution of the x positions of the dot during each simulation. In order to effectively combine the output of the simulations these histograms must overlap, as they do in the figure.

WHAM[104,105] is then employed to combine the results of the simulations. The principle underpinning this approach is that with a discrete number of states, a histogram may be created with discrete bins that give the relative probability of observing the different states along the reaction profile; using that relative probability, one can calculate the PMF profile along that reaction profile.

The probability of finding the system in a particular state (x) is related to that state's free energy:

$$P(x) \propto e^{-A(x)k_BT} \tag{3.10}$$

therefore the free energy of the unbiased system, $A(x)$, is given by:

$$A(x) = -k_BTlnP'(x) - U'(x) + F \tag{3.11}$$

where $P'(x)$ and $U'(x)$ represent the probability distribution and applied umbrella potential of the biased simulation respectively. The constant F is undetermined and depends on $U'(x)$; different simulation windows have different F values.

The WHAM equations[104–106] are:

$$P(x) = \frac{\sum_{i=1}^{N_{sims}} n_i(x)}{\sum_{i=1}^{N_{sims}} N_i exp([F_i - U_i'(x)]/k_BT)} \tag{3.12}$$

and

$$F_i = -k_BTln\left(\sum_{X_{bins}} P(x)exp\left(\frac{(-U_i'(x))}{k_bT}\right)\right) \tag{3.13}$$

where N_{sims} is the number of simulations, $n_i(x)$ is the number of counts in a histogram bin associated with the value x, U_i' and F_i are the biasing potential and free energy shift from simulation i respectively and $P(x)$ is the unbiased probability distribution. Both $P(x)$ and F_i are unknown.

The WHAM equations are solved by iteration to self-consistency using the WHAM software implementation[106].

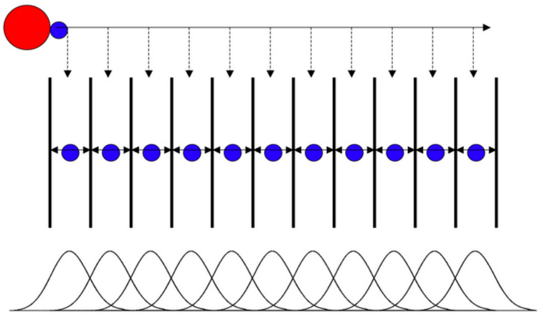


Figure 3.10: An example to illustrate calculating the PMF for the red dot moving from left to right using umbrella sampling[107].

3.5 Method: Enhanced sampling

Six initial bilayer configurations were prepared for all the ELBA1.0 simulations involving enhanced sampling: 1 gel and 1 L_α configuration for each of 3 hydration levels: 3 wpl, 10.56 wpl and 33 wpl. In order to try to ensure that each was fully relaxed in its configuration, these were prepared by 'annealing' bilayers constructed in the gel phase using the protocol described below. All simulations were performed using LAMMPS (Nov 2016) with the DLM integrator. A Langevin thermostat, water and lipids thermostatted separately, was used for the NPT and NVT portions together with a Berendsen barostat with anisotropic pressure coupling, xy coupled, pressure= 1 atmosphere, for the NPT portions.

In order to obtain the ELBA1.0 annealed gel phase bilayer, each gel conformation used in the standard MD simulations for the required hydration level was simulated for the sequence of ensembles/temperatures in table 3.2 and the output conformation used as the input 'gel' configuration for all the enhanced sampling techniques described as starting in the gel phase. This protocol was designed to ensure that the bilayer remained in the gel phase but had nonetheless been cycled through higher temperatures to ensure that it had had sufficient opportunity to relax fully. Therefore all periods of heating were in the NVT ensemble, followed by periods at lower temperature in the NPT ensemble.

As discussed above, the ELBA1.1 forcefield was not stable in the gel phase conformation at any hydration level/temperature tested and therefore it was not possible to construct annealed gel phase input configurations for the ELBA1.1 forcefield.

To obtain the annealed ELBA1.0 and ELBA1.1 L_α phase bilayers at each of the specified hydration levels, the conformations used in the standard MD simulations

Timestep / fs	Simulation time / ns	T / K	Ensemble
15	10	353	NVT
15	2	323	NVT
15	2	323	NPT
15	10	383	NVT
15	2	323	NVT
15	2	323	NPT
15	5	403	NVT
15	5	383	NVT
15	2	323	NVT
15	2	323	NPT
15	5	403	NVT
15	5	383	NVT
15	2	323	NVT
15	2	323	NPT
15	5	403	NVT
15	5	383	NVT
15	2	323	NVT
15	2	323	NPT
15	5	403	NVT
15	5	383	NVT
15	2	323	NVT
15	2	323	NPT
15	5	403	NVT
15	5	383	NVT
15	2	323	NVT
15	2	323	NPT

Table 3.2: Annealing protocol to produce gel conformation input

Timestep / fs	Simulation time / ns	T / K	Ensemble
15	10	353	NPT
15	4	323	NPT
15	10	383	NPT
15	2	323	NPT
15	2	323	NPT
15	5	403	NPT
15	5	383	NPT
15	2	323	NPT
15	2	323	NPT
15	5	403	NPT
15	5	383	NPT
15	2	323	NPT
15	2	323	NPT
15	5	403	NPT
15	5	383	NPT
15	2	323	NPT
15	2	323	NPT
15	5	403	NPT
15	5	383	NPT
15	2	323	NPT
15	2	323	NPT

Table 3.3: Annealing protocol to produce L_α conformation inputs

described above was simulated for the sequence of ensembles/temperatures in table 3.3 with either ELBA1.0 or ELBA1.1 forcefield and the output conformations used at the input 'L_α' configurations for the all the enhanced sampling techniques described as starting in the L_α phase.

3.5.1 Metadynamics (standard and well-tempered): Results

The MD engine used for the metadynamics simulations was LAMMPS (Nov 2016) with the PLUMED 2.4 plugin patch[102] to control the applied bias and for post processing. The annealed (see above for annealing protocol) bilayers were simulated in the NPT ensemble at either 323 K or 341 K with a timestep of 15 fs using a Langevin thermostat (lipids and water thermostatted separately) and Berendsen barostat applied anisotropically with xy coupling and a pressure of 1 atmosphere.

The CV used was the area of the xy plane of the simulation box, which is equivalent to the bilayer area when the bilayer remains parallel to that plane, with bias applied every 200 timesteps during the simulation, $W = 0.36$ kcal/mol and $\sigma = 13$ Å^2 (equivalent to 0.2 Å^2 per lipid) being half the variation of the CV over a non-biased simulation. In initial simulations it was found that the simulations were driven to very small xy areas where the bilayer structure was completely destroyed. An interval was therefore introduced into the Plumed input files which ensured that the Gaussian biases were not applied outside the range of xy areas 2560 - 4800 Å^2 ie an equivalent range of 40-70 Å^2 per lipid.

Additionally, the biasfactor (γ) specified for the well-tempered simulations was $\gamma = 8.0$K on the basis that the energy barrier is reduced by a factor equal to the bias factor[102].

The results of the metadynamics simulations were analysed to check for convergence. The biased trajectory was divided into 'strides' and the free energy calculated for each one. As the biased simulation converges, the free energy curve calculated for each stride becomes more similar. For example, figure 3.11 shows the free energy curves for a metadynamics simulation of a DSPC bilayer, 10.56 wpl, simulated at 323 K, with the initial configuration being the annealed gel phase ELBA1.0 bilayer. As the figure shows, strides 3 and 4 show very similar free energy curves over the span of interest.

The final free energy curve is shown in figure 3.12. This shows no indication of a stable or metastable state at an area around $49 - 50$ Å^2 per lipid. This is in contradiction to both the BRAHMS and LAMMPS simulations described above which showed the gel phase conformation of the bilayer persisting for lengthy unbiased simulations.

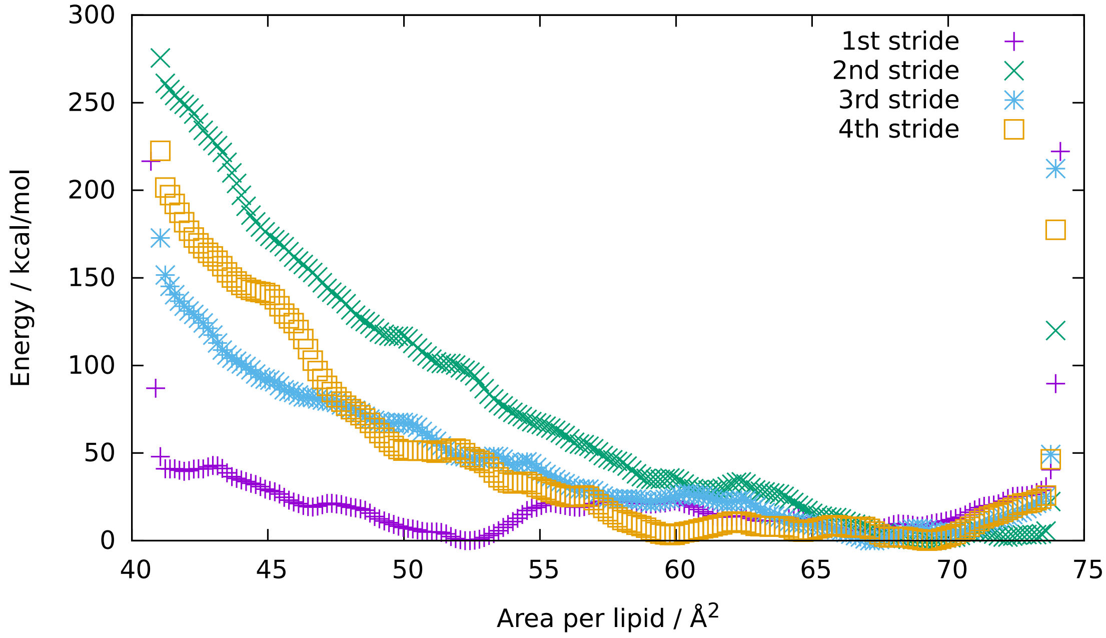


Figure 3.11: Free energy curves calculated every 100,000 Gaussian depositions for convergence.

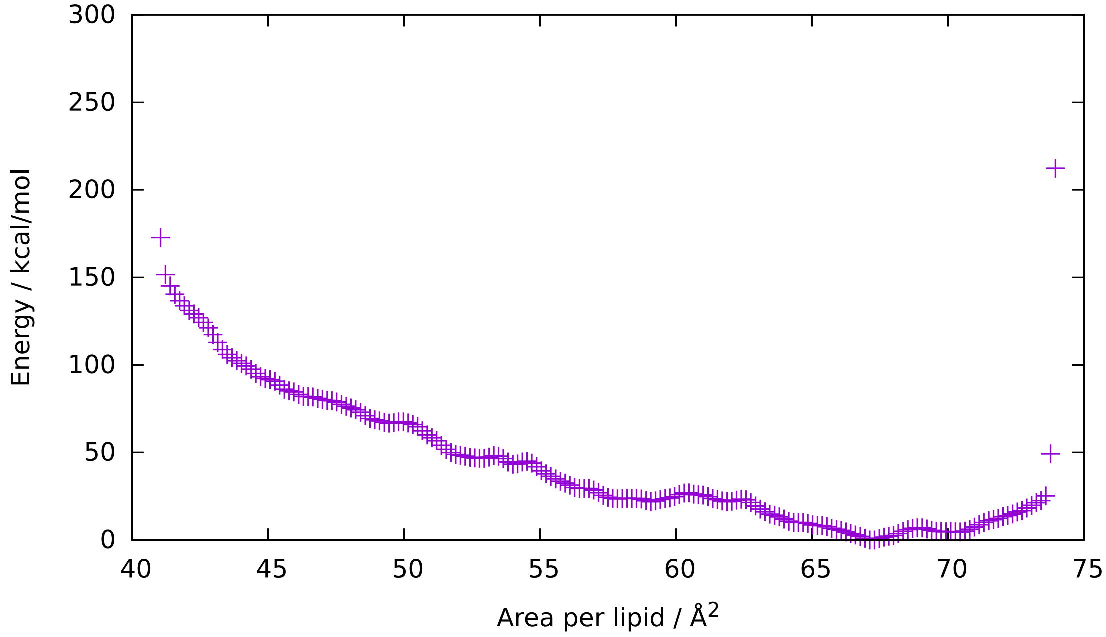


Figure 3.12: Final free energy surface calculated from the metadynamics simulation.

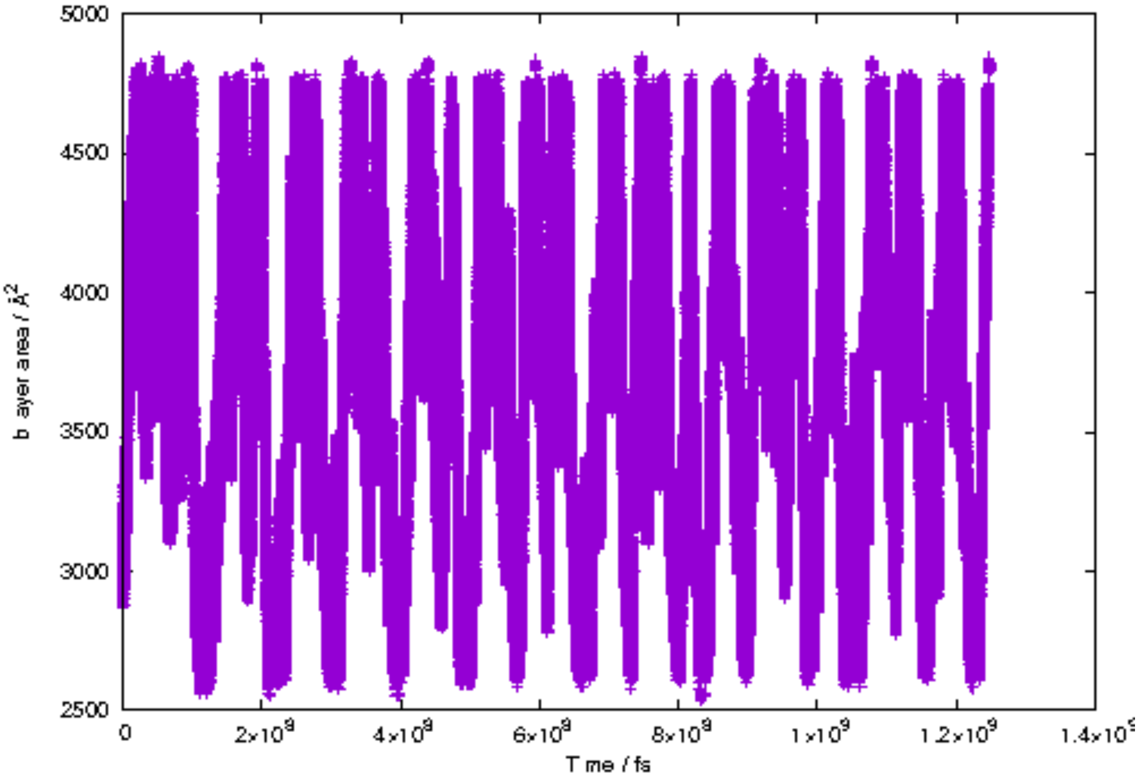


Figure 3.13: Evolution of CV during metadynamics simulation

An indication as to why the free energy curve calculated from the metadynamics simulation using bilayer area as a CV may not be an accurate representation is given by figure 3.13. This figure shows how the bilayer area changes during the simulation. Over the course of the simulation, the bilayer area oscillates several times between that associated with a gel phase and with an L_α phase as one would hope from a sufficiently well sampled simulation. However, the figure also shows some ‘trapping’ in both phases; there is no evidence of smoothly diffusive behaviour in the spread of bilayer areas across the simulation time as would be required to show that the bilayer area were a sufficient an appropriate CV alone to describe the relevant energy landscape.

Further indication that bilayer area or area per lipid is an ineffective sole choice of CV for this transition is shown in figure 3.14. This figure shows that, in the well tempered metadynamics simulation, the heights of the deposited Gaussians fall initially, then are restored to close to the initial value as the system escapes the energy basin of the unbiased/hidden CV.

There is therefore clearly an unbiased slow degree of freedom affecting the phase transition and it is not possible to say that the metadynamics simulations are

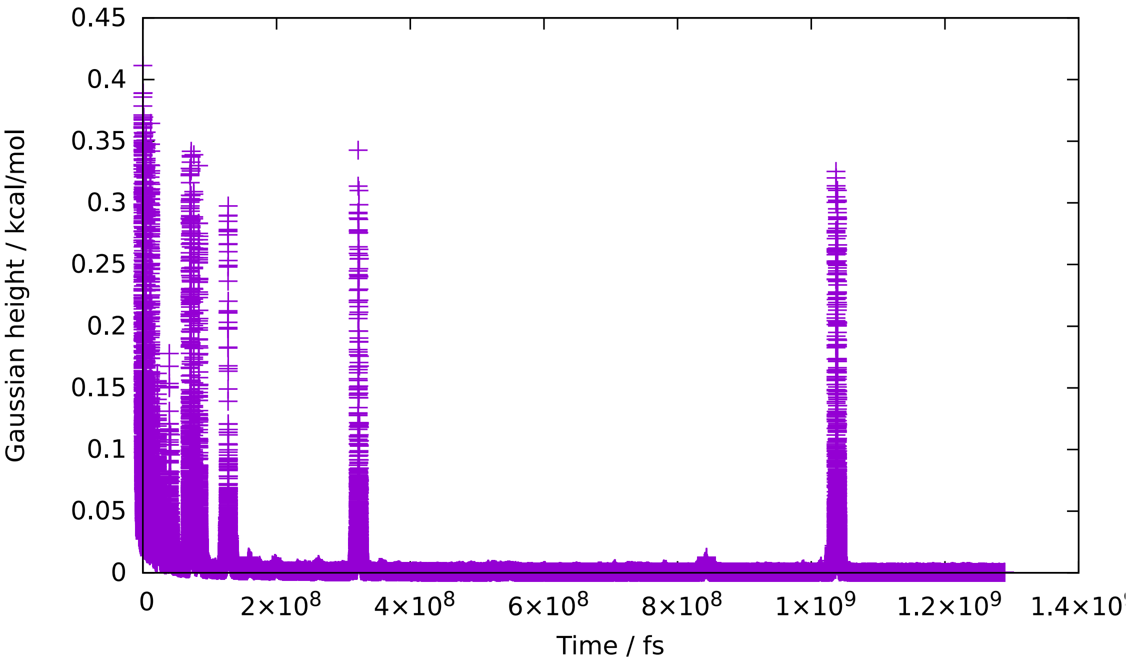


Figure 3.14: Heights of Gaussians deposited during a well-tempered metadynamics simulation. 10.56 wpl, 323 K simulation starting in the annealed gel conformation.

converged unless either this degree of freedom is equilibrated ie the simulations are run sufficiently long such that multiple transitions are observed or this slow degree of freedom is biased as an additional CV. Since the slow degree of freedom is unidentified, biasing it is impossible. Running a sufficiently long simulation to observe several transitions between the basins in each direction is likely to run into the same problem as observing L_α to gel transitions in the unbiased simulations - namely insufficient simulation time. Therefore a different enhanced sampling technique was attempted: Umbrella Sampling.

3.5.2 Umbrella Sampling: Results

The MD engine used for the umbrella sampling simulations was LAMMPS (Nov 2016) with the PLUMED 2.4 plugin patch to control the umbrella sampling bias. The annealed (see above for annealing protocol) bilayers were simulated in the NPT ensemble at either 323 K or 341 K with a timestep of 15 fs using a Langevin thermostat and Berendsen barostat applied anisotropically with xy coupling and a pressure of 1 atmosphere. For each of 40 separate simulations ("umbrella windows") the total xy area of the simulation box was restrained using a harmonic restraint, where $U(s)$ is the additional potential applied:

$$U(s) = \frac{\kappa}{2}(x - x_0)^2 \tag{3.14}$$

where x_0 is the xy area to which the xy plane was restrained. $\kappa = 0.001\text{kcal.mol}^{-1}.\text{Å}^2$. This value of κ was selected as being approximately equal to $\frac{k_B T}{\sigma^2}$ where σ^2 is the variance in the xy bilayer area in an unbiased simulation.

Umbrella windows were run for each bilayer/temperature combination restrained to total bilayer areas equivalent to areas per lipid between 40 and 80 Å^2. The initial set of umbrella windows were equally spaced throughout the range, at bilayer areas equivalent to 1 Å^2 per lipid apart. Each umbrella window was run for between 17 and 28 ns, with only the last 5 ns used in the subsequent analysis in order to allow any slow degrees of freedom to have equilibrated. The resulting bilayer areas were plotted to ensure the windows overlapped. Where there were gaps in the overlap of windows, additional simulations were run (for example, see figure 3.15). For all simulation sets, at least one such additional window was required and all additional windows were required between $\approx 3000\text{Å}^2$ and $\approx 3500\text{Å}^2$, which reflects an area per lipid of between $\approx 47\text{Å}^2$ and $\approx 55\text{Å}^2$ ie around the area per lipid associated with the gel-phase.

The output of the umbrella window simulations were analysed using WHAM (Weighted Histogram Analysis Method) using the Grossfield code implementation[106]. Figures 3.16 and 3.17 show the results of the WHAM analysis of the umbrella sampling simulations conducted at 323 K and initiated with annealed gel phase ELBA1.0 bilayers and annealed ELBA1.0 L_α phase bilayers respectively.

The experimental and unbiased simulation results summarised in figures 3.7 and 3.8 suggest that at 323 K a more stable gel phase, compared to the L_α, should be expected to be indicated for all the hydrations tested (3 wpl, 10.56 wpl and 33 wpl). As shown in figures 3.16 and 3.17, at the lowest hydration level simulated (3 wpl) there is only one minimum in the PMF profiles for both those bilayers initiated in the gel and L_α phases and for both this occurs at an area per lipid approximately equal to that associated with a gel phase. This accords with the results from the unbiased simulations (figures 3.7 and 3.8) where the gel phase remained stable under a similar temperature/hydration combination. The situation is slightly more complex for the 2 higher hydration levels analysed. From the unbiased simulations, it might have been expected

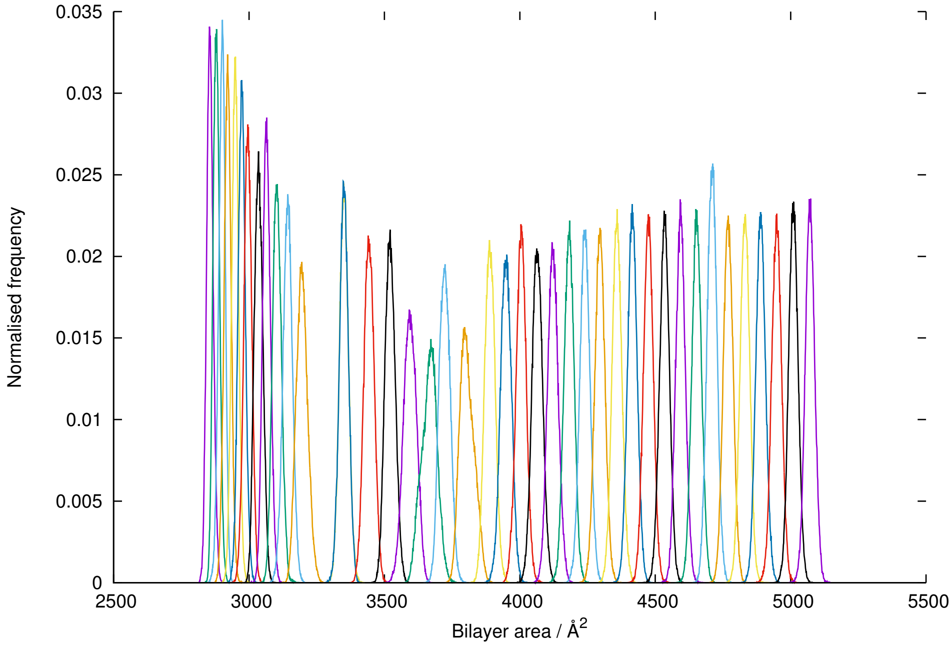


Figure 3.15: Histogram of umbrella windows 10.56 wpl, 323 K, initiated in gel phase showing a non-overlap in umbrella windows at a bilayer area of $\sim$ 3250 Å^2 ($\sim$ 51 Å^2 per lipid).

that these would also show only one minimum in the PMF, as at both hydration levels, the gel phase is stable in the unbiased simulations. However, for the umbrella sampling simulation initiated in the gel phase (figure 3.16) 2 low energy conformations in the profile are observed at areas per lipid associated with the gel and L_α phases. This indicates that the 'stability' of the gel phase in the unbiased simulations may be due to insufficient sampling ; whilst the PMF profiles indicate that L_α phase is more thermodynamically stable the unbiased simulation remain in the gel phase due to the kinetic barrier to transition.

The PMF curve obtained from the umbrella sampling simulations initiated in the L_α phase are less straightforward (figure 3.17). They do not show such an obvious minimum at the gel phase area per lipid. This may be further indication that at these hydration levels the gel phase is not a stable phase but merely a metastable artefact. For the simulations initiated in the L_α phase it was impossible to drive them back to a gel phase because that phase was not in fact thermodynamically stable. Alternatively, the lack of a minimum in the PMF at that point may also indicate that the umbrella sampling simulations were too short to equilibrate the slow degrees of freedom required to condense the bilayer into the genuine gel

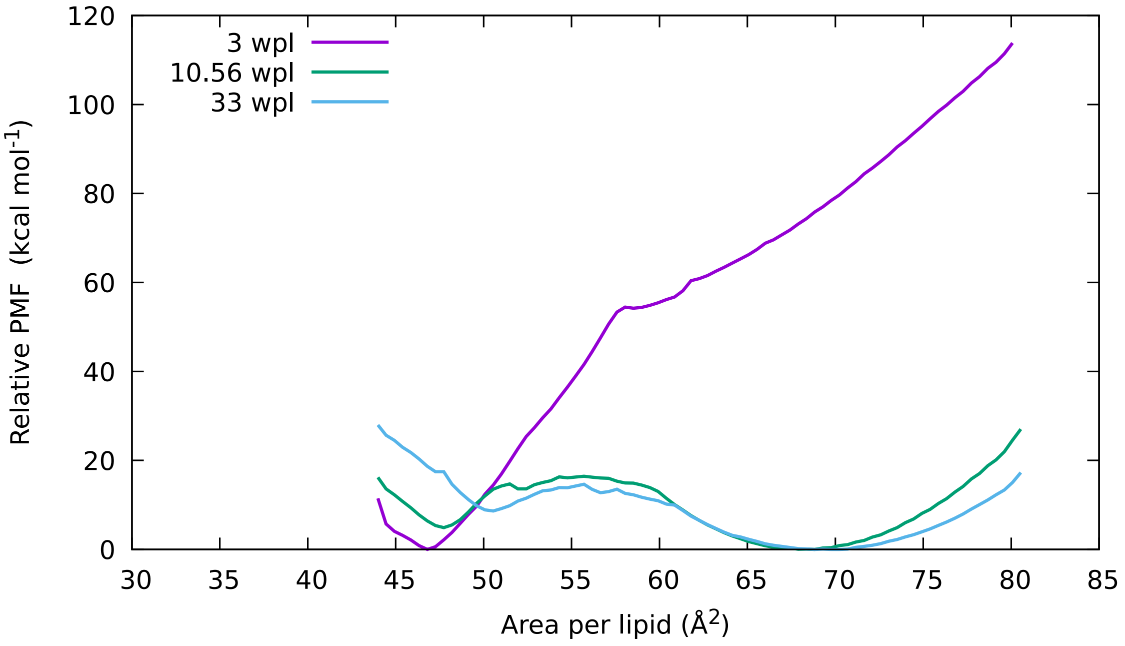


Figure 3.16: The PMF curves obtained from 3 sets of umbrella sampling simulations using the ELBA1.0 forcefield conducted at 323K, all initiated from annealed gel phase bilayers at 3 different hydration levels.

phase.

Figures 3.18 and 3.19 show the equivalent PMF curves obtained for the ELBA1.0 forcefield at the higher temperature of 341 K initiated from the annealed gel and annealed L_α conformations respectively. The unbiased simulations suggest that, at this temperature, the gel phase is the most stable for the lowest hydration and the L_α for the higher 2. This is confirmed by the results of the umbrella sampling analysis. Even for the simulations initiated in the annealed gel conformation, for the 2 higher hydrations, the stable phase is clearly the L_α at 341 K with no indication of any stable/metastable phase at lower area per lipid.

None of the sets of umbrella sampling simulations conducted using the ELBA1.1 forcefield (figures 3.20 and 3.21) show any sign of any stable phase with an area per lipid similar to that of the gel phase. This accords with the results seen in the unbiased simulations whereby even the equilibration process was sufficient to cause a gel to L_α transition. For the fully hydrated ELBA1.1 bilayers (33 wpl) the umbrella sampling simulations do seem to indicate a stable conformation at low area per lipid - albeit almost 10 Å^2 lower than one would expect from a

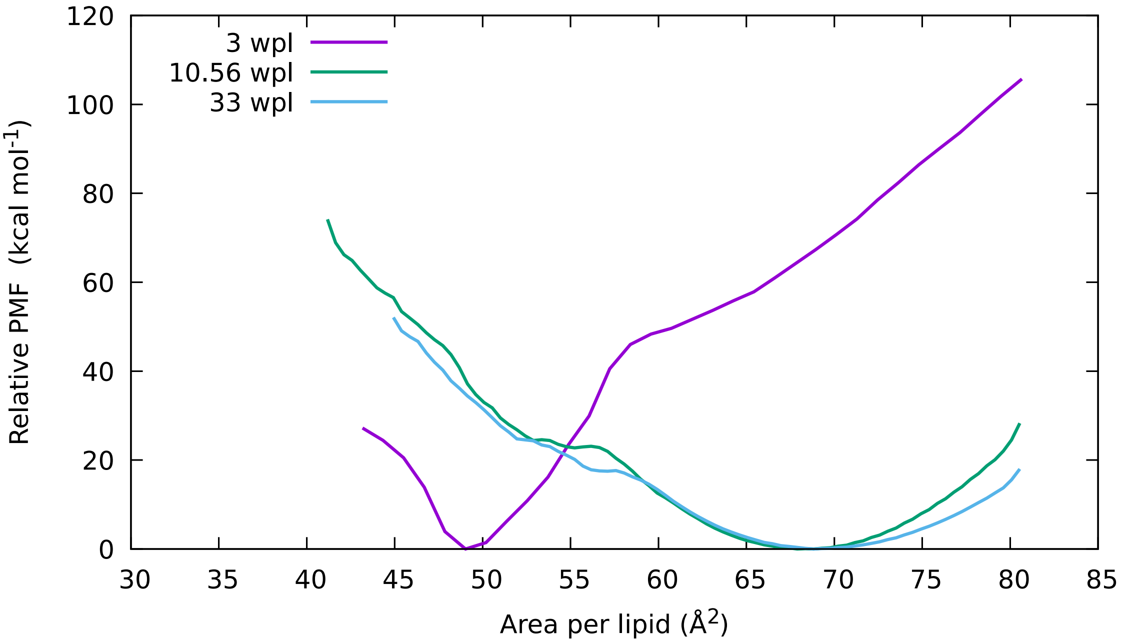


Figure 3.17: The PMF curves obtained from 3 sets of umbrella sampling simulations using the ELBA1.0 forcefield conducted at 323K, all initiated from annealed L_α phase bilayers at 3 different hydration levels.

gel phase. Visual examination of the trajectories from the umbrella sampling windows, see figures 3.22, indicates that this is not a gel phase but rather that the bilayer has broken down and therefore the xy area of the simulation box is not longer equivalent to the bilayer area nor proportional to area per lipid.

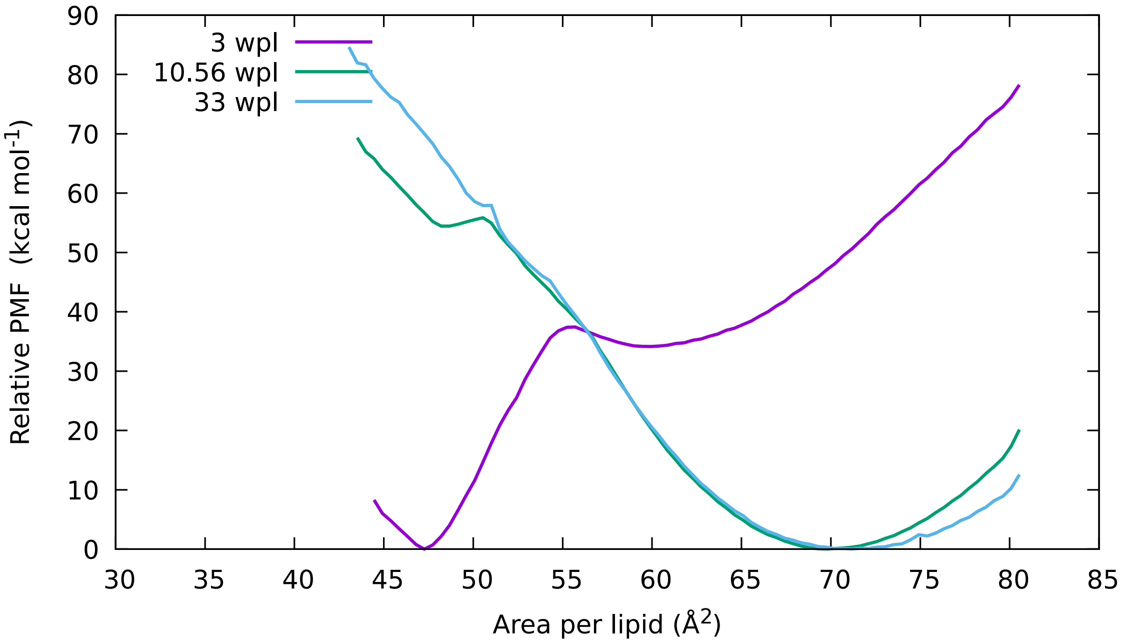


Figure 3.18: The PMF curves obtained from 3 sets of umbrella sampling simulations using the ELBA1.0 forcefield conducted at 341K, all initiated from annealed gel phase bilayers at 3 different hydration levels.

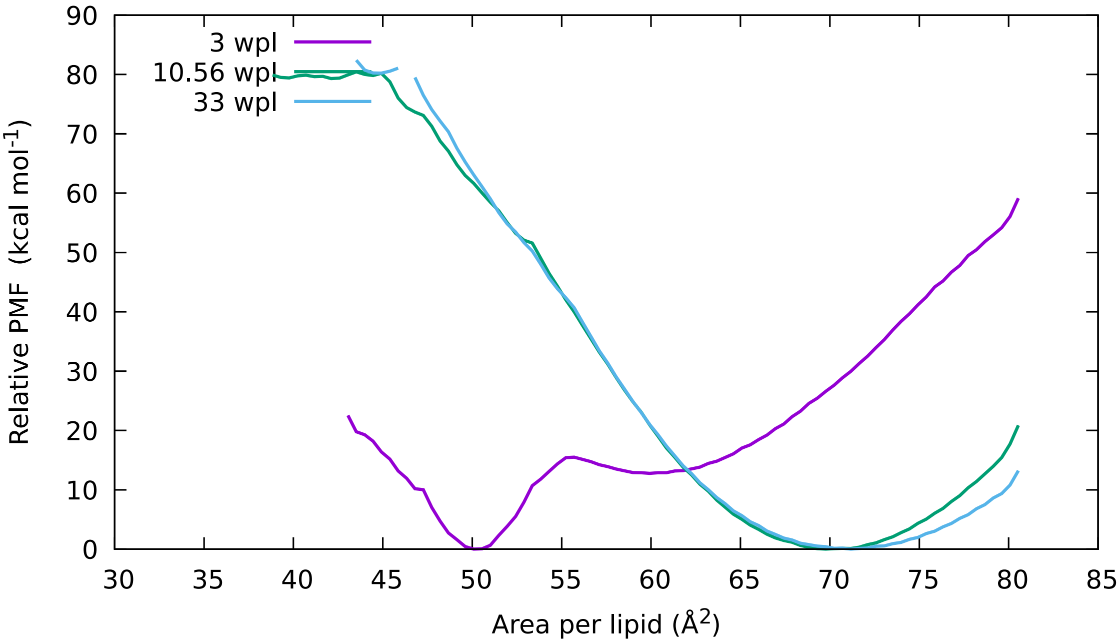


Figure 3.19: The PMF curves obtained from 3 sets of umbrella sampling simulations using the ELBA1.0 forcefield conducted at 341 K, all initiated from annealed L_α phase bilayers at 3 different hydration levels.

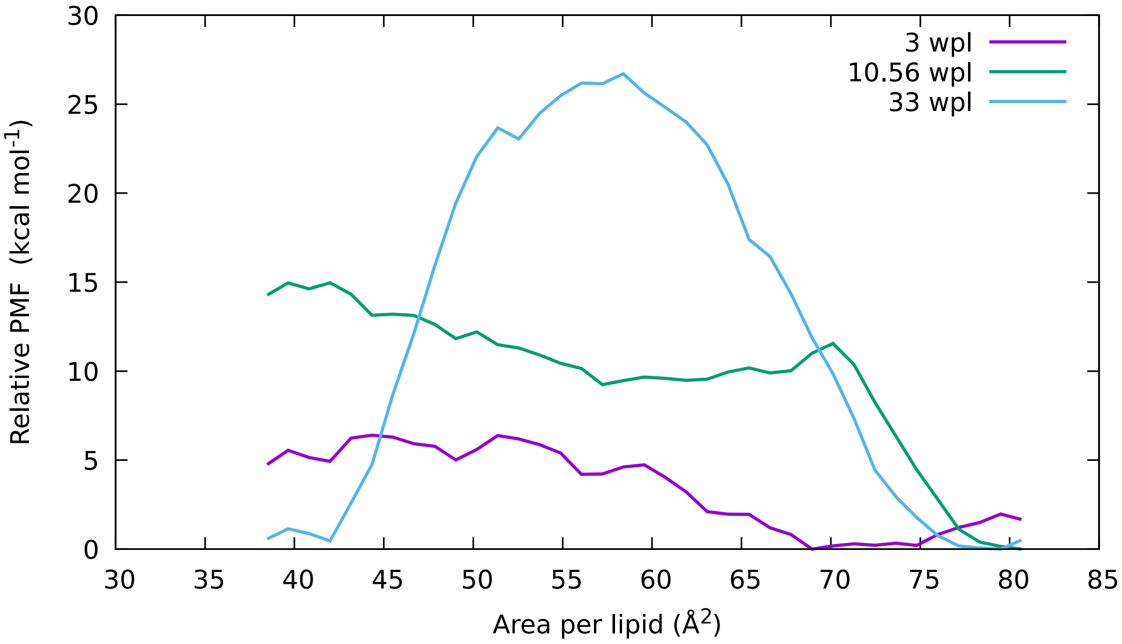


Figure 3.20: The PMF curves obtained from 3 sets of umbrella sampling simulations using the ELBA1.1 forcefield conducted at 323 K, all initiated from annealed L_α phase bilayers at 3 different hydration levels.

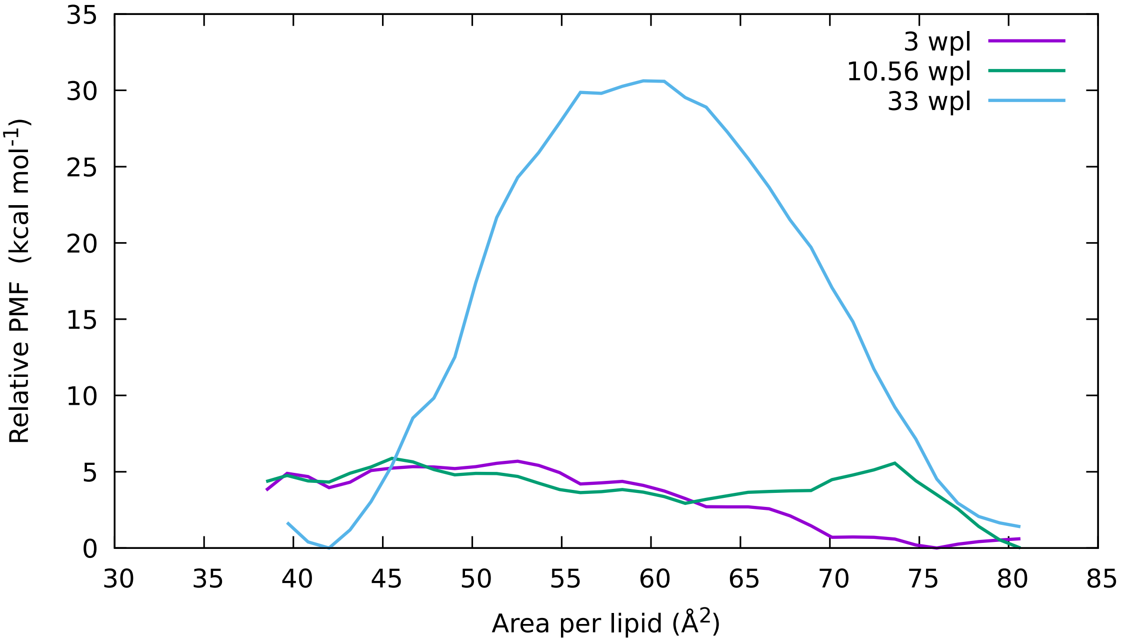
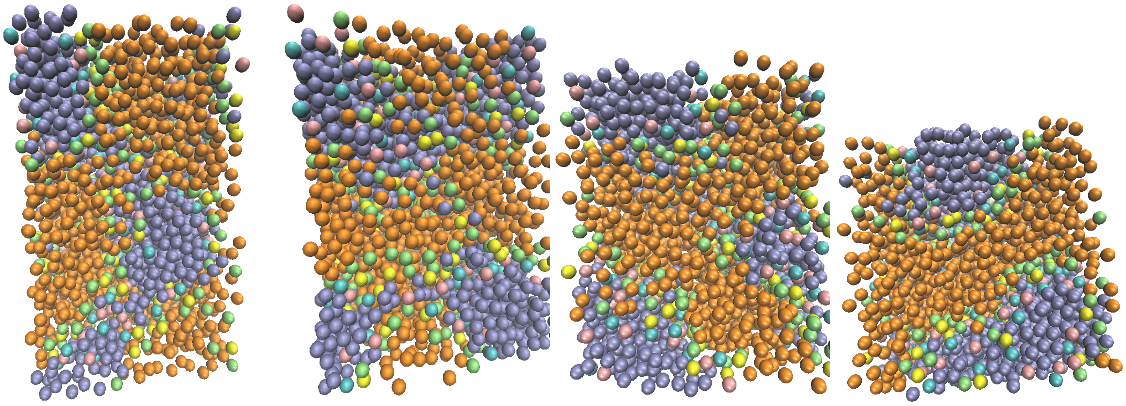


Figure 3.21: The PMF curves obtained from 3 sets of umbrella sampling simulations using the ELBA1.0 forcefield conducted at 341 K, all initiated from annealed L_α phase bilayers at 3 different hydration levels.



(a) xy plane of the simulation box = 2560 Å^2

(b) xy plane of the simulation box = 2880 Å^2

(c) xy plane of the simulation box = 3200 Å^2

(d) xy plane of the simulation box = 3520 Å^2

Figure 3.22: Four snapshots from end of the umbrella sampling windows for a bilayer simulated using the ELBA1.1 forcefield. Water beads are blue, lipids tails orange and lipid heads green/yellow.

3.6 Conclusions

Unbiased simulation of DSPC bilayers using the ELBA1.0 forcefield indicate that this forcefield reproduces the experimental phase diagram for this lipid well. On the other hand, the ELBA1.1 forcefield does not, despite the 2 forcefields being extremely similar. Using umbrella sampling techniques gives an indication that the phases and transitions observed using the ELBA1.0 forcefield are genuine thermodynamically stable phases rather than an artefact of insufficient sampling during the simulations. Similar umbrella sampling techniques also support the conclusion that the gel phase is not reproduced in the ELBA1.1 forcefield.

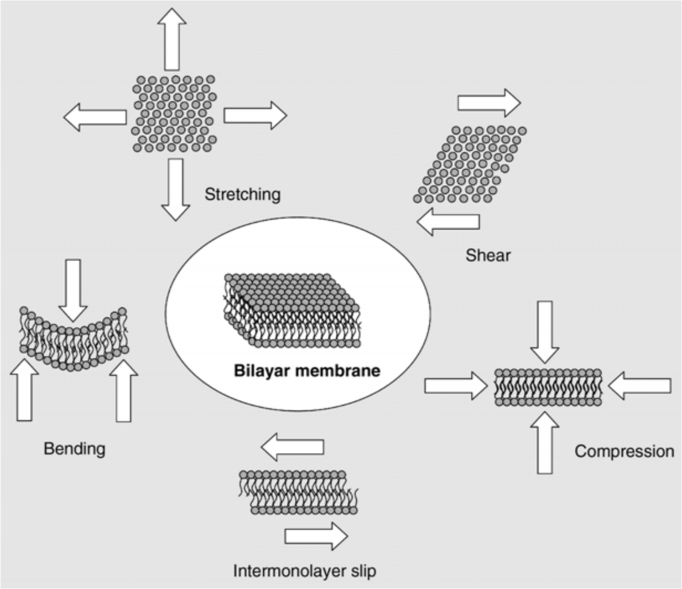
Area per lipid is not an appropriate CV for metadynamics simulations exploring lipid phase transitions as the area per lipid does not sample the slowest degrees of freedom involved in that phase transition. If lipid phase transitions are to be explored using metadynamics other CVs should be considered and tested, including head group packing as the gel phase is associated with hexagonal packing.

Chapter 4

Bending

As described in chapter 1 above, the lipid bilayer forms the underlying structure of cell membranes. In nature membranes are almost always asymmetric; there is a difference in composition between the inner and outer leaflets[108]. The degree of asymetry can vary, from slight differences in the distribution of lipids in some eukaryotic membranes to the outer membrane of Gram-negative bacteria which is almost completely assymetric between 2 very different lipid types[109–111]. It is thought that such asymmetry plays an important role in biological functions such as signal transduction[112,113], vesicle budding[114] and modulation of protein channel opening and closing[115]. These functions are believed to be influenced by altering the membranes' elastic properties[112,115], including its bending modulus[116]. The bending modulus is a macroscopic constant that describes the tendency of a particular material to resist bending under low stress. For a membrane, it is the energy required to deform the membrane from its intrinsic curvature - ie deform it from the equilibrium shape which is dependent on the shape of its constituent lipids, described in chapter 1.

In this chapter, we describe the experimental techniques employed to measure bilayer bending, the different analysis methods used to calculate bending modulus from simulation and describe the MD simulations we have undertaken to investigate the bending properties of both symmetric and asymmetric bilayers using different forcefields and analysis techniques.



.

Figure 4.1: Schematic of the different modes of membrane deformation[117]

4.1 The theory of membrane stretching and bending

Stretching any material or object requires energy. We can define that energy ($E_{stretch}$) as:

$$E_{stretch} = \frac{1}{2} K_{stretch} \frac{(A - A_0)^2}{A_0} \tag{4.1}$$

where A_0 is the area of the membrane under no external stress which is then stretched to an area A where $A > A_0$. $K_{stretch}$ is a proportionality constant.

The lateral stress or tension (Σ) is the derivative of that energy with respect to area:

$$\Sigma = \frac{\partial E_{stretch}}{\partial A} = K_{stretch} \frac{A - A_0}{A_0} \tag{4.2}$$

as the strain u is:

$$u = \frac{A - A_0}{A_0} \tag{4.3}$$

therefore:

$$\Sigma = k_{stretch} u \tag{4.4}$$

which is Hooke's law - stress is proportional to strain.

$k_{stretch}$ for lipid membranes can be measured experimentally using micro-pipette techniques[118,119] described further below. It has been found that lipid membranes obey Hookes law (they exhibit a linear stress/strain relationship) with a $k_{strain} = 250mN.m^{-1}$[120] up to a rupture limit of $u \approx 0.04$[119,120].

If considered on a length scale considerably larger than that of the individual lipids, it has been shown that, because its thickness is much smaller than its lateral dimension, the membrane can be considered as a two-dimensional elastic sheet when considering a its elastic properties. Thin 3-dimensional objects can be deformed not only by stretching but also by bending. They can be deformed in directions away from the direction(s) in which they are extended leading to substantial changes in shape for a much smaller applied stress than that required for stretching. For example a sheet of paper can be much more easily bent than stretched. If we consider bending a square thin sheet of side L, as it curves, the outside surface of the stretches a little and the inside compresses. Somewhere between the 2 is a region which is neither stretched nor compressed - this is the neutral surface. The energy required $E_{stretch}$, if we consider a 3D, uniaxial extension analagous to equation 4.2 is:

$$E_{stretch} = \frac{1}{2}Y\frac{(V - V_0)^2}{V_0} \tag{4.5}$$

where Y is the Young modulus and V and V_0 are the new volume element and equilibrium volume respectively.

The bending energy per area, assuming the sheet's thickness is h and the radius of bending is R, is:

$$\begin{aligned} e_{bend} &= \frac{E_{bend}}{L^2} \\ &= \frac{1}{L^2}\int_0^L\int_0^L\int_{-h/2}^{h/2}\frac{1}{2}Y\frac{((1+\frac{z}{r})dxdydz - dxdydz)^2}{dxdydz} \\ &= \frac{1}{2}Y\int_{-h/2}^{h/2}dz\left(\frac{z}{R}\right)^2 \\ &= \frac{1}{24}Y\frac{h^3}{R^2} \end{aligned} \tag{4.6}$$

Thus the bending energy is proportional to the cubic power of thickness - if the sheet is made thinner (or thicker) the energy required to bend it changes greatly. The 3D modulus Y can be expressed using the 2D stretching modulus $K_{stretch}$:

$$K_{stretch} = Yh \tag{4.7}$$

therefore:

$$e_{bend} = \frac{1}{24} K_{stretch} \left(\frac{h}{R} \right)^2 \tag{4.8}$$

Equation 4.8 forms part of the basis for calculating bending modulus according to the polymer brush method (discussed further below). However it incorporates 3 important assumptions; first that bending modulus (κ) is related to $K_{stretch}$ thus:

$$\kappa = \frac{1}{12} K_{stretch} h^2 \tag{4.9}$$

Secondly the effect of the Poisson ratio - ie the amount that a material under strain changes dimension in the direction perpendicular to the strain - is neglected, and finally we have only considered a single radius of curvature. A membrane is a 2D surface and therefore there are different types of bending that should be considered.

4.1.1 The Helfrich model of bending modulus

The bending modulus, K_c, described by Helfrich in 1973[121], characterises the bending of a membrane under low stress and relates to the energy required to deform the shape. It is a quantitative measure of the flexibility of a bilayer. The total bending energy of a membrane according to Helfrich theory is related to the bending modulus K_c as follows:

$$E_{bend} = \int \{ \frac{1}{2} K_c (H - C_0)^2 + K_G K \} dA \tag{4.10}$$

where H is the total curvature, C_0 is the spontaneous curvature and is 0 for a flat membrane), K_G is the Gaussian bending modulus or saddle splay modulus, K is the Gaussian curvature and A the membrane area. Gauss-Bonnet theorem states that the surface integral over the Gaussian curvature K_G, can be written as a topological constant, which means that the K_G term in equation 4.10 leads to a constant. This means that, in the case of a flat bilayer, ie $c_0 = 0$ and where the surface is represented by a height function $h = h(x, y)$ and taking into account the surface tension Σ:

$$E_{bend} = \frac{1}{2} \int_{plane} dxdy \left(\kappa (\Delta h)^2 + \Sigma (\nabla h)^2 \right) \tag{4.11}$$

4.2 Experimental Methods

There are 3 main experimental methods used to obtain the bending modulus for lipid bilayers, the development of which have been described in detail in recent reviews[122,123]. All 3 rely on measurement of displacement of the membrane from its average shape due to thermal fluctuations. Membranes readily undulate at room temperature because bending rigidity is $\approx 10 - 20 k_B T$.

In turn, the measurement methods rely on 2 different kinds of bilayer arrangement: giant unilamellar vesicles (GUVs) and bilayer stacks. GUVs are liposomes, up to $100\,\mu m$ in diameter, therefore directly visible using optical microscopy, consisting of a single bilayer. GUVs can be formed by various methods including by lipid-film hydration[124], electroformation using an alternating electric field[125] or phase transfer[126].

4.2.1 Bilayer preparation

The lipid film hydration technique involves spreading a thin film of lipids dissolved in organic solvent onto a plate and evaporating the solvent under vacuum to form a multilamellar stack of dry lipids on the plate. This stack is then hydrated and, as each layer of lipids hydrates, it swells and peels away from the stack forming a GUV. Electroformation similarly involves spreading a thin film of lipids and evaporating off the organic solvent, however, in the case of electroformation, the thin film is spread on an electrode. As the film is hydrated, often with a aqueous sucrose solution, an alternating electric field is applied to the electrodes for 1 or 2 hours. Initially the alternating current frequency is usually between 10 - 20 Hz and many protocols gradually reduce this frequency during the electroformation procedure. The amplitude that is used usually varies between 1.5 - 5 V, with a lower voltage used when charged lipids form part of the lipid mixture. Using an electric field during the hydration process, rather than simply allowing the vesicles to increases the yield of high quality GUVs[127].

Emulsion phase transfer methods of GUV production do not involve hydrating dry lipids at all. As shown in figure 4.2, they are based on passing lipid stabilised water-in-oil droplets through an oil - water interface, where it picks up a second lipid monolayer, resulting in formation of a completed GUV. This method can also be adapted to produce asymmetric membranes[128,129]. Further details of the microfluidic phase transfer method employed by Elani et al[126] are provided below.

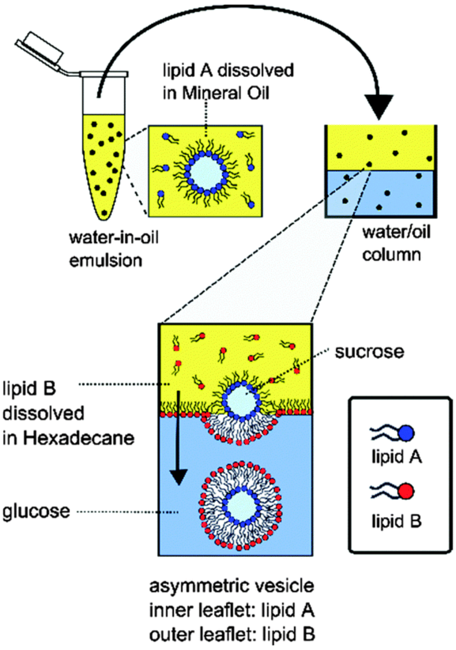


Figure 4.2: “Schematic of asymmetric GUV generation via phase transfer. A water-oil emulsion is first prepared with lipid A dissolved in the oil phase, encasing droplet in a monolayer. The emulsion is added to the water-oil column with lipid B dissolved in the oil phase, assembling as a monolayer at the interface. Droplets descend through the column under gravity (represented by the arrow) and are enveloped by a second monolayer. Asymmetric GUVs are thus formed, with lipid A in the inner leaflet and lipid B in the outer leaflet” Reproduced from DOI: 10.1039/C5CC00712G (Communication) Chem. Commun., 2015, 51, 6976-6979 - Published by the Royal Society of Chemistry.

GUVs are generally more often used than bilayer stacks, because the latter exhibit slightly different properties, postulated to be due the steric interactions between the layers[130] and discussed further below.

4.2.2 Fluctuation Analysis - Flicker

This method for measuring the bending modulus was originally developed in 1975[131] for erythrocytes and later extended to giant vesicles[132]. The analysis of shape fluctuations is based upon the collection of a time series of snapshots of the GUV obtained by optical microscopy. Fluctuations in the radius are measured and a Fast Fourier Transform (FFT) is performed. The analysis assumes that the vesicle is, on average, a spherical object that conserves both its volume and area. In order to increase contrast, and therefore visibility of the fluctuations, different sugar solutions are often used for the aqueous solutions within and outside the GUV, therefore leading to a hydrostatic pressure ($\bar{p}$).

The complex vesicle fluctuations can be described as a general function on a sphere. That function is decomposed into a series of orthogonal Laplace spherical harmonic functions, Y_n^m, using FFT in the same way that a general linear function can be decomposed into a series of sine and cosine functions. The mean squared magnitude of the thermal shape fluctuations for particular harmonics, $\langle |U_n^m(t)|^2 \rangle$, depend on 3 variables; the bending modulus (K_c), the surface tension (γ) and the hydrostatic pressure ($\bar{p}$).

$$\langle |U_n^m(t)|^2 \rangle = \frac{k_B T}{K_c} \frac{1}{(n-1)(n+2)(\gamma + n(n+1)) + 2(2\gamma - \bar{p})} \tag{4.12}$$

where k_B is the Boltzmann constant and T the temperature. Provided that the fluctuations are small in relation to the GUV's diameter, the last term of the denominator which includes the hydrostatic pressure can be omitted. The mean surface tension does not need to be directly measured because $\frac{k_B T}{K_c}$ is independent of harmonic mode and, if the fluctuations are calculated for several modes, simultaneous equations can be used to calculate K_c without ever calculating the surface tension.

The advantage of this method is its simple experimental setup, though it should be noted that, although larger GUVs make fluctuations easier to measure, the larger the GUV, the more sensitive the result to artefacts introduced by variations in membrane tension and osmotic effects[133]. As Faucon et al report, "...for a vesicle of radius $R = 10^{-5} m$ and $k_c = 10^{-19} J$ a tension of $\gamma = 2x10^{-5} mN/m$

leads to a two fold decrease of the amplitudes of second harmonics, $U_2^m(t)$. For a small vesicle of radius $R = 10^{-8}m$ the tension required to produce the same effect would be $\approx 20mN/m$; a value one order of magnitude higher than the reported membrane rupture tension. Due to the relation $p \approx 2\gamma$, the same two-fold decrease of the amplitude of the second harmonic could also be produced by osmotic effects of as low as $2x10^{-9}M$ concentration difference for a giant vesicle, while 2 M are necessary for a small (100 Å) one."

The fluctuation analysis method has an advantage that there is also no need to introduce other molecules eg fluorophores into the system. Because only measurements for low surface tension conditions are used, consideration of lipid tilt is not required[134]. The main disadvantage is that, since the membrane must exhibit visible fluctuations, the method cannot be applied to vesicles in the gel phase.

4.2.3 Micropipette Aspiration

This method was described by Evans and Needham[118]. It is a mechanical manipulation method which uses a micropipette to aspirate part of the membrane of the GUV into the small-calibre pipette. The vesicle is at first slightly deflated and then drawn into the pipette, the volume enclosed by the membrane remains constant.

The length of the vesicle projection, ΔL, inside the pipette is connected with a projected area expansion that has two contributions: a reduction of membrane undulations and an increase in the area per lipid.

It can be demonstrated that the total area expansion, defined as:

$$\Delta a = \frac{1}{2}\left(\left(\frac{R_p}{R_0}\right)^2 - \left(\frac{R_p}{R_0}\right)^3\right)\frac{\Delta L}{R_p} \tag{4.13}$$

where R_p is the pipette radius and R_0 the exterior vesicle radius, is a function of the membrane tension γ:

$$\Delta a = \frac{k_B T}{8\pi K_c} ln(1 + c\gamma A) + \frac{\gamma}{K_a} \tag{4.14}$$

where c is a coefficient which has different values depending on the approximations used, A is the total area of the membrane and K_a is the area compressibility modulus; the resistance of the membrane to compression. Thus, in the "low tension regime", the logarithmic term dominates and almost all the area expansion is due to the smoothing of thermal undulations. The slope of the logarithm of the tension applied versus the area expansion is inversely proportional to the bending modulus K_c.

The length of the bilayer aspirated into the pipette must be measured with high precision and, in order to achieve this, fluorescent dyes can be used. However, such dyes (like most other impurities) can alter membrane elastic properties. In addition, the theory is based upon the assumption that the membrane slides freely along the pipette walls as it is drawn in. A coating can be added to reduce adhesions of the membrane to the pipette, but effect on K_c of their addition remain another possible source of error.

Finally, in order to maintain the the constant volume which is required for the theory, significant amounts of sucrose and glucose are typically added to the vesicle solution. It has been suggested that these additional sugars soften the bilayer and therefore lower the bending rigidity; though more recent experiments employing the x-ray scattering technique in the presence of sucrose and glucose have failed to show that they have any effect at all[134].

4.2.4 X-ray Scattering

Using X-ray scattering to measure the bending modulus of lipid membranes dates back to 1995[135]. Stacks of bilayers are mounted on solid substrates (usually glass, mica or silicon) such that their normals are aligned along one axis. X-rays are shone on the stack, along that axis, and the scattered intensity measured.

The scattered beam intensity can be written:

$$I(q) = S(q)\,|F(q_z)|^2\, q_z^{-1} \tag{4.15}$$

where $\boldsymbol{q}_z$ is the vector difference between the incoming wave vector and the outgoing wave vector, $F(q)$ is a form factor ie the contribution of the scattering attributable to the type of atom in to which the scattering electron belongs and $S(q_z)$ a structure factor, the contribution of the scattering attributable to structure in which the atoms are arranged. In addition, the scattering intensity is also affected by the bending fluctuations of the undulating bilayer which is why we are able to use the measured scattering intensity to calculate the bending modulus. One proposed theory for fitting the experimental xray scattering data is the Modified Caillé Theory[136] which allows us to relate the scattering to the bending modulus K_c via a Caillé factor η:

$$\eta = \frac{q^2 k_B T}{8\pi\sqrt{K_c K_a}} \tag{4.16}$$

with K_a once again being the area compressibility modulus.

Nagle has attempted to rationalise the numerous discrepancies in the experimentally measured values of K_c found in the literature[137]. He noted that values obtained from fluctuation analysis tend to be higher than values obtained with micropipette aspiration by a factor of about two. However, some hesitancy remains about the results from micropipette aspiration, due to the uncertainty introduced by artefacts, discussed above. X-ray scattering and micropipette aspiration appear to agree on values of K_c that are smaller than the ones obtained by fluctuation analysis.

All of these methods rely on the assumption that the membrane is a thin sheet lacking any internal structure. This is an acceptable assumption when one looks at what happens at long wavelengths, but breaks when one has to deal with length scales comparable to the bilayer thickness. In this second case, the internal degrees of freedom of each single lipid become important (e.g. lipid tilt, defined as the orientation of the lipid with respect to the bilayer normal). Lipid tilt oscillations contribute to the undulation spectrum, adding a new term that scales as q^{-2} and thus dominates at short wavelengths.

For fluctuation analysis, the typical wavelengths of the bilayer undulations are of the order of $10\,\mu m$, and the surface tension remains low; in such cases, the effect of tilt on the fluctuation power spectrum should be negligible, and therefore an accurate value of K_c can be obtained without consideration of lipid tilt[134].

For measurements using X-ray scattering, the undulations of bilayers in a stack are constrained by neighbouring bilayers, and this reduces the amplitudes of the undulations when the wavelength exceeds the lateral correlation length ξ, which is typically of the order of 5 nm, comparable to the length scale at which lipid tilt becomes important (around 2 nm). Indeed, Nagle et al. observed small but systematic differences between the X-ray scattering data and fits made without taking into account lipid tilt. The differences become more important as the signal-to-noise ratio increases[134], confirming the relevance of lipid tilt in the analysis of X-ray data. The addition of a contribution due to lipid tilt would, in principle, increase the value of K_c making the results obtained from xray scattering experiments more consistent with those obtained from fluctuation analysis.

With respect to the micropipette aspiration technique, Nagle et al[134] analysed existing data using a refined theory that includes the contribution of lipid tilt. Their analysis showed that, in the low tension regime, including a tilt degree of freedom does not alter the value of K_c.

In summary, X-ray scattering and micropipette aspiration experiments tend to result in a smaller value for the bending modulus compared to those conducted using fluctuation analysis. For X-ray scattering, the difference could be explained by the contribution of lipid tilt to the undulation spectrum discussed above.

For micropipette aspiration, the difference between the results obtained using fluctuation analysis have been historically ascribed to the effect of sucrose/glucose added to the sample. However, as discussed above, this has not been supported by the results of x-ray scattering experiments[134] in the presence of sucrose and glucose.

4.3 Experimental results for asymmetric bilayers

The bending moduli of symmetric membranes have been studied both experimentally using various methods of both membrane preparation and measurement[119,122,138] as well as via atomistic simulation[139–144]. However it has only relatively recently been possible to prepare asymmetric GUVs using phase transfer methods and for the first time measure the bending modulus of asymmetric bilayers experimentally[126].

Elani et al.[126] prepared GUVs of 4 compositions: (1) pure DOPC (18:1 PC); (2) pure POPC (16:0-18:1 PC); (3) symmetric 50:50 DOPC/POPC mixture; and, for the first time, (4) asymmetric where one leaflet was DOPC and the other POPC. The bending moduli obtained for the pure and symmetric bilayers accorded well with previous results. Surprisingly however, the bending modulus for asymmetric bilayers was found to be almost three times higher than for the symmetric; this increased stiffness was found whichever of the two lipids comprised the inner or outer leaflet (see figure 4.3 reproduced from Chem. Commun., 2015,51, 6976-6979 - published by the Royal Society of Chemistry).

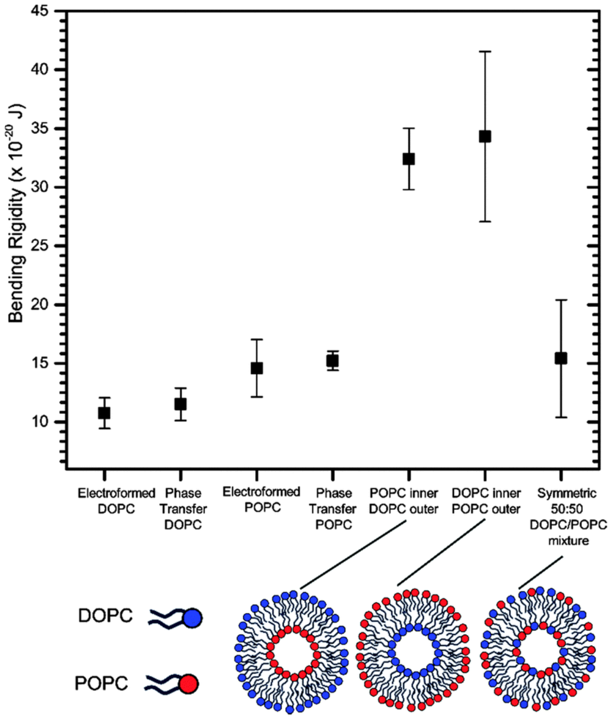


Figure 4.3: "Bending rigidities of symmetric and asymmetric GUVs formed via electroformation and phase transfer. For symmetric vesicles, the values obtained from the two methods correspond to one another within error. The bending rigidity is significantly higher ($p < 0.01$) for asymmetric vesicles compared to their symmetric counterparts. Error bars = standard deviation, where each vesicle is considered a single data point." Reproduced from DOI: 10.1039/C5CC00712G (Communication) Chem. Commun., 2015, 51, 6976-6979 - Published by the Royal Society of Chemistry.

The only other experimental study we have found which measured the bending modulus of asymmetric bilayers prepared GUVs in a similar way, but in this case, measured their bending rigidity using the micropipette aspiration technique[145]. They report bending moduli for asymmetric bilayers 50% higher than the pure or symmetric bilayers prepared using the same method. However, the bilayers they considered had quite different constituents from those considered by Elani et al. or indeed by the simulations in this chapter. The 2 constituents Lu et al.[145] considered were DMPC and DOPC and moreover, their experiments were conducted at 22.5°C, below the main phase transition temperature they quote for DMPC of 23.4°C. This means that they considered an asymmetric saturated gel phase lipid and unsaturated L_α phase lipid system, whereas Elani et al. and we are considering an asymmetric system where both leaflets are unsaturated and in the L_α phase.

4.4 Computational Methods

Four algorithms for the determination of the bending modulus of a lipid bilayer from simulation data are discussed in this section, all of which can be applied to trajectories from standard MD simulation and none of which require application of external forces: (1) a method based on polymer brush theory, in which the bending modulus is directly dependent on the area compressibility modulus of the bilayer; (2) and (3) two methods dependent on spectral analysis, involving 2D discrete Fourier transforms; and (4) a method dependent upon the distribution of splay angles between individual lipid molecules.

4.4.1 Polymer Brush method (PBM)

This model was proposed by Rawicz et al.[119] based on fitting experimental data derived from micropipette aspiration of GUVs prepared by simple rehydration. The model is commonly used due to its simple application and dependence on properties which are easily measured from simulation. The theoretical basis for the model is the assumption that the surface pressure in a fluid bilayer is dominated by confinement of chain entropy - neglecting van der Waals interactions between chains and specific interactions between the headgroups - and the treatment of the hydrocarbon chains as short, freely jointed polymers. The derivation of the model is based on the free energy of the hydrocarbon chains, assumed to be dependent on only the chain extension and the number of 'statistical' segments ($\approx$ 8 C-C bonds)[146]. This derivation (which can be found in full at the appendix to Rawicz et al.'s publication[119]) gives the expression for the bending modulus:

$$K_c = \frac{K_A h^2}{24} \tag{4.17}$$

where K_A is the area compressibility modulus (related to the ability to laterally compress the bilayer) and h is the hydrophobic thickness of the bilayer, given by

$h = h_{pp} - h_0$, where h_{pp} is the separation between the phosphates of the 2 leaflets and h_0 is a constant, commonly 1 nm. The value of h_0 is derived from fitting to the experimental data.

For a non-interacting pair of leaflets, the constant factor in the denominator is derived to be 48, reducing this to 24 results in the calculated bending rigidity being twice that of an unbonded pair of elastic layers reflecting the fact that the 2 leaflets do have some interaction which inhibits bending.

The area compressibility modulus, K_A can be calculated[147]:

$$K_A = A_L \left(\frac{\partial \gamma}{\partial A_L} \right)_T = \frac{2 K_B T A_L}{N_L \langle \delta A_L^2 \rangle} \tag{4.18}$$

where γ represents surface tension, A_L the mean area per lipid, $\langle \delta A_L^2 \rangle$ the mean square fluctuation in the area per lipid and N_L the total number of lipids in the system.

Determining accurate values of the bending modulus using this method can be challenging, due to the difficulty in obtaining correct values for the area compressibility modulus. The issue lies in the direct dependence of the modulus on the mean square fluctuation in the membrane area, and therefore, an accurate measurement requires very well sampled data[148] resulting in a requirement for large system sizes and/or long simulations in order to obtain accurate results.

4.4.2 Helfrich-Canham theory

A different approach to the calculation of the bending modulus is based on Helfrich-Canham theory[121,149], discussed in section 4.1.1 above; treating the membrane as a thin (by comparison to its lateral dimensions), structureless and homogeneous fluid sheet. The shape of the bilayer is quantified by its height field, $h(x, y) = h(\mathbf{r})$, which is determined by the displacement of the bilayer in the z-direction from a reference plane - usually the bilayer centre. Under conditions of vanishing surface tension, the power spectrum of the height field is[150]:

$$\langle |h_q|^2 \rangle = \frac{k_B T}{K_c q^4} \tag{4.19}$$

where q represents the Fourier wavenumber.

This equation is valid only where the underlying presumption that the bilayer can be treated as a thin, structureless sheet holds i.e. it is only valid in the limit where q tends towards 0. Higher values of q correspond to shorter wavelength undulations. Therefore large systems must be simulated. It has been estimated that the lateral box size would need to be approximately ten times larger than the bilayer thickness for accurate use of this method, corresponding to a system size of approximately 5000 lipids for DPPC[151].

Analogous to the improvement of the experimental measurements achieved using x-ray scattering by accounting for lipid tilt, it has been suggested that the poor performance of the HC method at wavelengths with a magnitude similar to, or greater than, the lateral dimensions of the simulation box is also caused by neglecting the effects of lipid tilting. An amendment to equation 4.19 has therefore been proposed[144,152].

$$\langle |h_q|^2 \rangle = K_B T \left(\frac{1}{K_c q^4} + \frac{1}{K_\theta q^2} \right) \tag{4.20}$$

where the tilt modulus, K_θ, is calculated using the method same algorithm employed by the SALO method below.

4.4.3 Spectral analysis of lipid orientations (SALO)

A different spectral analysis model has also been developed[142,144,151] to explicitly account for the effect of lipid tilt. It is described as being based "... on a reformulation of [the Helfrich Canham model] by changing basis within the theory to replace the bilayer height field with the longitudinal component of the vector field associated with lipid orientation."

In this method, an orientation vector, $\mathbb{N}_j$ is assigned to each lipid, j, defined as the vector between the midpoint between the end carbons of the 2 tails and the phosphate atom in the head group. $\mathbb{N}_j$ is projected, by considering the x and y components only, onto a 2D real space grid. The vector contents of each grid square are considered separately to give a realspace vector, $\hat{n}^\alpha$, for each square for each monolayer. The superscript denotes the upper or lower monolayer. Two dimensional (FFT) then yields a vector $\hat{n}^\alpha_q$ in Fourier space (hence the q

subscript) which is scaled by L/M^2 where M is the number of discrete points used in each dimension for the FFT and L is the x (or y since they are coupled) dimension of the bilayer.

The vector:

$$\hat{n}_q = \frac{1}{2}\left(\hat{n}_q^1 - \hat{n}_q^2\right) \tag{4.21}$$

is decomposed into longitudinal and transverse components. The power spectra of these properties are related to the bending modulus of the bilayer, as well as the lipid tilt, K_θ, and twist, K_{tw} moduli.

$$\langle \left|\hat{n}_q^{\parallel}\right|^2 \rangle = \frac{k_b T}{K_c q^2} \tag{4.22}$$

$$\langle \left|\hat{n}_q^{\perp}\right|^2 \rangle = \frac{k_b T}{K_\theta + K_{tw} q^2} \tag{4.23}$$

This method for calculating K_c is claimed to be much more accurate than HC, or even ‘modified-HC’ (ie using the Helfrich Canham formulation but including an extra term to take account of lipid tilt) in analysing trajectories from small bilayers. It has been reported that, using this method, an accurate value for the bending modulus was obtained from MD simulation of a bilayer composed of only 144 DPPC lipids per leaflet[144].

4.4.4 Real-space fluctuation analysis (RSF)

In recent years, another model has been developed, based on the calculation of lipid splays, as well as their tilt angles with respect to the bilayer normal[153–156]. The lipid splay, S_i, is defined as the covariant derivative of the vector field $\mathbf{n}-\mathbf{N}$ (where n represents a lipid director and N represents the bilayer normal) along one direction on the membrane interface:

$$S_i(r) = \lim_{h\rightarrow\ 0} \frac{n_i(\boldsymbol{r} + h\mathbf{e_i}) - n_i\mathbf{r} + N_i((r) + h\mathbf{e_i}) - N_i(\mathbf{r})}{h} \tag{4.24}$$

where $\mathbf{e_i}$ is a vector tangent to the membrane, and $n_i(r)$ and $N_i(r)$ are the components along $\mathbf{e_i}$ of the lipid director and membrane normal vector fields at position $\boldsymbol{r}$ on the membrane surface. Splays are calculated for all lipid pairs within 1 nm of each other, where at least one of the lipids is tilted by less than 10° with respect to the membrane normal. It can then be shown that the distribution of weakly correlated splays is given by equation 4.25:

$$P(S_i) = Ce^{-\frac{\chi_{ij} S_i^2 A_L}{2k_B T}} \tag{4.25}$$

where C is a constant and χ_{ij} is the splay modulus between the lipid species i and j present in the membrane. A Gaussian curve is fitted to the probability distribution, in order to determine the mean value, μ, and standard deviation, σ of the distribution, then, using the extracted μ and σ, a quadratic curve is constructed which is the potential of mean force (PMF) for lipid splay over the range of $(\mu - \sigma)$ to $(\mu + \sigma)$.

$$-k_B T ln(P(S_i)) = a + bS_i^2 \tag{4.26}$$

From this, the splay modulus can be extracted via $\chi_{ij} A_L = 2b$. For single component membranes, there is one value for the splay modulus, which is equal to the monolayer bending modulus, K_c, however, for multicomponent membranes, the monolayer bending modulus must be obtained as follows:

$$\frac{1}{K_c} = \frac{1}{\phi_{tot}} \sum_{i,j} \frac{\phi_{ij}}{\chi_{ij}} \tag{4.27}$$

where ϕ_{ij} is the number of splays calculated between species i and j, and ϕ_{tot} is the total number of splays calculated.

It has been reported that, using this method, accurate results can be obtained from systems containing only 128 lipids, with a difference of less than 1%, compared to systems of 400-500 lipids[154].

In order to attempt to replicate and investigate the results obtained by Elani et al[126] and to validate the alternative algorithms described above for deducing the membrane bending modulus from MD simulations, a variety of MD simulations were conducted and analysed as described below.

4.5 Molecular dynamics simulations

All simulations in described in tables 4.1 and 4.2 were carried out using the CHARMM36 lipid force field[47,48] and all those in tables 4.3 and 4.4 were conducted using the SLipid forcefield[40,49]. Unless otherwise noted, all simulations were conducted and analysed by Sophia Wheeler. The CHARMM36 forcefield was selected because CHARMM is a widely used forcefield for atomistic molecular dynamics simulations - both for membrane only systems and protein - membrane systems. SLipids was selected as an alternative forcefield to compare the CHARMM results to as it is, although less widely used than CHARMM, specifically parameterised for lipid systems. In addition, the atom naming conventions in SLipids are similar to those used in CHARMM and it is therefore convenient to use the CHARMM membrane builder to initialise the atom positions and use in-house scripts to convert them as required to use SLipids.

In order to compare our results with with the work of Elani et al[126], we conducted simulations using same 2 lipids as they used - DOPC and POPC. These are both PC lipids, described further in chapter 1. DOPC has 2 identical hydrophobic tails, both with 18 C atoms in total, with a single cis-double bond between C9 and C10. POPC has 2 different tails, 1 fully saturated with 16 C atoms and the other 18 C long with a single cis-double bond between C9 and C10. In addition, in order to investigate trends in the values of the bending modulus calculated by the different methods available, we also simulated pure DMPC and DPPC bilayers. DMPC has 2 fully saturated tails, each 14 carbons long. DPPC has 2 fully saturated tails, each 16 carbons long. Because DMPC and DPPC are both fully saturated lipids, their melting temperatures (T_m) from gel to L_α are both higher than for the partially unsaturated lipids. The temperatures used for the simulations are therefore also different in order to ensure that all the bilayers were in the same phase. The T_m for all lipids considered in this chapter is set out in table 3.1.

Systems were constructed, and input files generated, using the CHARMM-GUI Membrane Builder[157,158]. All simulations employed the TIP3P water model[159] and were run using the GROMACS simulation package, versions 5.0.4 or 5.1.1[160–165] with systems equilibrated using the standard protocol given by the CHARMM Membrane Builder, set out in detail at Appendix A. All production simulations were run in the NPT ensemble. For the production simulations, a semi-isotropic Parrinello-Rahman barostat[72,166] was used to maintain the pressure of the system at 1 bar, with a compressibility of $4.5\times10^{-5}\,\text{bar}^{-1}$, whilst the system temperature was maintained via the Nosé-Hoover thermostat[68,69] at the temperatures specified in the tables. The barostat and the thermostat employed coupling constants of 1 ps and 5 ps, respectively. At distances greater than 12 Å, electrostatic interactions were handled by the Particle Mesh Ewald (PME) method[167,168] and van der Waals interactions were not calculated at distances greater than this. Over a range of 8 to 12 Å, the Lennard-Jones interactions were scaled to zero via the force-based switching method[169]. Bonds involving hydrogen atoms were constrained via the LINCS algorithm[170,171].

We constructed 3 different asymmetric systems. For all of them one leaflet was purely DOPC and the other, POPC. For the first type, the area of the upper and lower leaflets was matched and therefore, because the area per lipid of DOPC and POPC are different, there were different numbers of lipids on the upper and lower leaflets. In order to calculate the numbers of lipids required, the area for DOPC and POPC was obtained by taking an average of the area per lipid in the pure DOPC and POPC simulations; for the second type, the number of lipids in the upper and lower leaflet was matched and finally, we explored the effect of tension on the bilayer by deliberately mismatching the area of the upper and lower leaflets i.e. having many more DOPCs (larger area per lipid) on one leaflet than POPCs (smaller area per lipid) on the other.

Analysis of the trajectories resulting from both the CHARMM36 and SLipids simulations were conducted using code written in-house (by Marley Samways and Jonathan Shearer) for the calculation of K_c by the HC, modified HC, SALO and PBM methods. For the RSF method, the code implementation released by the authors of the original paper was employed[153].

Lipid Type	Lipids per leaflet	Temperature / K	Simulation time / ns	Equilibration time / ns	Number of repeats
DPPC	64	323	200	10	3
DPPC	324	323	200	10	3
DOPC	64	298	200	10	3
DOPC	324	298	200	10	3
DOPC	1024	294	100	20	2
POPC	64	298	200	10	3
POPC	324	298	200	10	3
POPC	1024	294	100	20	2
DMPC	324	308	200	10	3

Table 4.1: Simulation details for each of the single-lipid bilayer systems simulated in this chapter using CHARMM36. All bilayers containing 64 lipids per leaflet, all DMPC bilayers and one repeat of each of the POPC and DPPC 324 systems were constructed and simulated by Marley Samways. The total production simulation time for each is given by the *Simulation time* column; the time given in the *Equilibration time* column represents the length of time discarded at the beginning of the production simulation, before using the remainder of the trajectory to perform the K_c calculations.

	Upper leaflet		Lower leaflet					
	Lipid Type(s)	No.	Lipid Type(s)	No.	Temp. / K	Sim. time / ns	Equil. time / ns	Reps
A	DOPC POPC	512 512	DOPC POPC	512 512	298	100	20	2
B	DOPC	1024	POPC	1102	298	100	20	2
B	DOPC	512	POPC	551	298	200	10	1
C	DOPC	512	POPC	512	298	200	10	1
D	DOPC	324	POPC	220	298	200	10	1

Table 4.2: Simulation details for each of the mixed lipid bilayer systems simulated in this chapter using CHARMM36. The first column refers to the system type; A being symmetric, mixed bilayer systems; B asymmetric systems where the total areas of the upper and lower leaflets have been matched using the areas per lipid calculated from the relaxed single component systems; C asymmetric systems where the number of lipids on the upper and lower leaflets have been matched; and D asymmetric systems where the numbers of POPC have been deliberately reduced to exaggerate the mismatch in area between the leaflets. The total production simulation time for each is given by the *Simulation time* column; the time given in the *Equilibration time* column represents the length of time discarded at the beginning of the production simulation, before using the remainder of the trajectory to perform the K_c calculations.

Lipid Type	Lipids per leaflet	Temperature / K	Simulation time / ns	Equilibration time / ns	Number of repeats
DOPC	1024	294	100	20	2
POPC	1024	294	100	20	2

Table 4.3: Simulation details for each of the symmetric bilayer systems simulated in this chapter using the Slipid forcefield. The total production simulation time for each is given by the *Simulation time* column; the time given in the *Equilibration time* column represents the length of time discarded at the beginning of the production simulation, before using the remainder of the trajectory to perform the K_c calculations.

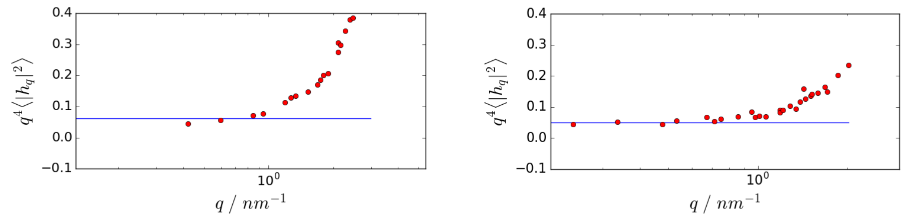
	Upper leaflet		Lower leaflet					
	Lipid Type(s)	No.	Lipid Type(s)	No.	Temp. / K	Sim. time / ns	Equil. time / ns	Reps
A	DOPC POPC	512 512	DOPC POPC	512 512	298	100	20	2
B	DOPC	1024	POPC	1072	298	100	20	2
B	DOPC	512	POPC	536	298	150	10	1

Table 4.4: Simulation details for each of the mixed lipid bilayer systems simulated in this chapter using the SLipid forcefield. The first column refers to the system type; A being symmetric, mixed bilayer systems and B asymmetric systems where the areas of the upper and lower leaflets have been matched using the areas per lipid calculated from the relaxed single component systems. The total production simulation time for each is given by the *Simulation time* column; the time given in the *Equilibration time* column represents the length of time discarded at the beginning of the production simulation, before using the remainder of the trajectory to perform the K_c calculations.

4.6 Results

4.6.1 Convergence criteria

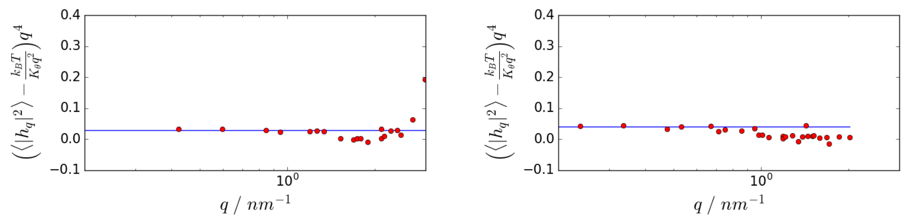
The calculation using the HC method only converged for the largest system size due to the size dependency of this methods discussed above, see figure 4.4. The other spectral methods (SALO and mod HC) did not yield results for the smallest system size because both methods involve discretising lipid vectors onto a grid and require that no adjacent grid spaces are empty. For very small sizes, this is not possible. However, they were applicable to system sizes of at least 648 lipids as the graphs in figures 4.5 and 4.6 show.



(a) HC convergence for 648 lipid system

(b) HC convergence for 2048 lipid

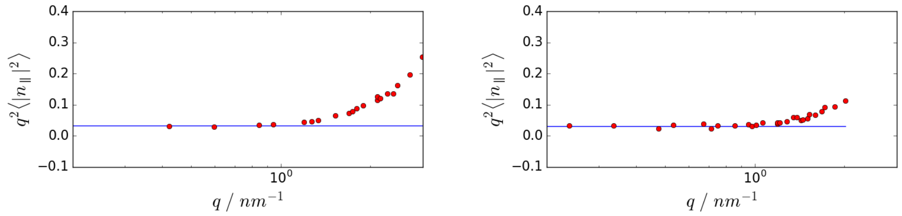
Figure 4.4: Following equation 4.19, if the calculation is converged, a graph of $q^4\langle|h_q|^2\rangle$ against q should converge to a line of gradient 0 and y-intercept $\frac{k_BT}{K_c}$ for at least the smallest 4 wave numbers. As shown in the graphs above, this is not the case for the 648 lipid system, but is for the 2048 lipid system.



(a) modHC convergence for 648 lipid system

(b) modHC convergence for 2048 lipid

Figure 4.5: Following equation 4.20, if the calculation is converged, a graph of $\langle|h_q|^2\rangle - \frac{k_BT}{K_\theta q^2}$ against q should converge to a line of gradient 0 and y-intercept $\frac{k_BT}{K_c}$ for at least the smallest 4 wave numbers. As shown in the graphs above, this is the case for both the 648 lipid system, and the 2048 lipid system.



(a) SALO convergence for 648 lipid system

(b) SALO convergence for 2048 lipid

Figure 4.6: Following equation 4.22, if the calculation is converged, a graph of $\langle\left|\hat{n}_q^{\|}\right|^2\rangle$ against q should converge to a line of gradient 0 and y-intercept $\frac{k_BT}{K_c}$ for at least the smallest 4 wave numbers. As shown in the graphs above, this is the case for both the 648 lipid system, and the 2048 lipid system.

4.6.2 Symmetric bilayers (CHARMM36)

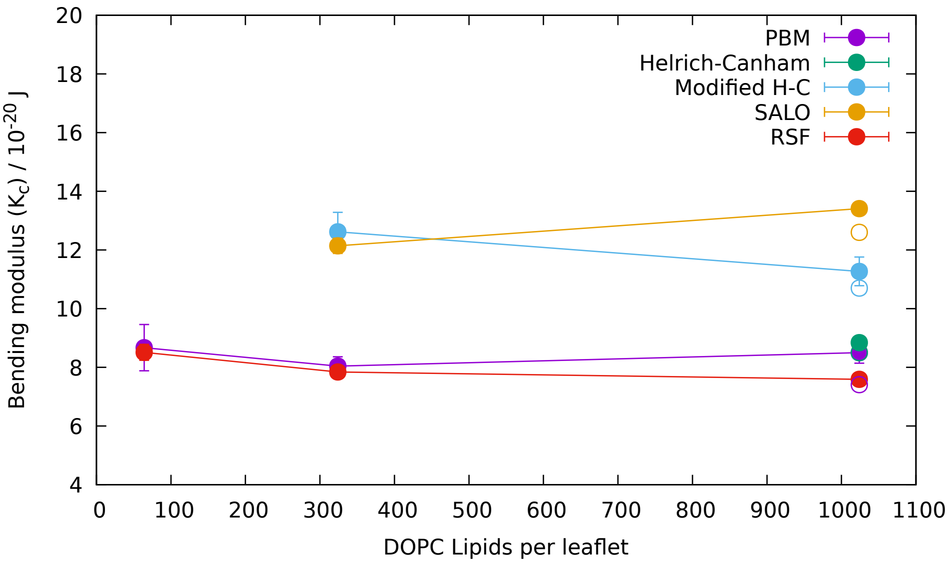


Figure 4.7: The results of the bending modulus calculations from the CHARMM36 simulations (closed symbols) and Slipid (open symbols) of DOPC (tabulated in table 4.5 and table 4.7 respectively).

As can be seen from the results in table 4.5 and figure 4.7, all the calculation methods yield a value for K_c for pure DOPC bilayers between $7.6 \pm 0.1 \times 10^{-20}$ J and $13.4 \pm 0.1 \times 10^{-20}$ J. For those methods where it was possible to calculate a value for the bending modulus, the results obtained were not affected by system size. Published experimental measurements for the DOPC bending modulus cover a similar range of values, depending upon the measurement method employed.

		Bending modulus K_c / $10^{-20}J$				
Lipid Type	Lipids per leaflet	PBM	Helfrich -Canham	SALO	modified HC	RSF
DPPC	64	6.6±0.4	-	-	-	15.6 ± 0.2
DPPC	324	7.4±0.2	-	15.9±0.1	-	15.6 ± 0.2
DOPC	64	8.7±0.8	-	-	-	8.5 ± 0.3
DOPC	324	8.0±0.3	-	12.1±0.3	12.6±0.7	7.8 ± 0.1
DOPC	1024	8.5±0.4	8.8±0.2	13.4±0.1	11.3±0.5	7.6 ± 0.1
POPC	64	9.4±0.4	-	-	-	10.5 ± 0.3
POPC	324	8.3±0.7	-	13.6±0.5	14.5±1.3	10.6 ± 0.2
POPC	1024	6.4±0.2	8.9±0.8	15.5±0.6	12.6±1.3	10.1 ± 0.2
DMPC	324	5.7±0.3	-	13.5±0.2	14.0±1.1	14.2 ± 0.1

Table 4.5: The results of bending modulus calculations for the simulations of pure single lipid bilayers simulated using the CHARMM36 forcefield for various system sizes. Errors are a standard deviation calculated over the simulation repeats.

For x-ray scattering experiments[134], values between $5.53 \pm 0.3 \times 10^{-20}$ J and $6.84 \pm 0.5 \times 10^{-20}$ J; for micropipette aspiration[119,172] between $5.0 \pm 0.2 \times 10^{-20}$ J and $9.2 \pm 0.1 \times 10^{-20}$ J; for fluctuation analysis between 10.8×10^{-20} J and 12.0×10^{-20} J.

As can be seen from the results in table 4.5 and figure 4.8, all the calculation

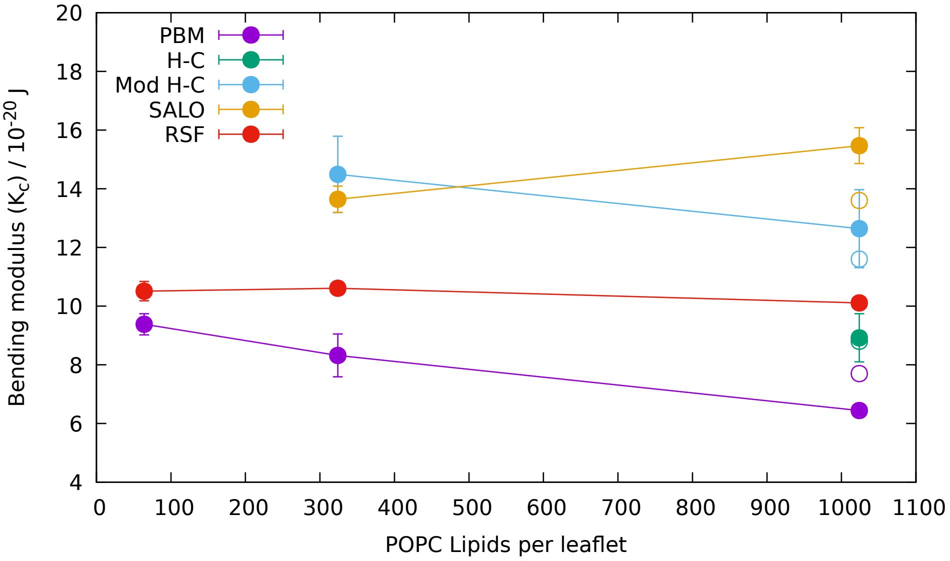


Figure 4.8: The results of the bending modulus calculations from the CHARMM36 simulations (closed symbols) and SLipid (open symbols) of POPC (tabulated in table 4.5 and 4.7 respectively).

methods yield a value for K_c for pure POPC bilayers between $6.4 \pm 0.2 \times 10^{-20}$ J and $15.5 \pm 0.6 \times 10^{-20}$ J. Like DOPC, for those methods where it was possible to calculate a value for the bending modulus, the results obtained were not materially affected by system size. Published experimental measurements for the POPC bending modulus cover a similar range of values, depending upon the measurement method employed. For x-ray scattering: values between 7.7×10^{-20} J and 8.53×10^{-20} J; for micropipette aspiration between 2.46×10^{-20} J and 21.1×10^{-20} J; for fluctuation analysis between 10.0×10^{-20} J and 21.1×10^{-20} J.

Given the spread of published experimental results and the spread of results obtained by calculation, we considered the comparison of the different methods may be clearer if we investigated the trends in measured K_c and calculated K_c as lipid tail length increases (DMPC to DPPC) and as tail saturation decreases (DPPC to POPC to DOPC) rather than comparing absolute values. Results from x-ray scattering experiments, whether including a correction for lipid tilt or not, and fluctuation analysis show a significant difference and we have therefore considered these separately. These trends are shown in figs 4.9 and 4.10. We have, as far as possible, only compared results from one experimental group for each trend line so that artefacts introduced by the preparation methods (often not reported in detail) employed for the GUVs or bilayer stacks are minimised.

The graph in figure 4.9 shows the effect on bilayer bending modulus of increasing

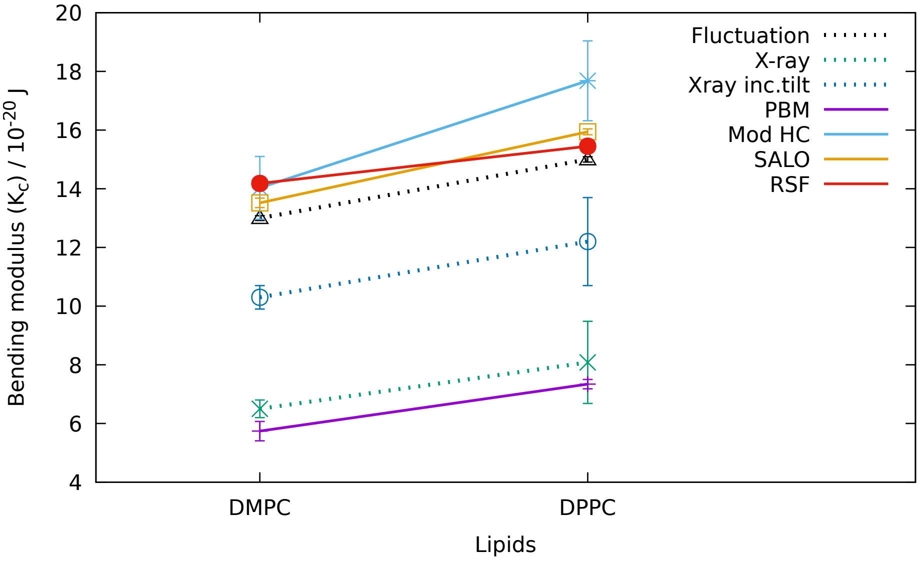


Figure 4.9: Graph showing calculated (solid lines) and experimentally measured (dashed lines) K_c values for DMPC and DPPC. CHARMM36 simulations of 324 lipids per leaflet were used for these calculations.

the tail length of the constituent saturated lipids. For reasons of computational availability, the system sizes simulated for these results was the smallest for which all the analysis methods were available. As seen in figures 4.7 and 4.8, the dependency of the results on system size is small for all the analysis methods. Both fluctuation analysis[173] and x-ray scattering experiments[174] show an increase in K_c as chain length increases for saturated lipids. All simulation analysis methods investigated showed the same trend with RSF and SALO methods giving results in closest agreement with fluctuation analysis.

The graph in figure 4.10 shows the effect on bilayer bending modulus of decreasing the tail saturation of the constituent saturated lipids. Both fluctuation analysis[126,173] and x-ray scattering experiments[174] show an decrease in K_c as tail saturation decreases. With the exception of PBM, all simulation analysis methods investigated showed the same trend with the SALO and mod H-C methods giving results in closest agreement with fluctuation analysis. The RSF method, whilst initially generating results in good agreement with fluctuation analysis for the saturated lipid (DPPC), as tail unsaturation increases, seems to increasingly underestimate K_c.

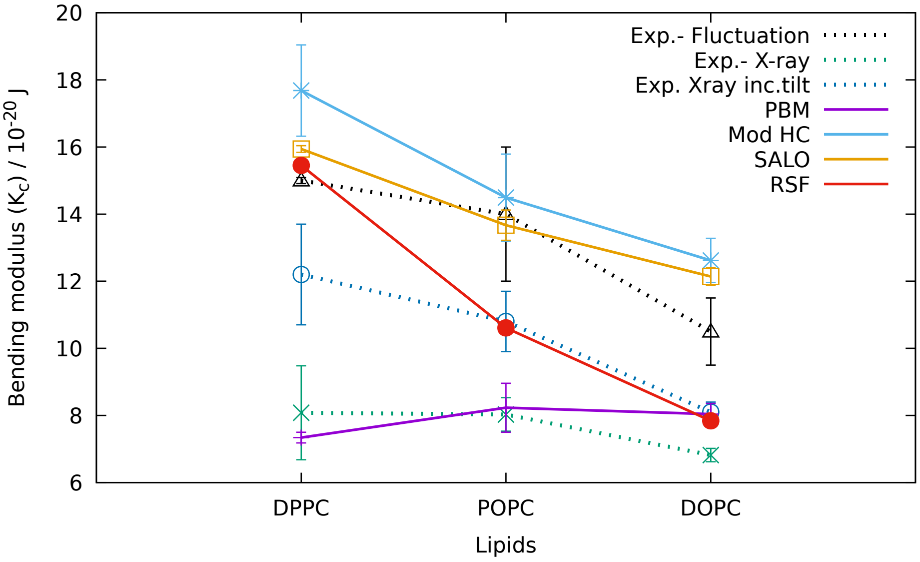


Figure 4.10: Graph showing calculated (solid lines) and experimentally measured (dashed lines) K_c values for DPPC, POPC and DOPC. CHARMM36 simulations of 324 lipids per leaflet were used for these calculations.

4.6.3 Mixed symmetric and asymmetric bilayers (CHARMM36)

There has been much discussion recently as to the most appropriate methods to ensure that asymmetric bilayers in simulations are indeed tensionless and therefore are appropriate systems on which to use the analysis methods described above; all of which include the assumption that the membrane is tensionless in their derivation. Rather than employing methodologies to ensure each monolayer was tensionless and thereby the resultant bilayer was tensionless[175], we firstly followed the widely used protocol[176] of simulating systems where we matched the total areas of the upper and lower leaflets as closely as possible.

First, we needed to establish which areas should be used for DOPC and POPC. There are some discrepancies in the areas per lipid to be attributed to DOPC and POPC experimentally. Kucerka et al.[177] report 68.3 Å^2 for POPC and 72.4 Å^2 for DOPC. However, in a later paper, the same authors revised the area for POPC to 62.7 Å^2 and that for DOPC to 67.4 Å^2[178], explaining the discrepancy by a ‘limited data set’ in the first paper. Then Herberle et al.[179] presented an improved set of area per lipid data based upon scattering density profile measurements rather than stand alone scattering experiments and confirmed the areas per lipid given in the later Kucerka et al. paper[178]. The CHARMM Membrane builder uses 68.3 Å^2 and 69.7 Å^2 for POPC and DOPC respectively as standard.

Given these differences, rather than applying either the area per lipid quoted in experimental papers for these lipids, or the areas assigned by the CHARMM Membrane Builder software, we analysed the production data from our own simulations of the pure DOPC and POPC systems to obtain a simulated area per lipid for each species and then using those simulated areas per lipid, we constructed the matched-area asymmetric bilayers (type B further described below) by specifying in the CHARMM Membrane builder the number of DOPC and POPC molecules in the upper and lower leaflet to ensure the total area of each leaflet was as similar as possible. In addition, in order to investigate what effect membrane tension would have on the methods, we simulated membranes (type C and D further described below) where we knew by the area mismatch between the areas of the leaflets that tension was definitely present.

Table 4.6 and figure 4.11 show the results for the bending modulus calculations of the mixed DOPC/POPC systems. We ran 3 different asymmetric systems; for type B the area of the upper and lower leaflets was matched as described above; for type C, the number of lipids in the upper and lower leaflet was matched and for type D, we explored the effect of tension on the bilayer by deliberately greatly mismatching the area of the upper and lower leaflets by placing more DOPC (larger area per lipid) in the upper leaflet than POPC in the lower.

Table 4.6 and figure 4.11 show the K_c results we obtained for the pure and mixed symmetric systems using the CHARMM36 forcefield. Those calculated for the pure DOPC and POPC systems using the SALO and mod-HC methods are consistent with those reported by Elani et al of $\approx$ 11 to 12 $\times$ 10^{-20}J and $\approx$ 15 to 16 $\times$ 10^{-20}J respectively (see figure 4.3). Even for these large system sizes, not only do the HC and PBM methods underestimate the values of K_c but they also both fail to reproduce the trend shown in Elani et al's data that the K_c of a pure POPC bilayer is higher than that of a pure DOPC bilayer. For the mixed symmetric systems, the results we found using the mod-HC and SALO analyses accords with Elani et al reported values of $K_c = 15 \pm 5 \times 10^{-20}$J but once again, the PBM and HC methods produce a somewhat lower value. We were not be able to reproduce the dramatic increase in K_c reported by Elani et al in any of the asymmetric simulations using any of the analysis methods. In addition our results on the type C and D systems show that unexpected tension in the system is unlikely to be the cause of the lack of increase in K_c. Comparing the

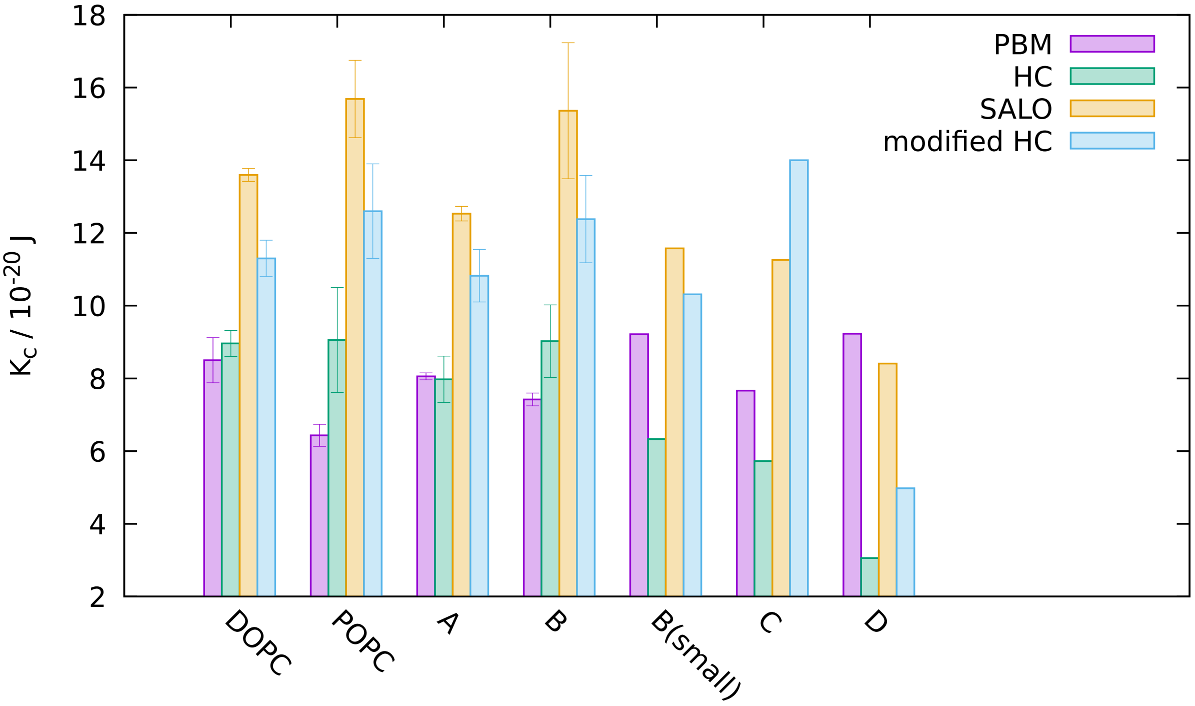


Figure 4.11: Bending moduli calculated from our pure DOPC, pure POPC, symmetric 1:1 mix DOPC:POPC (type A) and asymmetric systems simulated using the CHARMM36 forcefield.

results for each analysis method across the type B, type C and type D simulations, ie increasing the known tension in the bilayer, we found that increasing tension has either no effect on the calculated K_c, as in the PBM method results or, for the other analysis methods, causes the calculated value of K_c to fall.

		Bending modulus K_c / $10^{-20}J$			
Sim. Type	Total lipids	PBM	Helfrich -Canham	SALO	modified HC
DOPC	2048	8.5±0.6	9.0±0.3	13.7±0.7	11.3±0.5
POPC	2048	6.4±0.3	9.1±0.5	15.6±1.0	12.6±1.3
A	2048	8.1±0.2	8.0±0.6	12.5±0.2	10.8±0.7
B	2126	7.4±0.2	9.0±1.0	15.8±1.5	12.3±1.2
B	1063	9.2	6.3	11.8	10.3
C	1024	7.7	5.7	11.3	14.0
D	548	9.2	3.1	8.41	4.98

Table 4.6: Bending moduli data from the CHARMM36 simulations. The *Sim. Type* column refers to whether the bilayer is pure (DOPC and POPC); mixed-symmetric (type A); asymmetric, area of inner and outer leaflet matched (type B); asymmetric, number of lipids on inner and outer leaflet matched (type C) and; asymmetric, with an increased number of POPC compared with DOPC (type D)

4.6.4 SLipid simulations

Though the K_c results obtained from simulations using the SLipid forcefield set out in table 4.7 and figure 4.12 are slightly lower overall than those obtained using the CHARMM36 forcefield, the results for the pure and mixed symmetric systems are again consistent with those reported experimentally. Once again, we have not be able to reproduce the increase in K_c in any of the asymmetric systems simulated. For this forcefield, increasing the tension of the bilayer did not appear to have any effect on the calculated K_c values for any of the calculation methods.

		Bending modulus K_c / $10^{-20} J$			
Sim. Type	Total lipids	PBM	Helfrich-Canham	SALO	modified HC
DOPC	2048	7.41±1.6	8.5±0.4	12.6±0.3	10.7±0.4
POPC	2048	7.7±0.1	8.8±0.2	13.6±0.2	11.6±0.1
A	2048	8.7±0.3	7.5±0.4	10.8±0.8	9.2±0.8
B	2096	9.3	7.1	12.1	11.2
B	1048	8.7	5.0	11.8	9.8
C	2048	6.4	7.8	12.1	9.7
C	648	10.9	6.14	11.9	12.6

Table 4.7: These are the results from our pure DOPC, pure POPC, symmetric 1:1 mix DOPC:POPC (type A) and asymmetric systems simulated using the SLipid forcefield. Two different asymmetric systems were run; for type B the area of the upper and lower leaflets was matched; for type C, the number of lipids in the upper and lower leaflet was matched.

4.6.5 MARTINI simulations

Having been unable to reproduce in simulations the increase in bending modulus observed experimentally, it was hypothesised that this may be due to the atomistic systems being too small to observe the effect. However, simulating larger systems atomistically would be too computationally expensive and it was therefore decided to employ a CG forcefield, in this case, MARTINI, to explore larger system sizes. Three different system sizes were prepared for each of: pure POPC, DOPC, symmetric mixed (type B) and asymmetric (type C) bilayers using the MARTINI Maker website[157,180,181], the smallest being the same size as the largest atomistic systems simulated in order to aid comparison of results. For the asymmetric systems, areas of the leaflets were matched by calculated the areas per lipid

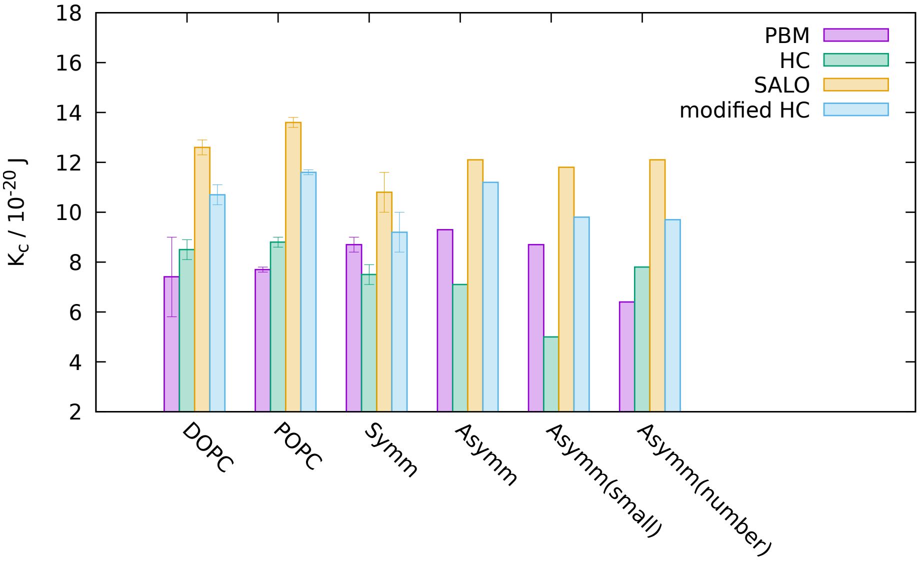


Figure 4.12: These are the results from our pure DOPC, pure POPC, symmetric 1:1 mix DOPC:POPC (type A) and asymmetric systems simulated using the SLipid forcefield. We ran 3 different asymmetric systems; for type B, the areas of the upper and lower leaflets were matched; for type C, the number of lipids in the upper and lower leaflet was matched.

of the pure systems and using those areas to calculate the numbers of lipids required for each of the leaflets. All simulations were equilibrated using the protocol in Appendix B. Production simulations were then run for 200 ns, with a timestep of 20 fs, at 298 K, with the temperature of lipids and water thermostated separately using a Berendsen thermostat. Pressure was maintained using a semi-isotropic Berendsen barostat at a pressure of 1 atmosphere. The results of these MARTINI simulations are shown in table 4.8 and figures 4.13, 4.14 and 4.15.

As can be seen in figure 4.13, where the results of the MARTINI simulations are directly compared to those from the atomistic simulations, the K_c values calculated by the SALO method for the MARTINI bilayers are somewhat lower for those bilayers containing POPC than the K_c values calculated for atomistic systems. In contrast the values for K_c calculated via the PBM method from the MARTINI systems are somewhat higher. Nonetheless, even taking these quantitative differences into account, there is no indication in the results from the medium and large sized MARTINI systems that simulated asymmetric bilayers have significantly larger K_c than symmetric bilayers.

Fowler et al[182] have also reported bending moduli for POPC bilayers calculated

		Bending modulus K_c / $10^{-20} J$			
Sim. Type	Total lipids	PBM	Helfrich-Canham	SALO	modified HC
DOPC	2048	10.59	8.30	11.00	9.58
DOPC	4096	8.15	9.87	10.66	12.30
DOPC	8192	7.64	11.12	13.09	11.60
POPC	2048	11.55	9.66	13.01	11.55
POPC	4096	10.08	11.19	14.28	12.34
POPC	8192	8.27	12.77	15.63	13.44
A	2048	9.10	9.13	12.03	10.72
A	4096	10.47	9.00	11.47	9.71
A	8192	11.00	10.33	12.44	10.76
B	2196	9.97	9.25	12.30	10.75
B	4231	9.41	10.91	13.78	11.90
B	8399	7.84	11.89	14.66	12.45

Table 4.8: These are the results from our pure DOPC, pure POPC, symmetric 1:1 mix DOPC:POPC (type A) and asymmetric systems (type B) simulated using the MARTINI forcefield.

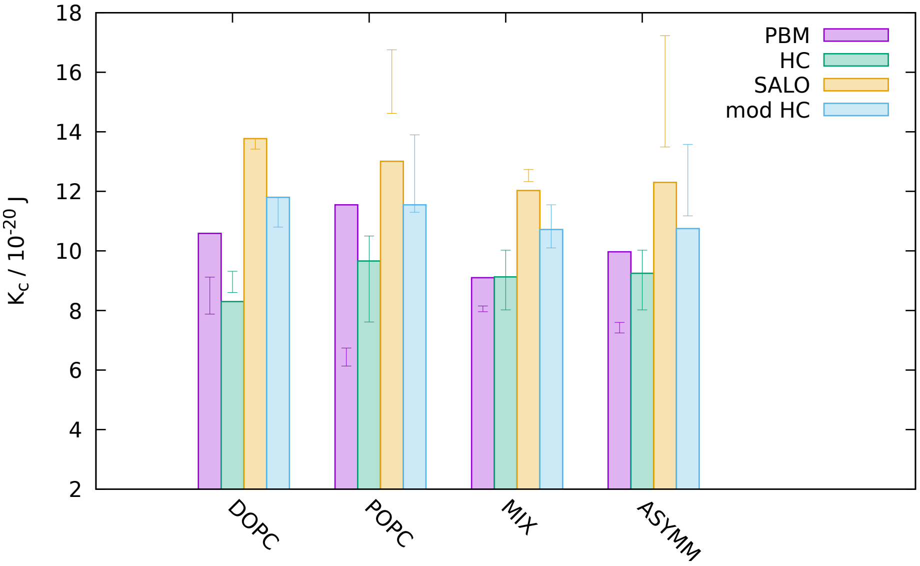


Figure 4.13: These are the results from the smallest (*circa* 2000 lipid) CG systems of pure DOPC, pure POPC, symmetric 1:1 mix DOPC:POPC (type A) and asymmetric systems simulated using the MARTINI forcefield as tabulated in table 4.8. The error bars plotted on the graph are the error bars from the atomistic simulations of the same type and size for ease of comparison.

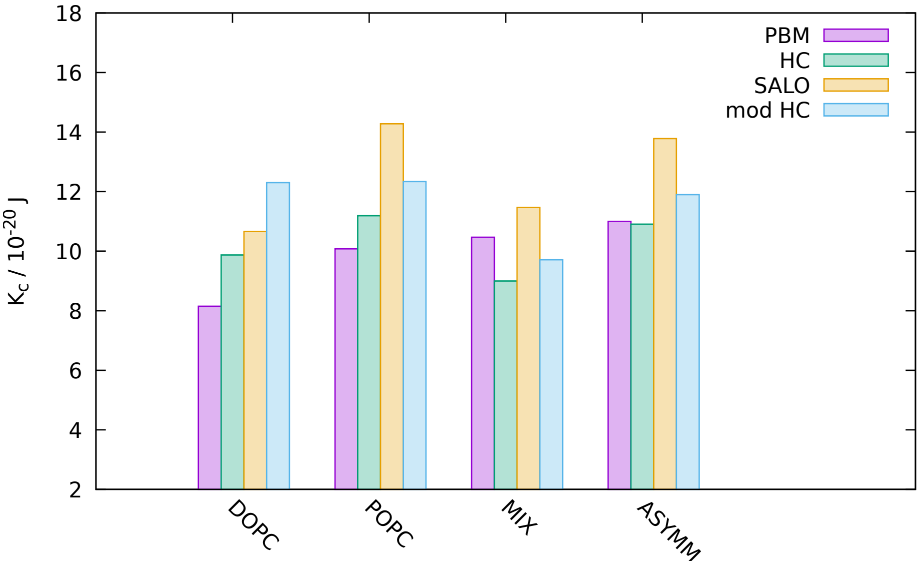


Figure 4.14: These are the results from the medium sized (*circa* 4100 lipids) MARTINI systems pure DOPC, pure POPC, symmetric 1:1 mix DOPC:POPC (type A) and asymmetric systems simulated using the MARTINI forcefield as tabulated in table 4.8

using the Helfrich-Canham method. They simulated bilayers both atomistically (CHARMM36, approx 1500 lipids) and coarse-grained (MARTINI, circa 55000

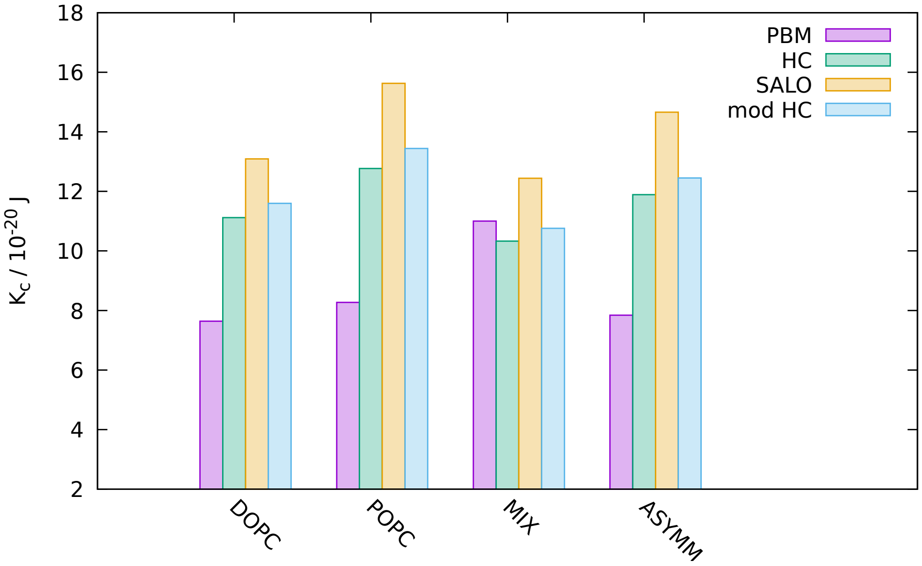


Figure 4.15: These are the results from the largest (*circa* 8200 lipids) systems pure DOPC, pure POPC, symmetric 1:1 mix DOPC:POPC (type A) and asymmetric systems simulated using the MARTINI forcefield as tabulated in table 4.8

lipids). For both sets of simulation, they report a bending modulus of approximately 30 $k_B T$ ie ≈12 × 10^{-20}J which accords well with the values we calculate for POPC bilayers.

4.7 Discussion

There has been discussion in the literature[122,134,174] recently regarding the differences in experimental measurements for K_c. Across different groups results from pipette aspiration and x-ray scattering are significantly lower, in some cases by a factor of 2 or more, than those from fluctuation analysis[134,174]. Some hypotheses for the differences have been postulated, including the presence of sucrose/glucose in the pipette aspiration experiments or the influence of neighbouring bilayers in x-ray scattering experiments, both possibly resulting in artificially low K_c results when compared with those from fluctuation analysis.

However Nagle has presented results and analysis[137] which suggest that any effect caused by the presence of sucrose/glucose in the pipette aspiration experiments is insufficient to explain the difference between those results and those from fluctuation. In addition, he showed that the inclusion of a factor representing the lipid tilt modulus in the analysis of x-ray scattering results raised the resulting K_c but not sufficiently to approach the values from fluctuation analysis. Neither fluctuation analysis nor pipette aspiration methods require the addition of a lipid tilt factor. The source of the different results between experimental methods therefore remains elusive, but the implication seems that the results from fluctuation analysis are 'correct' and those from the other methods, even when lipid tilt are included, ia 'artificially low' due to a hitherto unconsidered artefact.

Turning to the analysis methods applied to simulation trajectories in this chapter, we have shown that, with the exception of Helfrich-Canham, each of the methods result in similar values for K_c whether the bilayer size is around 300 lipids per leaflet or around 1000. The Helfrich-Canham, when modified to include lipid tilt, SALO and RSF methods yield results which are in closest agreement with fluctuation analysis experiment for saturated lipids. However the RSF method appears to underestimate K_c when considering lipids with a degree of unsaturation in their tails. The results obtained with all the methods are consistent across the atomistic CHARMM36 and SLipid forcefields and also across the CG MARTINI forcefield - with the slight caveat for the latter that the SALO method used with MARTINI does somewhat underestimate the K_c value for bilayers containing POPC.

With respect to the results of bending modulus calculation for mixed and asymmetric bilayers, our results agree well for the mixed symmetric bilayers but we have been unable to reproduce the significant increase in K_c reported by Elani et al. for the asymmetric bilayers. It has been suggested[183] that the increase is due to the way that the spontaneous curvature of the DOPC and POPC lipids is considered. It is said that, in a symmetric system, only the spontaneous curvature of a single lipid is considered (even in mixed systems) whereas an asymmetric system should more appropriately be considered as 2 monolayers with independent spontaneous curvature values which should be added, as should the bending modulus for each monolayer. A symmetric bilayer is a lower energy system due to less curvature elastic stress and therefore has a lower bending rigidity compared to that of an asymmetric bilayer.

None of the methods we have used consider the 2 monolayers separately; PBM considers the bilayer as a 'modified single slab' whilst SALO and mod-HC effectively take an average of lipid-tilt vectors for the upper and lower leaflet at intermediate stages in the calculation. Adjustments to these methods should be investigated, treating each monolayer independently and summing the resultant K_c values in order to ascertain whether such an approach allows the retention of the accuracy achieved for symmetric bilayers, whilst also capturing the anomalous increase in K_c seen in asymmetric bilayers.

Consideration of the leaflets separately has been used by Sreekuman and Lipowsky[184] to calculate the bending modulus of asymmetric leaflets. They considered a membrane, modelled using a DPD forcefield, consisting of a mixture of lipids with small and large headgroups. For their asymmetric systems, they maintained the same total number of lipids in the upper and lower leaflets in all simulations but varied the mole fraction of the larger head group lipid from 0 to 1. They noted that the bilayers so produced were under tension; in the same way that those asymmetric systems simulated in this thesis where the areas of the upper and lower leaflets were mismatched were under tension. They calculated the bending modulus from the first moment of the stress profile for each monolayer and combined them to find the bending modulus for the bilayer. They did not show any increase in bending modulus for the asymmetric tensioned bilayers they simulated.

4.8 Conclusions

We have confirmed that the original analysis method, Helfrich-Canham, used to calculate K_c from simulation data requires a large system size. However, we have shown that for systems of at least 600 lipids, more recently developed methods which include a correction for lipid tilt and/or splay display little system size dependency. These methods also correctly reproduce the trends in K_c for increasing lipid tail length and decreasing lipid tail saturation. Establishing quantitative agreement between the values given by the different calculation methods and experimental methods has proved difficult because both the experimental results and the results from simulation display wide variability depending on the measurement and or calculation algorithm employed. Nonetheless, our the calculated K_c values are in broad agreement with those reported by Elani et al. save that for asymmetric systems, we have been unable to reproduce the 2 to 3 fold increase in K_c that they observed.

Chapter 5

Binary cholesterol and phospholipid mixtures

5.1 Cholesterol and phospholipid binary mixtures

It is believed that cholesterol plays a vital role in the function of mammalian cells[185–188]. The importance of cholesterol is hypothesised from its ubiquity - cholesterol, or other sterols such as ergosterol or phytosterols, is universally present in eukaryotic cell membranes in amounts up to 40 mol%, whilst being universally absent in prokaryotes[188].

It is known that cholesterol and other sterols modify lipid membrane material properties, such as fluidity[189] and bending rigidity[190,191]. In particular, it is hypothesised that, in the absence of cholesterol, a cell wall is required in the prokaryotic cell envelope in addition to two (for Gram-negative bacteria) or one (for Gram-positive bacteria) lipid membranes because the absence of cholesterol means that lipid membranes alone are too flexible and permeable. In the absence of cholesterol, an extra, more rigid container must therefore be included in the cell envelope. Depending on the prokaryote, such cell walls may consist of polysaccharides and/or glycoproteins[17].

The origin of the modification of bilayer properties is believed to be because the introduction of cholesterol into a phospholipid bilayer induces cholesterol-concentration dependent phase behaviour in binary mixtures[192]. Experiments in model systems have established[192–198] that above the phase transition temperature of the phospholipid and at low (< 10 mol%) cholesterol concentration, binary phospholipid/cholesterol mixtures are in the liquid disordered (L_d) or fluid phase; characterized by disordered lipid tails, fast lipid mobility, and relatively thin bilayers. In contrast, above the phospholipid main phase transition temperature and at high ($> 30\%$) cholesterol content, they are in a "liquid ordered" (L_o) phase. In the L_o phase, the lipids retain their fast diffusion, but, as in gel phase membranes, the bilayer is thicker, with reduced area per lipid, and straight, ordered hydrocarbon lipid tails. At intermediate cholesterol concentrations, there is an argument that these liquid disordered and liquid ordered phases co-exist in phase separated region of the phase diagram[25,199], see figures 5.6 and 5.7. However, there is also a hypothesis that there is no phase separated region above the main phase transition temperature and the transition is smooth across the concentration range[196,200,201].

In addition, it has long been proposed that the presence of cholesterol compartmentalises cell membranes into lipid domains[16]. In particular, some studies have identified cholesterol enriched functional microdomains, or "rafts" in multi-component model membranes containing both phospholipids and sphingomyelins[16,37,187,202,203]. One of the roles proposed for these rafts has been that by providing dynamic platforms, where particular raft-specific membrane-associated signalling proteins can aggregate, they play a critical role in cell signalling[37,204–207].

The existence of rafts as such in real cell plasma membranes remains controversial but it is nonetheless clear that the presence of cholesterol has an effect on the physical properties of membranes, and therefore any forcefield intended to be used for simulations of biological membranes is incomplete without an accurate representation of cholesterol. In this chapter, we describe the parameterisation and testing of cholesterol in the ELBA forcefield and compare the results with established atomistic and CG forcefields.

5.1.1 Cholesterol condensation

Cholesterol does not form stable single component structures in water. Therefore the approaches used to parameterise phospholipids in ELBA (described in chapter 1.1) were not appropriate for parameterisation of cholesterol. However, when included in a bilayer with phospholipids, cholesterol does cause a quantitatively measurable effect on the area of the bilayer - due to the phase change behaviour discussed above - the 'cholesterol condensation effect'[13,208,209].

Because a membrane can be seen as a 2D liquid, a first estimate of how the area per molecule would change upon the addition of cholesterol would be to assume ideal mixing, where the area per molecule is simply a weighted average of the areas of each pure component, and a graph of area-per-molecule against mole fraction of cholesterol would be a simple straight line see figure 5.1. However, as long ago as 1925, Leathes showed that when including cholesterol in a monolayer, instead of ideal mixing, significant non-ideal behaviour is observed[13]. This non-ideal behaviour has been called the condensing effect[208] because the area per molecule is much lower compared with what would expect assuming ideal mixing; the graph of area-per-molecule against mole fraction of cholesterol curves below the ideal mixing line (see figure 5.1).

If it were assumed that the area per phospholipid is constant notwithstanding the presence of cholesterol, the condensing effect is such that when cholesterol is included, the area of the bilayer reduces so much that the area per cholesterol molecule appears to be negative[140,210,211]. The effect has also been shown experimentally in a number of studies[193,212] as well as in atomistic simulations[211,213]. The condensation effect is believed to occur as a result of the ordering of the phospholipid tails; in fact experimental evidence has shown that the presence of cholesterol induces a phase transition in the phospholipids to a meso-phase, the liquid ordered phase, whereby the phospholipid tails have short-range orientational order (as in the gel phase) but long range translational disorder (as seen in the liquid disordered phase)[197,199,214,215].

Different models have been proposed to explain this ordering behaviour. For example, the condensed-complexes model[212,216,217] and the umbrella model[218,219]. The former explains the condensation effect by assuming that cholesterol induces the reversible formation of a stoichiometric cholesterol - lipid complex. In such a complex, the lipid acyl chains are more ordered. The umbrella model argues that the hydrophilic part of cholesterol is small in comparison to that of the phospholipid and therefore the lipid head groups must act as an ‘umbrella’ ie contribute to the screening of the cholesterol molecules from hydrophobic interactions with water. A phospholipid can only create this shield if its tail straightens, ie orders, to make space for cholesterol under the ‘umbrella’ head group. Notwithstanding which model is more appropriate, which is not considered here, both include the ordering of the phospholipid tails as a cause of the condensation effect.

As can be seen from figure 5.1, when results are plotted to compare them directly

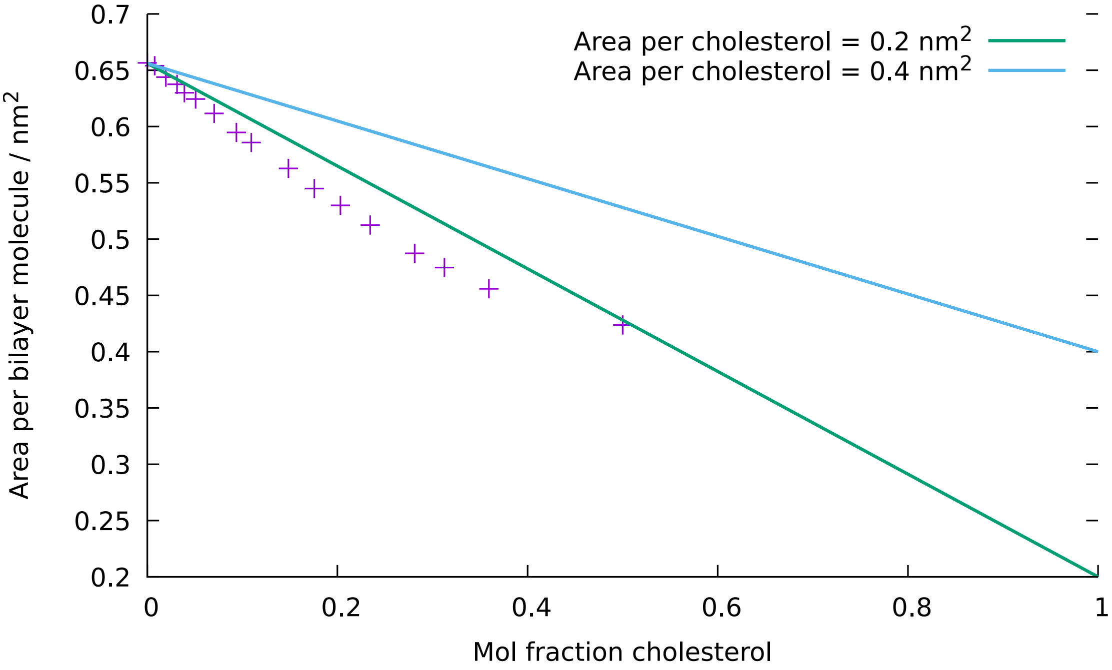


Figure 5.1: An ideal mixing graph for a DSPC/cholesterol bilayer. The straight lines show the areas expected if DSPC and cholesterol behaved as an 'ideal mixture' ie that the area per molecule is a simple weighted average of the area per DSPC and the area per cholesterol molecule. Each straight line represents a different area per cholesterol. The degree of non-ideality of mixing evident from this analysis depends on the area per cholesterol defining the ideal-mixing line. The points are the results of atomistic simulations described in section 5.4.1.

to an ideal mixing curve, the degree to which the system's behaviour deviates from 'ideal mixing' is very dependent on what value is ascribed to the area of pure cholesterol ie what point is chosen as the right hand end of the ideal mixing line. Therefore, when considering the effectiveness of the ELBA parameterisation below, we instead considered a partial molar area approach[210] to quantify the cholesterol condensation:

In this chapter, as proposed by Edholm[210], the mole fraction of cholesterol, x, is defined as:

$$x = \frac{N_{chol}}{N_{DSPC} + N_{Chol}} \tag{5.1}$$

where N_{chol} and N_{DSPC} are the numbers of cholesterol and DSPC molecules respectively.

The area per molecule ($a(x)$) for the system is:

$$a(x) = \frac{A(x)}{N_{molecules}} = \frac{A(x)}{N_{DSPC} + N_{Chol}} \tag{5.2}$$

Then Edholm[210] considers partial specific areas for a bilayer consisting of $i = 1, ..., m$ types of molecules, where A is the total area of the bilayer, X is the set of all mole fractions x_i of the $i = 1, ..., m$ components and the partial derivative is taken with N_j constant for $i \neq j$:

$$a_i(X) = \left(\frac{\partial A(N_i, ..., N_m)}{\partial N_i}\right)_{Nj \neq i} \tag{5.3}$$

and

$$a(X) = \frac{1}{N} A(X) = \frac{1}{N} \sum_i^m a_i(X) N_i = \sum_i^m a_i(X) x_i \tag{5.4}$$

where $N = \sum N_i$

Applying this to 2-component mixtures of DSPC and cholesterol, equation 5.4 can be rearranged:

$$a(x) = (1 - x) a_{DSPC}(x) + x a_{chol}(x) \tag{5.5}$$

which is then rearranged:

$$\frac{a(x)}{1 - x} = \frac{A(x)}{N_{DSPC}} = a_{DSPC}(x) + \frac{x}{1 - x} a_{chol}(x) \tag{5.6}$$

and $a_{chol}(x)$ can be estimated as the gradient of a plot of $\frac{A(x)}{N_{DSPC}}$ against $\frac{x}{1-x}$. The cholesterol condensation effect, ie a negative partial area per cholesterol is therefore obvious and quantifiable as a negative gradient to the graph at low cholesterol concentrations.

5.1.2 Cholesterol tilt

As well as considering the effect of the presence of cholesterol on the properties of the phospholipid molecules present in a binary bilayer, the properties of the cholesterol molecules have been shown to themselves change as cholesterol concentration increases[220].

Previous experimental and simulation work had shown that the preferred position of cholesterol molecules when in high concentrations in saturated phospholipid membranes is only slightly tilted with respect to the bilayer normal, with the hydroxyl group positioned in the lipid headgroup region[221,222]. However, more recent experiments have shown that, in contrast to the results for binary systems with fully saturated phospholipids, in the presence of polyunsaturated fatty acids[223], or mono-unsaturated phosphatidylcholine lipids[224], some cholesterol molecules, even with their hydrophilic hydroxy head group, can be found entirely inside the hydrophobic hydrocarbon core of the bilayer.

Furthermore, not only are the orientation and position of the cholesterol molecules dependent on the types of other lipids present, but it has also been shown in simulation studies that the tendency of the cholesterol molecules to tilt in the membrane is dependant on the concentration of cholesterol present[220]. Khelashvili et al. showed that, for low cholesterol concentrations, there was a significant probability for the cholesterol molecules to 'lie down' even in saturated PC membranes. However, they showed that that tendency gradually diminished with increasing cholesterol concentration, such that at higher concentrations ($\approx$ 30 mol %), the molecules were strongly aligned with the normal to the bilayer plane[220].

In a similar approach to the method described by Khelashvilli at al.[220], for the simulations and analysis of cholesterol tilt described in this chapter, average cholesterol tilt was calculated by considering the cosine of the angle between the bilayer normal and the vector joining each cholesterol's C3 and C17 atoms as shown in figure 5.2. For the CG cholesterol systems, the vector used was that between the beads containing C3 and C17.

5.1.3 Effect of cholesterol on bending rigidity

As well as having an effect on the conformation and alignment of the individual types of lipid in a bilayer, the presence of cholesterol has been reported to affect

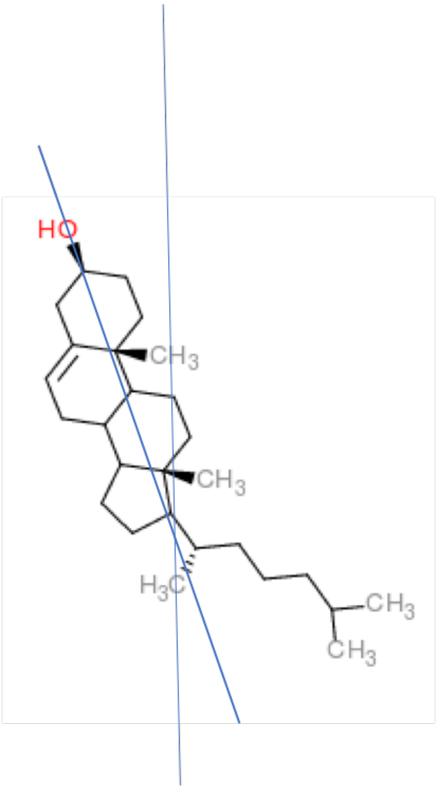


Figure 5.2: The cholesterol tilt is quantified by considering the angle between the bilayer normal and the vector joining C3 and C17.

macroscopic bilayer properties, such as its bending modulus, particularly in bilayers containing only saturated phospholipids[190,191]. Pan et al.'s low angle xray scattering experiments show that, for the saturated phospholipids measured (DMPC), the bending modulus of the bilayer increased rapidly, approximately 4-fold over the range of cholesterol concentrations $0 - 20$mol%, levelling off at a cholesterol concentration of ≈ 30mol%[191].

Various methods for calculating the bending modulus from simulation are described in further detail in this thesis in chapter 4. As further discussed in that chapter, most of the methods considered there require large systems sizes to achieve accurate results. The systems considered in this chapter are too small for many of them and therefore, only 2 of those methods were employed here; SALO and polymer brush. As far as we are aware, this is the first time the SALO method has been used to calculate the bending modulus of cholesterol containing bilayers and the first time it has been employed to measure the bending modulus of CG bilayers.

As described in chapter 4, the SALO method analyses the tilt vectors of the molecules in the bilayer, projects those vectors onto a 2D grid and analyses that projection using FFT. It is therefore necessary to define the tilt vector for the cholesterol molecules as well as the phospholipids. Merely ignoring the cholesterol molecules would lead to non-physically realistic holes in the 2D grid which would cause the FFT analysis to fail. The vector used for DSPC molecules is that between the phosphorous atom and the mid-point of the terminal tail carbons. Therefore, by analogy, the vector considered in this chapter for atomistic cholesterol molecules is that between C3 and C25. For the CG cholesterols, it is the vector between the beads including those carbons.

As described in chapter 4, for the standard polymer brush method, the only measurements required from the trajectory are the bilayer thickness (calculated from the phospholipid P-P headgroup to headgroup distance) and the variation in total bilayer area during the simulation. No extra variation or measurements are therefore required for the bilayers containing cholesterol; the thickness of the bilayer remains the DSPC P-P distance. The 'standard' polymer brush equation is:

$$K_c = K_A(2D_c)^2/24 \tag{5.7}$$

where K_c is the bilayer bending modulus, K_A is the area compressibility modulus and D_c the monolayer thickness.

However, Pan et al.[191] reported a breakdown in the standard polymer brush method of calculation (equation 5.7) and report a proposed an adjustment to it for bilayers containing cholesterol.

They report that Evans (the author of the original polymer brush model) proposes a variation to the polymer brush equation set out in equation 5.7 for those bilayers containing cholesterol on the basis that cholesterol has a rigid ring structure and therefore cannot be compared to a polymer brush.

They instead propose:

$$K_c = K_A(\delta)^2/12 \tag{5.8}$$

where δ is the thickness of the stiff cholesterol ring system. In their experiments, they measure K_c and K_A for SOPC and DOPC and, using equation 5.8 thereby calculate δ. They note that for the greatest concentrations of cholesterol measured, 50 mol%, the values for δ calculated are "close" (their calculated values are between $10.1 - 12.2$ Å) to the actual length of the cholesterol ring system (9 Å). However, they also note that this calculation method nonetheless contains flaws - as the cholesterol concentration decreases, the calculated δ value becomes larger than half the bilayer thickness which is obviously unphysical. As they also note, it is also unclear how the crossover between this theory and the standard polymer brush theory works; clearly standard polymer brush theory should apply when the cholesterol concentration is zero and they suggest that their model applies when cholesterol concentration is 'high'; but they do not propose how to merge the 2 models for concentrations between $0 - 30$ mol% cholesterol.

5.2 Coarse grained modelling of cholesterol

As discussed above, cholesterol's essential function in cells is intimately connected to molecular-scale interactions through modulation of membrane structure and dynamics. Many of these effects occur on micro- to millisecond time scales which are not yet accessible to atomistic simulations. It is therefore essential to develop accurate CG forcefields that can reach these time scales. By far the most commonly used CG model for cholesterol is the MARTINI CG forcefield[50,51]. As discussed in previous chapters, the MARTINI force field is a CG forcefield particularly designed for MD simulations of biochemical systems. It has been parametrized using atomistic simulations to derive the bonded parameters and fitting to experimental partitioning free energies for the non-bonded. Generally MARTINI uses a 4 to 1 mapping scheme and each bead is one of only 4 different interaction types, polar, non polar, apolar and charged. Each of these is types has a number of subtypes. For the ring structures (including that of cholesterol) the mapping is slightly different to reproduce the geometry of a ring. It is based on a 3:1 scheme with special ring beads with lower interaction strengths. MARTINI cholesterol uses 8 beads - 6 for the body and 2 for the tail. When introduced, the only validation of the cholesterol model provided was a simple area per molecule vs mole fraction of cholesterol plot (similar to figure 5.1 above which showed a small deviation from ideal mixing. However, since, MARTINI cholesterol has been used extensively and successfully in biological simulations including the effects of cholesterol on pore formation[225], cholesterol interaction with serotonin receptor sites[226] and conformational plasticity in GPCR dimers[227] and has been shown to reproduce lipid phase segregation in multi-component systems[36].

MARTINI cholesterol is not without its limitations however. It has limited numerical stability[228] and has been reported to incorrectly reproduce the fluidity of L_o domains[229,230] and although there has been some success in qualitatively reproducing the phase diagram for DPPC-cholesterol systems[231] there remains room for improvement. Therefore the development of MARTINI and indeed new CG parameterisations for cholesterol remains a focus. An updated MARTINI cholesterol model was published in 2015[228] which includes virtual sites and extends the model to other sterols. A different approach to reparametrising the MARTINI cholesterol model was adopted by Daily et al[232]. They noted that the original MARTINI model fails to capture the thickening and ordering effects of cholesterol on POPC bilayers and 'systematically reparameterised both the cholesterol and POPC bonded parameters' in order to improve its performance, though even their revised parameterisation over estimated the cholesterol ordering effects at higher concentrations.

Hadley and McCabe[233] took another different approach to parameterising a CG cholesterol. They started from the triclinic structure of crystalline cholesterol and used the unit cell parameters (which differ at room temperature and physiological temperature) and melting temperature as targets for their parameter optimisation and produced a successful CG model to this end - though to date, it appears not to have been tested in mixed lipid systems.

MacDermid et al. also developed an entirely new parameterisation[234] using a similar approach to that taken in the development of MARTINI but but resulting in an 11 bead model. Bond, angle, and torsional terms were derived from atomistic simulations and non-bonded interaction potentials derived from macroscopic measurements (eg density, surface tension, partition coefficients) of small molecules such as cyclic or branched alkanes that are in a liquid phase at room temperature. This model produced areas per lipid in good quantitative agreement with CHARMM36 simulations across the cholesterol concentration range but underestimates bending rigidity by a factor of 2 or more.

All the models described so far in this section have been at a similar level of abstraction. Several even more coarse grained, ie representing a whole cholesterol molecule by a single bead, models have been developed[235–237]. Though obviously very much faster to simulate, generally only qualitative agreement with experiment with regard to non-ideal phase separation and non-ideal mixing is achieved.

It is clear that whilst there are many CG models for cholesterol, there is still room for improvement in reproducing all the complex aspects of cholesterol's impact on the bilayer whilst achieving the computational speed-up available from a CG forcefield.

5.3 Methods

A series of MD simulations were conducted (further described below): fully atomistic, dual-resolution, fully CG (MARTINI) and 2 different fully CG ELBA parameterisations. For each different forcefield, a total of 16 different bilayer compositions were simulated with the ratio of DSPC:Cholesterol ranging from 1:1 to 512:0. The compositions were not equally spaced throughout the composition range; a greater number of different low concentrations were simulated because the condensation effect is believed to take effect at low concentrations of cholesterol. No concentrations above 50 mol% cholesterol were simulated as it has been found previously, both experimentally and computationally, that above a concentration of ≈ 60 mol% cholesterol, cholesterol is not fully soluble in phospholipid bilayers[238]. Where cholesterol was included in the bilayer, it was arranged randomly in the xy plane, with the hydroxyl group towards the water, and equal numbers of cholesterol molecules were initially included in each leaflet.

For the atomistic bilayers referred to in this chapter, a series of atomistic (CHARMM36) bilayers, each consisting of a total of 512 lipids (DSPC and cholesterol combined) and 16810 TIP3P waters were independently constructed using the CHARMM Membrane Builder[158]. Details of the CHARMM36 forcefield including cut-offs, long distance corrections etc are described in chapter 2. A series of energy minimisation and equilibration NVT and NPT simulations with increasing timesteps were performed on each in accordance with the protocol suggested by the CHARMM Membrane Builder and described in Appendix A, after which, a 50 ns production run was conducted for each system. All the production simulations were conducted in the NPT ensemble using GROMACS 4.6.5[162–164], at a temperature of 341 K (lipids and water thermostatted separately with a Nosé-Hoover thermostat[239]), and a pressure of 1 atmosphere (semi-anisotropic pressure coupling, Parinello-Rahman barostat[240]) with a timestep of 2 fs. All results reported in this chapter are derived from the 50 ns production runs.

To create the dual-resolution bilayers; ie those where DSPC and water were represented using the ELBA1.1 forcefield (described in detail in chapter 2) and atomistic cholesterol molecules represented using the CHARMM36 forcefield, the PyCGTool software[241] was applied to the atomistic trajectories described above. PyCGTool, described further below, was used in mapping mode only in this case ie purely to calculate the required positions of the associated CG beads.

These simulation were conducted in LAMMPS (Jan 2015)[242]. Once again, these systems were energy minimised and equilibrated in a series of short (10000 step) NVT simulations, with increasing timesteps from 0.15 fs to 2 fs, followed by NPT simulations using a timestep of 2 fs, followed by 50 ns "production" runs. The production runs were conducted using the rRESPA dual-timestep method (described in chapter 2 with a timestep of 2 fs for the 'inner' loop and 6 fs for the 'outer'. CG-CG non-bonded interactions are calculated in the outer loop and atomistic-CG and atomistic-atomistic in the inner loop. Water, DSPC and cholesterol were all thermostatted separately at 341 K using a Langevin thermostat. Semi-anisotropic pressure coupling was applied using a Berendsen barostat at 1 atmosphere.

Similar systems of DSPC/cholesterol were set up using a CG MARTINI representation using the MARTINI membrane builder[180]. They were energy minimised and equilibrated in accordance with the protocol suggested by the MARTINI membrane builder software, set out in Appendix B, and thereafter 50 ns production runs conducted in the NPT ensemble using GROMACS 5.0.4[162–164] at a temperature of 341 K with lipids and water thermostatted separately using v-rescale thermostat, and a pressure of 1 atmosphere (semi-anisotropic pressure coupling, Parinello-Rahman barostat[240]) with a timestep of 20 fs. Because the phospholipids simulated using the MARTINI forcefield have a 15 - 20 K lower T_m than atomistically simulated phospholipids[88] and the condensation effect is at its peak close to T_m[230], a further set of MARTINI simulations were conducted at 326K, ie 15 K lower than the atomistic simulations we conducted.

Two separate sets of ELBA DSPC/cholesterol simulations using different proposed parameterisations of cholesterol were also conducted. For the first, the cholesterol molecule, with the exception of the OH head group) was divided into identical spherical beads each consisting of 3 carbon atoms as shown in figure 5.3 and

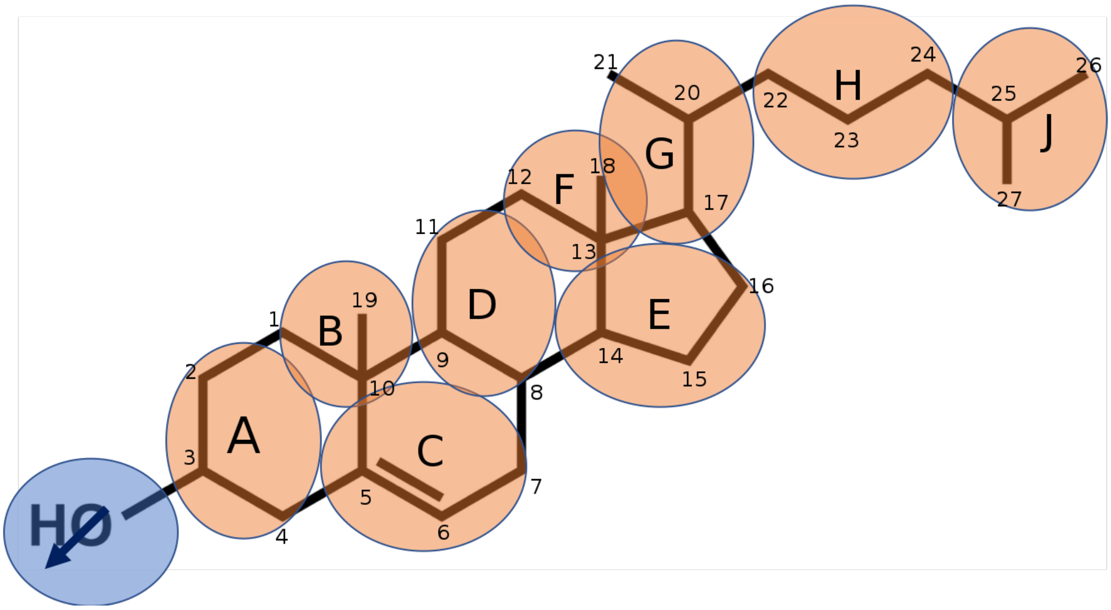


Figure 5.3: The mapping of ELBA cholesterol 1 (orange beads labelled with letters) compared with atomistic cholesterol (labelled with numbers). All ELBA beads are spherical and represent 3 carbon atoms.

described further in table 5.1. For the second, additional beads were included for the methyl groups protruding from the cholesterol ring system. This second parametrisation was attempted as there have been suggestions in the literature that the cholesterol condensation effect requires a contrast between the "rough" and "smooth" faces of the molecule[243–245]. The second parametrisation is further described in figure 5.4 and table 5.2.

PyCGTool[241] was used on the dual-resolution trajectories described above to calculate the intra-molecular constants required for the ELBA cholesterols. Together with the appropriate atomistic trajectories, the mappings described in figures 5.3 and 5.4 and tables 5.1 and 5.2 were input into PyCGTool. A bond input files was also created listing the pairs of beads connected by bonds. An angle input file was not created, instead relying on PyCGTool's default behaviour of evaluating all valid triplets of beads. Using these input files, PyCGTool works by creating a 'pseudo-CG' representation for each frame of the trajectory based on the centre of mass of the constituent atoms. Mean values and standard deviations are calculated for all of the internal co-ordinates calculated over all of the frames of the trajectory. Rather than requiring the generation of constraints which are not an existing part of the ELBA forcefield, PyCGTool was required to produce information for all the possible dihedral combinations. Dihedral potentials with

ELBA bead	Cholesterol carbon atoms
A	2 3 4
B	1 10 19
C	5 6 7
D	8 9 11
E	14 15 16
F	12 13 18
G	17 20 21
H	22 23 24
J	25 26 27

Table 5.1: Allocation of atomistic carbon atoms into each ELBA CG bead for the first cholesterol parameterisation.

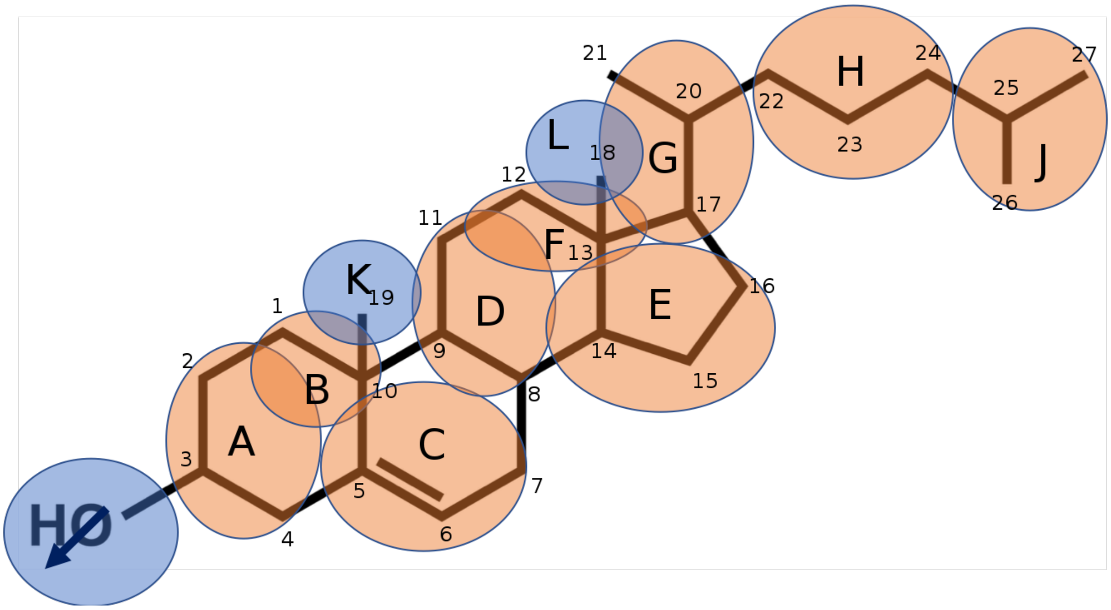


Figure 5.4: The alternative proposed mapping of ELBA cholesterol (orange and blue beads labelled with letters) compared with atomistic cholesterol (labelled with numbers). All ELBA beads are spherical.

ELBA bead	Cholesterol carbon atoms
A	2 3 4
B	1 10
C	5 6 7
D	8 9 11
E	14 15 16
F	12 13
G	17 20 21
H	22 23 24
J	25 26 27
K	19
L	18

Table 5.2: Allocation of atomistic carbon atoms are into each ELBA CG bead for the second cholesterol parameterisation.

a force constants less than 10 kJ.mol^{-1} were discarded from the resulting cholesterol parameterisation which is listed in full in Appendix C and Appendix D for the ELBA 1 cholesterol and the ELBA 2 cholesterol with extra methyl beads respectively.

For the inter-molecular Lennard-Jones parameters, ϵ and σ, the values for the new multi-carbon beads (coloured orange in figures 5.3 and 5.4), were selected to reproduce the radial distribution functions (rdf) obtained from atomistic simulations of liquid propane (see figure 5.5). Increasing the sigma value shifts the rdf curve to the right, whereas increasing the epsilon value increases the height of the first peak and decreases the well-depth before the second. The atomistic liquid propane simulations were conducted in GROMACS 5.0.4 using the OPLS-aa forcefield[159] at a temperature of 200K (Nosé-Hoover thermostat), pressure of 2MPa (Parinello-Rahman barostat with isotropic coupling) using PME for long-range electrostatics and a cut-off of 10 Å with potential shift function. The timestep and length of the production simulation from which bulk properties were calculated were 1 fs and 20 ns respectively. The temperature and pressure were selected to ensure that the propane was in the liquid phase[246].

σ	ϵ	density
4.224	0.3613	0.6305
5.034	0.4659	0.176
4.629	0.4136	0.517
4.629	0.34	0.468
4.629	0.361	0.4828
4.629	0.38	0.497
4.224	0.38	0.651
4.224	0.3613	0.6305
4.224	0.37	0.6400
4.224	0.38	0.6506
4.224	0.375	0.6452

Table 5.3: This table shows the sets of σ and ϵ values tested to replicate the density calculated from OPLS-aa simulations.

In addition, the propane density from the atomistic simulations was calculated (643.07 kg.m^{-3}) and the values of σ and ϵ adjusted to replicate as closely as possible. The experimentally measured value for propane at this temperature and pressure is 617.4 kgm^{-3} - a difference of 2%[246]. Table 5.3 shows the effects of different LJ parameters on the density calculated from separate 30 ns CG simulations of 'ELBA propane' - a LJ fluid - with these parameters.

The σ = 4.224 Å and ϵ = 0.375 kcal/mol finally selected were chosen to match the atomistic simulation density value as far as possible, whilst also closely reproducing the rdf curve.

The standard Lorentz-Bertholet rules were applied to calculate interactions with pre-existing ELBA bead types. An h-factor=1.5 for the cholesterol OH bead when calculating the ϵ values for interactions with the ELBA phosphate, glycerol and ester beads was included, in the same way as an h-factor is included for interactions with ELBA water beads (see chapter 2).

The OH head group was represented using an ELBA water bead type with the dipole magnitude, direction and force constant calculated from analysis of the trajectories containing atomistic cholesterol.

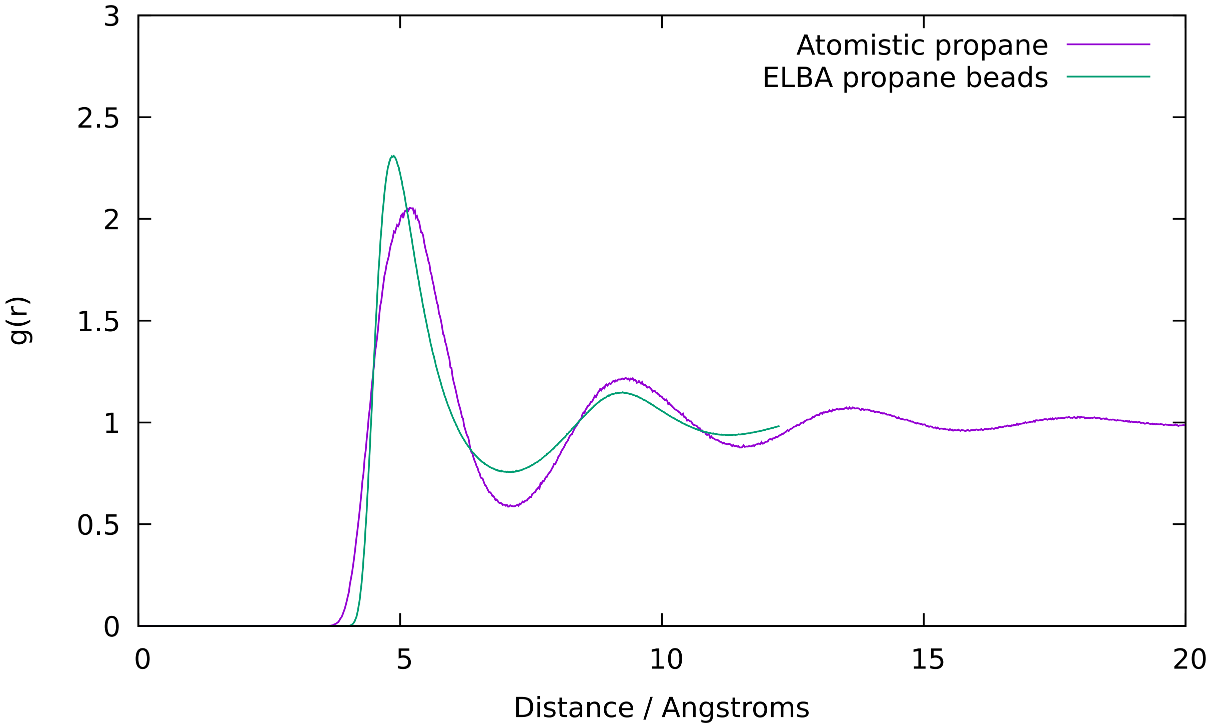


Figure 5.5: Radial distribution plots calculated for the central carbon atoms in the simulation of OPLS-aa propane (labelled ”atomistic”, compared with radial distribution plots for ELBA beads with $\sigma = 4.224$ Åand $\epsilon = 0.375$ kcal/mol.

The ELBA simulation were conducted in LAMMPS (Jan 2015)[242]. Once again, these systems were energy minimised and equilibrated in a series of short (10000 step) NVT simulations, with increasing timesteps from 0.15 fs to 2 fs, followed by NPT simulations using a timestep of 2 fs, followed by 50 ns “production” runs with a timestep of 6 fs. Water, DSPC and cholesterol were all thermostatted separately at 341 K using a Langevin thermostat. For these simulations, semi-isotropic pressure coupling was applied using a Berendsen barostat. Results are not reported for either ELBA cholesterol parameterisation above 31 mol% cholesterol as the simulations were not stable above this concentration for either parameterisation.

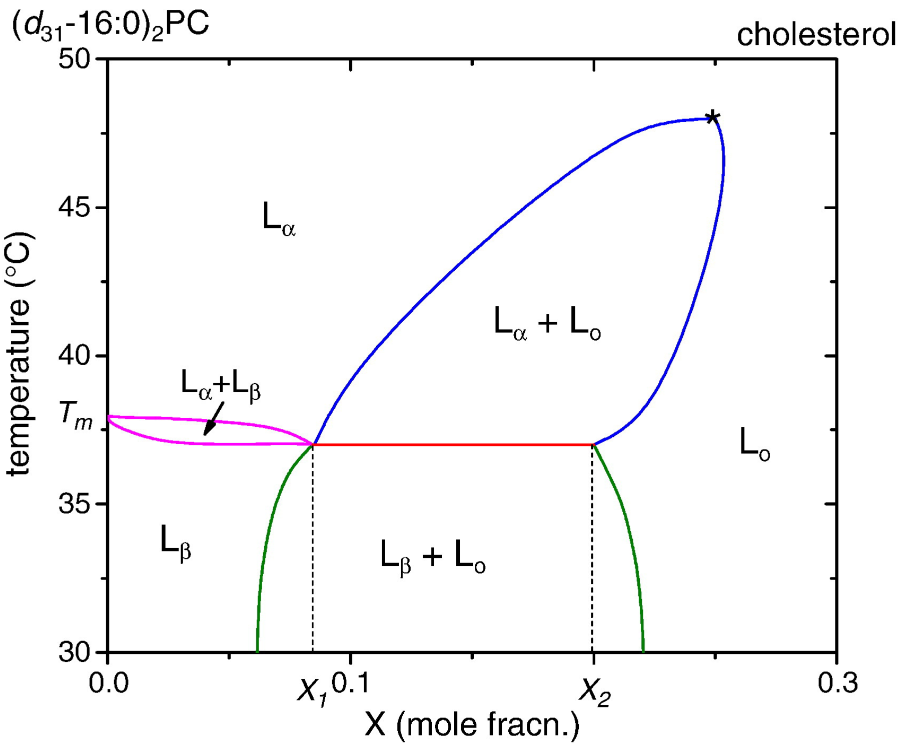


Figure 5.6: The model phase diagram presented by Marsh[199]. Reproduced with permission. The numeric values relate to a DPPC - cholesterol mixture.

5.4 Results

5.4.1 Cholesterol Condensation

Marsh reviewed the experimental data with regard to cholesterol induced phase transitions in saturated PC bilayers[199] and gives an model phase diagram for cholesterol - PC mixtures (reproduced with permission in figure 5.6) which predicts a co-existence region for L_o and L_α (or L_d) phases. He also shows the experimental data for DSPC - cholesterol mixtures (reproduced with permission in figure 5.7). This experimental data indicates that a L_o - L_α phase co-existence region might be expected in the region $x \approx 0.1$ to $x \approx 0.3$ corresponding to 51 - 144 cholesterols in the 512 lipid systems simulated in this chapter.

Figures 5.8, 5.9 and 5.11 show the partial molar area plots from the simulations we conducted using the atomistic, dual-resolution and CG forcefields respectively. For the first 2, where the cholesterol molecules are represented atomistically, the graph shows a negative gradient at low cholesterol concentrations ($x <$

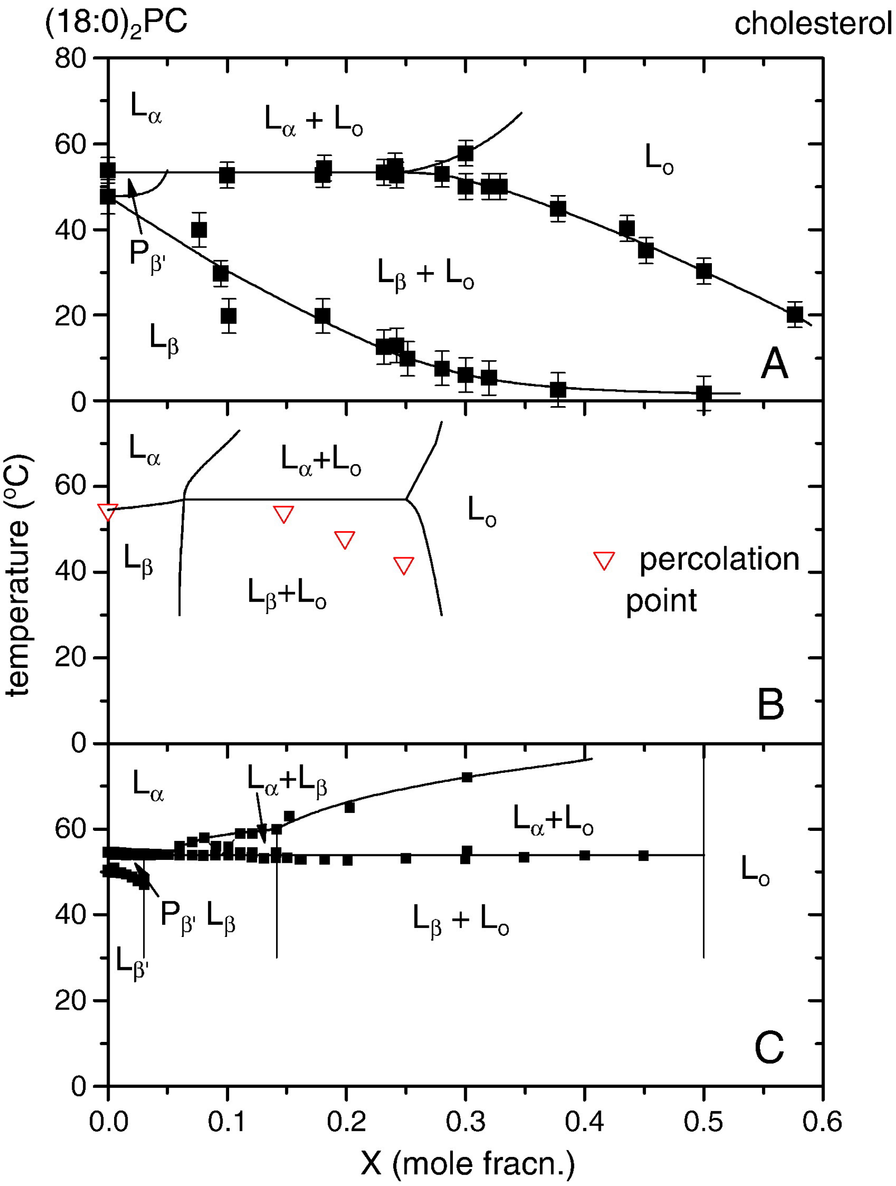


Figure 5.7: The experimentally measured phase diagrams measured for DSPC - cholesterol mixtures presented by Marsh[199]. Reproduced with permission. The data in A are obtained from ^{13}C-NMR, B from EPR spectra of spin-labelled PC and C solely from differential scanning calorimetry results.

0.15) which represents the cholesterol condensation effect. At higher cholesterol condensations, the gradient is positive. This is reflects the situation where all the phospholipids have ordered as much as possible and therefore addition of further cholesterol molecules does not cause any additional tail ordering, corresponding to the L_o region in the phase diagrams in figures 5.6 and 5.7; the difference in bilayer area upon replacement of a DSPC molecule by a cholesterol molecule simply reflects the difference in area-per-lipid between them.

As can be seen from the partial molar area equation 5.6 above, the gradient of the partial molar area graphs in the L_o region represents a value for area per cholesterol. The gradient of the straight line (r^2=0.99) fitted to the portion of the partial molar area graph (figure 5.8) above $x = 0.35$ is 28.4 Å^2, thereby giving a value for area per cholesterol calculated from the atomistic simulations as 28.4 Å^2. Similarly the gradient of the straight line fitted ($r^2 = 0.99$) to the same concentration region of the dual resolution partial molar area graph (figure 5.9) is 32.3 Å^2.

As discussed above, cholesterol does not form stable bilayers alone and therefore an experimental value for area-per-cholesterol in a bilayer is not readily available, however, the values obtained from the atomistic and dual-resolution studies here are comparable to the area for cholesterol calculated from monolayer spreading experiments of 35 Å^2 [247], those calculated from WAXS experiments of 31 Å^2 from binary cholesterol/diC22:1PC systems and 34 Å^2 from binary cholesterol/SOPC systems[191], and those obtained for cholesterol monohydrate crystals of 36.2 Å^2 [248].

As can be seen in figure 5.8, the points obtained from simulation do not fit simply fit 2 straight lines. There is an interim portion in the region $x > 0.14$ to $x < 0.28$, corresponding to $\frac{x}{(1-x)} = 0.16$ to $\frac{x}{(1-x)} = 0.39$. This may accord with the range of concentrations where the L_o and L_c phases co-exist - not all of the DSPC molecules are fully ordered shown in figures 5.6 and 5.7.

To test the hypothesis that this portion of the partial molar area plot represents a 2-phase coexistence region, we calculated the tail order parameters for each molecule in the final frame of each production run and plotted the results as histograms in figure 5.10, one histogram per simulation. If there were a 2 phase region between $x = 0.14$ and $x = 0.28$ (corresponding to $\frac{x}{(1-x)} = 0.16$ to $\frac{x}{(1-x)} = 0.39$), one might expect to see a bimodal distribution of the tail ordering in the histograms in figure 5.10(h) to 5.10(l). However, the data from the atomistic simulations is not definitive; possibly due to small system size.

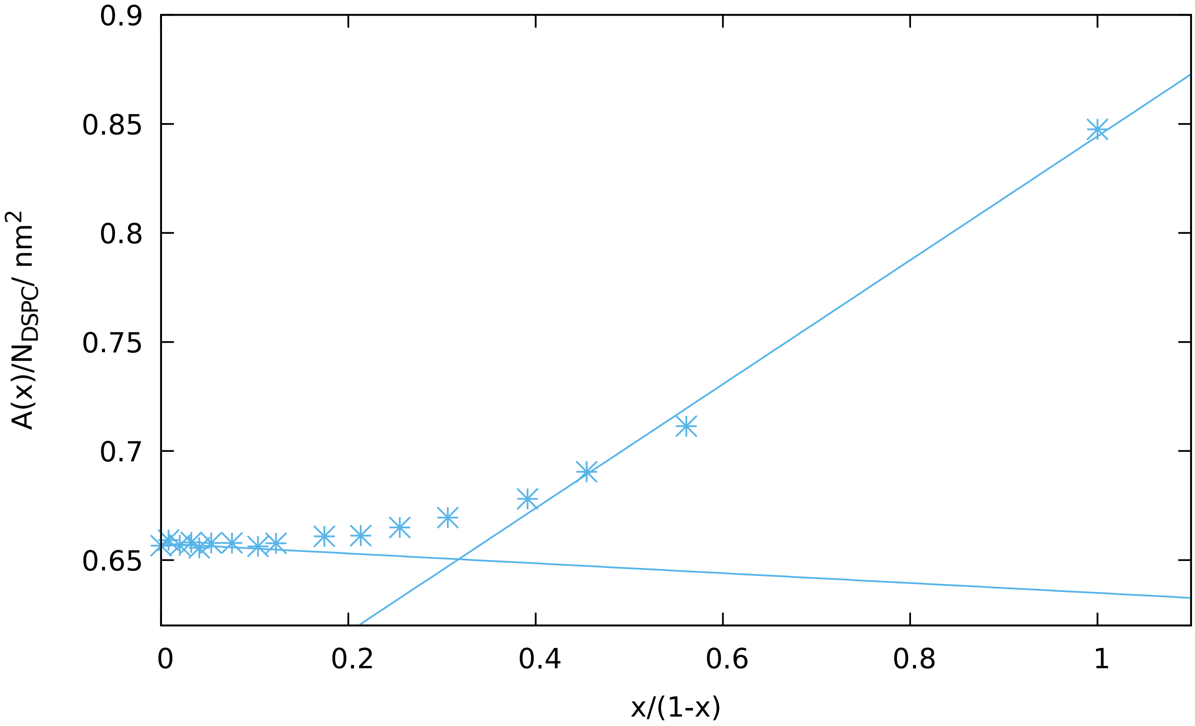


Figure 5.8: Partial molar area plot for atomistic (CHARMM36) simulations. r^2 =0.96 for the straight line ($y = 0.284x + 0.56$) fitted for the higher cholesterol concentrations, corresponding to $\frac{x}{(1-x)} > 0.39$. $r^2 = 0.99$ for the straight line ($y = -0.02 + 0.657$) fitted for the lower cholesterol concentrations, corresponding to $\frac{x}{(1-x)} < 0.14$.

Figure 5.11 shows that none of the CG simulations exhibited cholesterol condensation if defined by a negative gradient at the lower cholesterol concentration of the partial molar area plot; there is a reduced, albeit still positive, gradient at low cholesterol concentrations for both of the MARTINI simulations, thereby indicating that the DSPC and cholesterol mixing is non-ideal, however there is no negative gradient at low cholesterol concentrations at either temperature and therefore no cholesterol condensation effect is observed. Taking the gradient of the high cholesterol concentration portion (defined as $x > 0.35$ as before) of the graph for the MARTINI simulations gives a calculated value for the area per cholesterol of 19.3 Å^2 for the MARTINI simulations conducted at 341 K and 20.2 Å^2 for those conducted at 326 K. Both are significantly different from both the experimentally calculated values quoted above and the values calculated from our atomistic and dual-resolution simulations and also significantly different from the area calculated by Waheed et al.[230] from their DPPC and cholesterol simulations

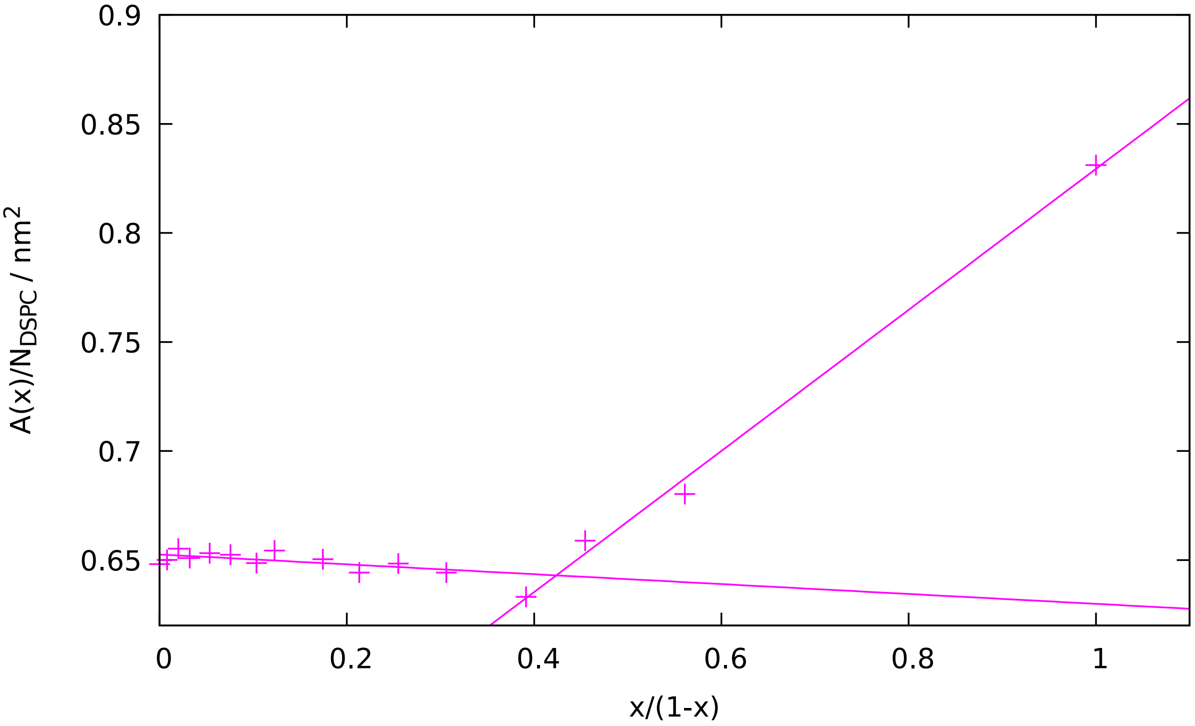
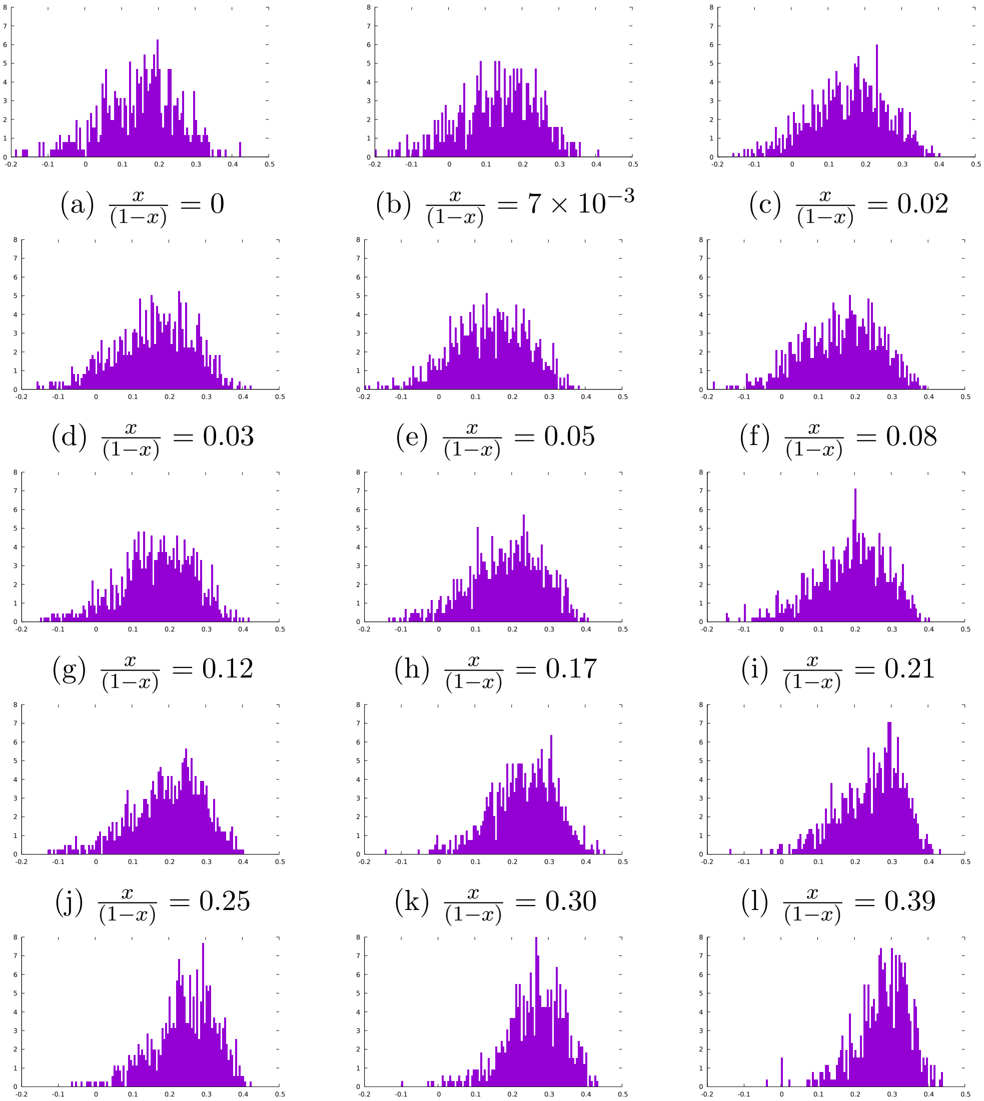


Figure 5.9: Partial molar area plot for dual-resolution (CHARMM36 and ELBA) simulations. r^2 =0.99 for the straight line ($y = 0.323x + 0.50$) fitted for the higher cholesterol concentrations, corresponding to $\frac{x}{(1-x)} > 0.39$. $r^2 = 0.76$ for the straight line ($y = -0.02x + 0.650$) fitted for the lower cholesterol concentrations, corresponding to $\frac{x}{(1-x)} < 0.14$.

using MARTINI.

It is of particular note that the lower temperature MARTINI simulations do not exhibit cholesterol condensation. Condensation, as defined by a negative portion of the partial molar area plot, has been reported for simulations of DPPC and cholesterol systems using the MARTINI forcefield at temperatures close to the DPPC T$_m$ in MARTINI ie lower than either the atomistic or experimental T$_m$ for DPPC[230], however, we have not been able to reproduce these results either with DPPC/cholesterol systems (not shown here) or (as seen in figure 5.11) with DSPC/cholesterol systems. It should be noted that Waheed et al's[230] set of simulations were equally spaced across the concentration range and therefore they simulated fewer cholesterol concentrations in the low cholesterol concentration portion of the graph where the crucial negative gradients are observed, see figure 5.12. For the curves plotted for 287 K and 300 K, the negative gradient and therefore cholesterol condensation claimed, relies on only one data point at each



(a) $\frac{x}{(1-x)} = 0$ (b) $\frac{x}{(1-x)} = 7 \times 10^{-3}$ (c) $\frac{x}{(1-x)} = 0.02$

(d) $\frac{x}{(1-x)} = 0.03$ (e) $\frac{x}{(1-x)} = 0.05$ (f) $\frac{x}{(1-x)} = 0.08$

(g) $\frac{x}{(1-x)} = 0.12$ (h) $\frac{x}{(1-x)} = 0.17$ (i) $\frac{x}{(1-x)} = 0.21$

(j) $\frac{x}{(1-x)} = 0.25$ (k) $\frac{x}{(1-x)} = 0.30$ (l) $\frac{x}{(1-x)} = 0.39$

(m) $\frac{x}{(1-x)} = 0.45$ (n) $\frac{x}{(1-x)} = 0.57$ (o) $\frac{x}{(1-x)} = 1$

Figure 5.10: Normalised histograms of the tail ordering of the DSPC lipids in the final frame of each atomistic CHARMM36 simulation.

temperature and for the 295 K curve, on 2 data points.

Not only do neither of the sets of ELBA simulations plotted in figure 5.11 exhibit any cholesterol condensation, they do not even show any non-ideal mixing behaviour for cholesterol and DSPC. For both the partial area graphs fit a straight line well over the whole concentration range. Using the gradient of this straight line fit to calculate the area per cholesterol as before, gives a value for the ELBA 1 forcefield of 16.6 Å^2 and for ELBA 2 of 13.1 Å^2. Both are even more significantly different that MARTINI from both the experimentally calculated values quoted above and the values calculated from our atomistic and dual-resolution simulations.

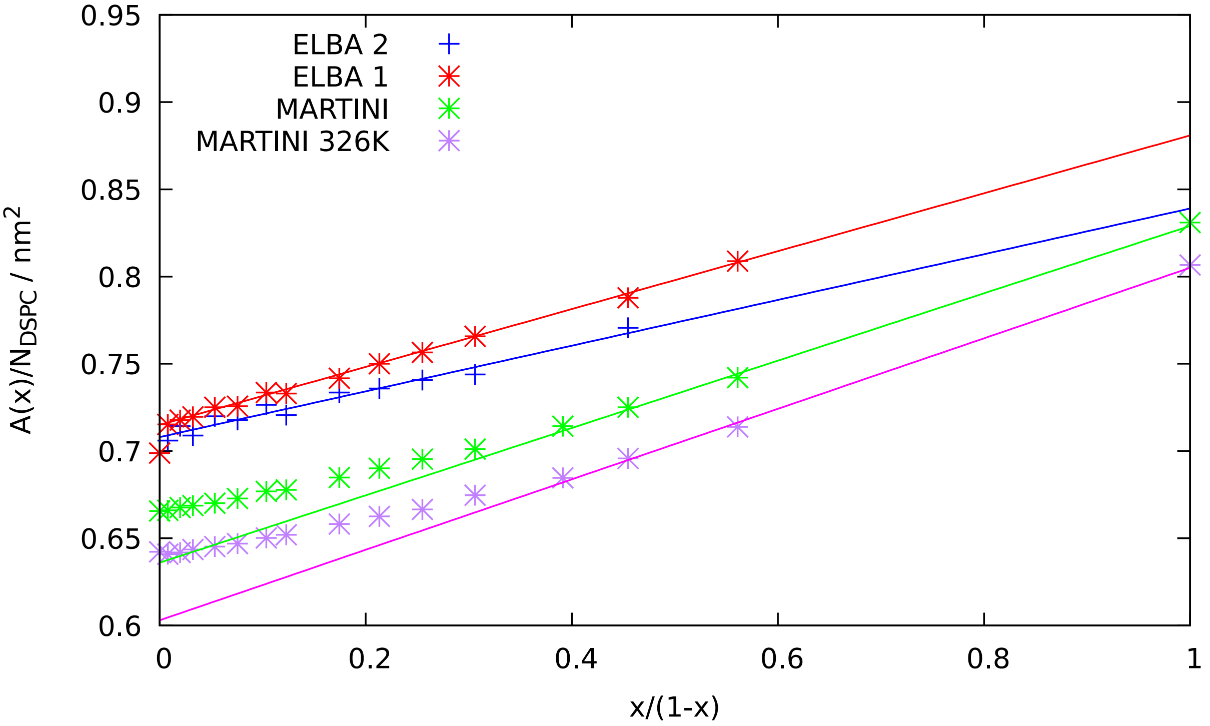


Figure 5.11: Partial molar area plot for CG simulations; MARTINI and both ELBA cholesterol parameterisations. Straight line fits are plotted for both MARTINI sets of simulations over the range $x > 0.3$; for MARTINI at 341 K, $y = 0.193x + 0.636$, r^2=0.99; MARTINI at 326 K, $y = 0.202x + 0.603$, r^2=0.99. For both ELBA simulations, straight line fits are plotted over the whole concentration range; for ELBA 1 $y = 0.1659x + 0.698$, r^2=0.99 and for ELBA 2 $y = 0.131x + 0.7$ and r^2=0.96.

5.4.2 Tail ordering

As discussed above, the origin of the cholesterol condensation effect is believed to be due to the increased ordering of the phospholipid tails. Analysis was therefore conducted as to the ordering effect on the DSPC tails of increased cholesterol concentration. With the exception of the atomistic simulations, all the DSPC molecules in the simulations in this chapter were represented in CG forcefields and therefore carbon-carbon segmental order parameters (S^k_{mol}) were calculated and averaged down both the 5 tail bead chains:

$$S^k_{mol} = \frac{1}{2} \left\langle \, 3 \, cos^2\eta - 1 \, \right\rangle \tag{5.9}$$

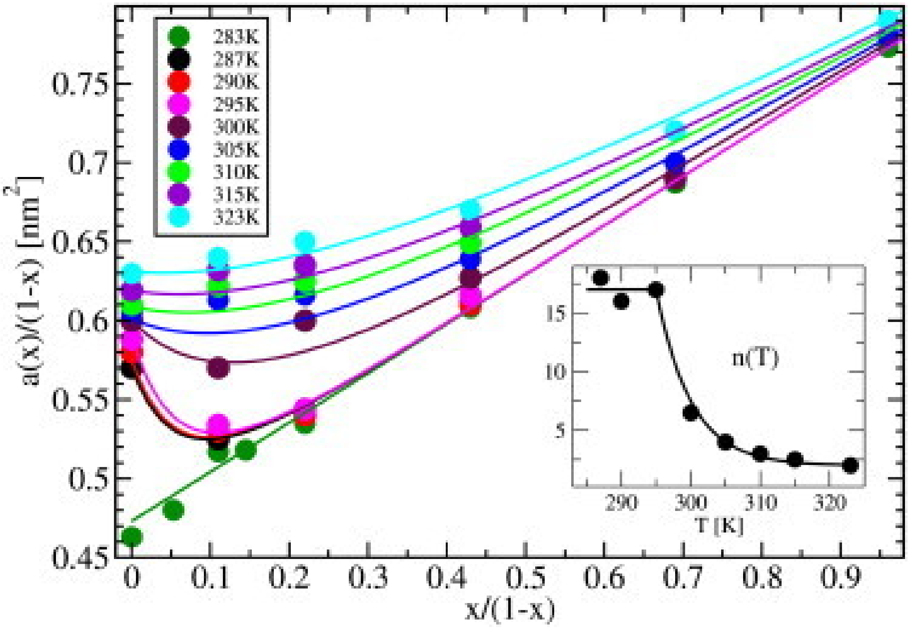


Figure 5.12: Waheed et al.'s[230] results showing cholesterol condensation of DPPC bilayers modelled with the MARTINI forcefield. Reproduced with permission.

Where η is the instantaneous angle between the bilayer normal and the bond vector between the k^{th} and the $(k-1)^{th}$ tail bead.

If the same analysis were to be conducted for the atomistic simulations using each carbon atom along the tail, the results would not be directly comparable to the results from the CG analysis due in part to the 'odd-even' effect[249]. Therefore, in order to directly compare the atomistic and CG tail ordering, for the atomistic simulations S^k_{mol} was calculated considering η to be the instantaneous angle between the bilayer normal vector and the vector between the carbon tail atoms: C2-C6, C6-C10, C14-C18; thereby dividing each of the atomistic tails into a number of segments similar to the number of beads present in the CG models. We tested different groupings to divide the atomistic tails into 3 segments and the results presented here were not affected by the specific groupings chosen.

Figure 5.13 shows the ordering effect of cholesterol on the DSPC tails across all the different forcefields. The ordering effect as cholesterol concentration increases is most pronounced in the dual-resolution simulations. This accords with the more pronounced cholesterol condensation effect seen in the partial molar area graph (figure 5.9) which shows a very slightly more negative partial molar area for cholesterol at lower cholesterol concentrations. However the dual resolution simulations results are also much 'noisier' than the other simulations. Further

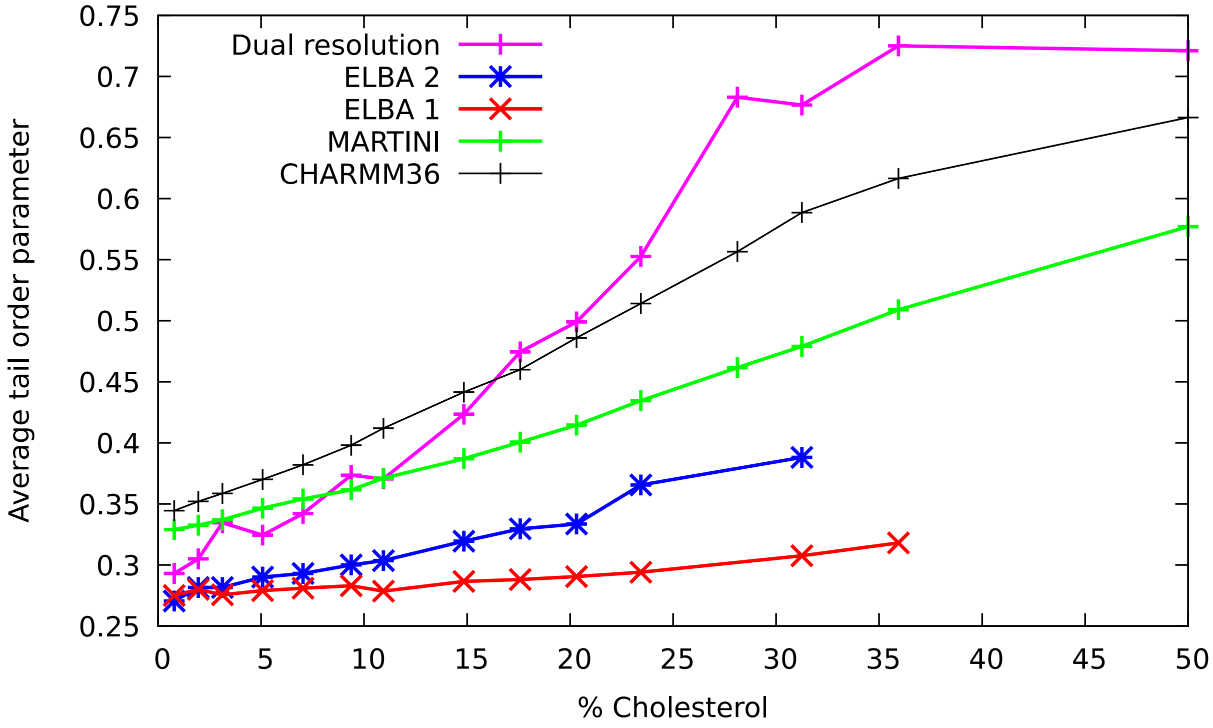


Figure 5.13: Average tail order over the production simulation trajectory. Averages taken over 5 tail beads/segments, over both tails.

work should be done either repeating these simulations from a different bilayer configuration or extending these to allow an estimate of the errors in these results. The ordering effect is also seen in the atomistic simulations where cholesterol condensation was also seen most significantly in the partial molar area analysis. To a much lesser extent, it also appears in the MARTINI simulations which accords with the non-ideal mixing (though not cholesterol condensation) seen in figure 5.11.

5.4.3 Cholesterol tilt

Figure 5.14 shows the results of the cholesterol tilt analysis performed on each of the simulation sets. With the exception of ELBA parametrisation 1, all the forcefields show that the average absolute value of $cos\ \theta$ tends to be greater as cholesterol concentration increases ie there is a tendency for the cholesterol molecules to be more aligned with the bilayer normal as the concentration of cholesterol is increased. However, for the atomistic and dual resolution simulations, ie in both cases where the cholesterol molecules are represented atomistically, not only is the increase in $cos\ \theta$ greater over the concentration range but at high concentrations the average cholesterol is much closer to parallel to the bilayer normal ($cos\ \theta \approx 1$) than in any of the CG representations. This suggests that the average cholesterol tilt is dependent on a property of the cholesterol molecule which is captured in an atomistic representation but not in any of the CG representations.

The shaded areas on figure 5.14 represent the width of the variation of the absolute values of $cos\ \theta$ over the simulations. They represent a single standard deviation plotted on either side of the average for that simulation. This plot shows that for all the forcefields, at low cholesterol concentrations, the cholesterol molecules adopt a wide variety of tilt angles over the course of the simulations whereas at higher concentrations, the cholesterol molecules are much more likely to be found in a smaller range of tilt angles. The spread of tilts is particularly small for the atomistic simulations, suggesting that the CG representations of the DSPC do not enforce the correct tilt in the cholesterol molecules as much as the atomistic representation does; though the effect of coarse-graining the cholesterol molecules themselves is much greater.

5.4.4 Bilayer Bending modulus

Figure 5.15 and figure 5.16 show the results of the polymer brush and SALO bending modulus calculations respectively for all the forcefields. Neither of the ELBA parametrisations, using either calculation method, show an increase in bending modulus as cholesterol concentration increases. The atomistic, dual-resolution and MARTINI all do. Though the shape of the curves is different for each calculation method, the magnitude of the bending modulus increase they exhibit both both is approximately 5 fold in each case which accords well with the magnitude seen experimentally by Pan et al. (albeit for a shorter chain phospholipid)[191]. However Pan et al. observed the increase in bending modulus at much lower cholesterol concentrations ($0 - 20$ mol%). In addition, the increase in bending modulus observed in our simulations occurs in a portion of the concentration range with relatively few data points (4). Therefore, whilst the results are encouraging that the polymer brush and SALO methods may be appropriate for exploring bending moduli in cholesterol containing bilayers, it is suggested that further work is needed; for example (1) reproducing Pan et al's experimental curve using the same lipid (DMPC) in atomistic simulations, should be demonstrated before

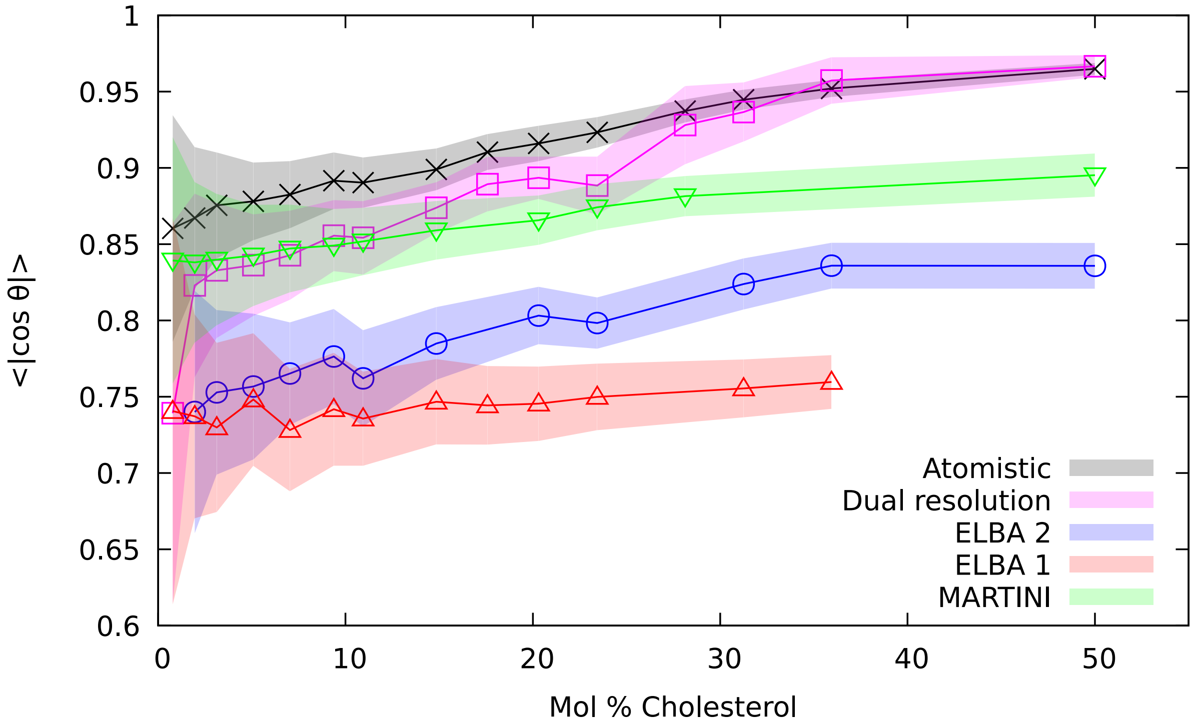


Figure 5.14: Average cholesterol tilt over each trajectory over all cholesterol molecules. The lines and points show the average $\langle|cos\theta|\rangle$ with matching coloured shaded region being one single standard deviation either side of that average.

relying on either calculation method; and (2) conducting more atomistic and CG simulations of the types considered in this chapter in the concentration range $20-50$mol% cholesterol.

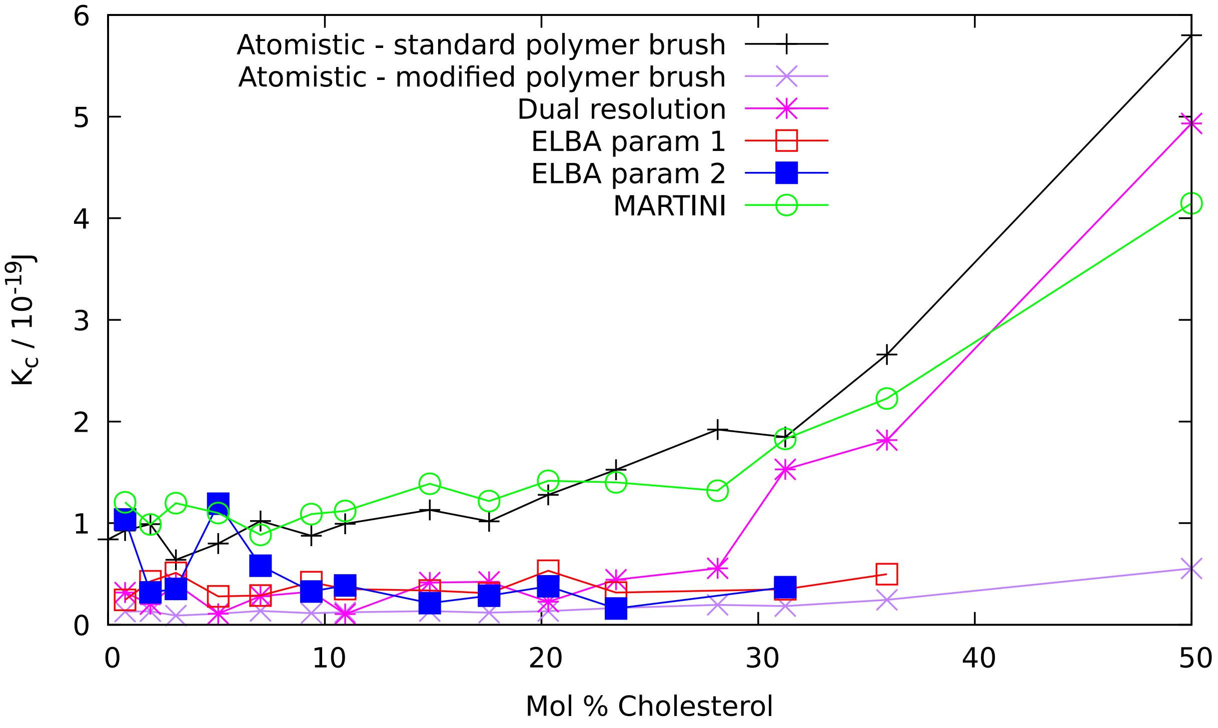


Figure 5.15: Bilayer bending calculated using the polymer brush methods.

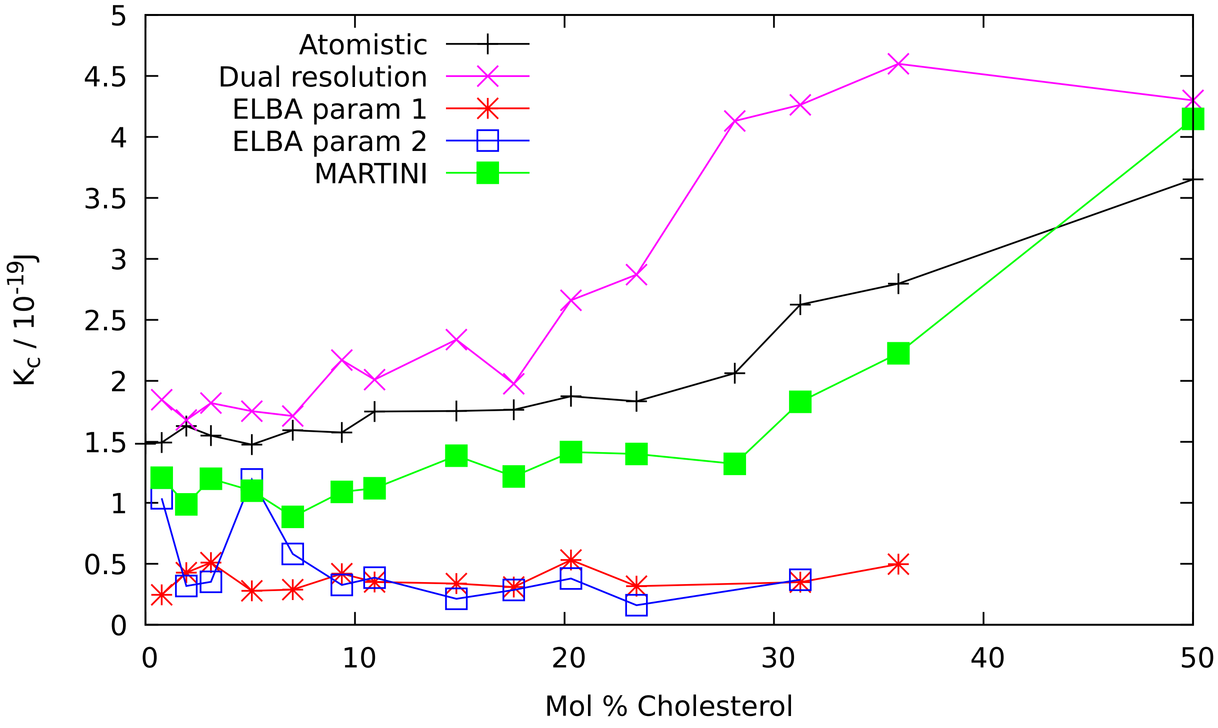


Figure 5.16: Bilayer bending modulus calculated using the SALO method.

5.5 Conclusions and future work

Both the dual-resolution and atomistic sets of simulations performed gave an apparent negative partial molar area for cholesterol at low concentrations and therefore showed cholesterol condensation. They also showed similar levels of cholesterol-induced DSPC tail ordering. By contrast, none of the CG simulation data showed cholesterol condensation when analysed in the same way. Only the simulations using the MARTINI CG forcefield displayed any indication of non-ideal mixing and they also showed the most tail ordering of the CG forcefields - though much less than the simulations involving atomistically modelled cholesterol. Taken together, these results accord with previous reported results that the cholesterol condensation effect is due to the ordering (and phase transition) of the phospholipid component in such binary systems.

In addition, the results from the cholesterol tilt analyses follow similar pattern. For simulations where the cholesterol is modelled atomistically, even where the the phospholipids are not, the average tilt angle of the cholesterol molecules at higher cholesterol concentrations reaches a similar value, reflecting the cholesterol's high tendency to adopt conformations parallel to the bilayer normal. At lower cholesterol concentrations though, the dual-resolution simulations appear to overestimate the cholesterol tilt as compared to fully atomistic simulations.

None of the CG cholesterol representations were able to match such tilt angles, though both the MARTINI and ELBA 2 representations at least displayed the correct trend - a decrease in cholesterol tilt at increasing concentration. This suggests that contrast between the 'rough' and 'smooth' faces of cholesterol improves the cholesterol tilt behaviour in the ELBA forcefield.

Khelashvilli et al.[220] propose that a particular double bond's presence (or absence) in the cholesterol ring system accounts for its tendency to line up with the bilayer normal in binary bilayers. They describe 7DHC, the immediate precursor to cholesterol in *in vivo* cholesterol synthesis, which only differs from cholesterol in having an additional double bond in the cholesterol ring system between C7 and C8 (marked on figure 5.17). They report that 7DHC does not exhibit the same concentration preference for orientations parallel to the bilayer normal. They do not report the effect of 7DHC on either cholesterol condensation or phospholipid tail ordering. Such a small difference would be difficult to capture in a CG forcefield and if this double bond is indeed required for the full cholesterol

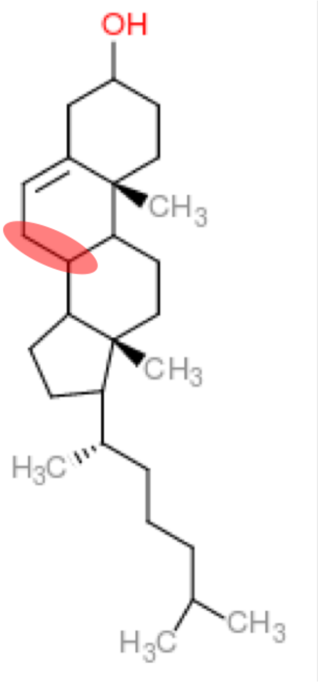


Figure 5.17: The location of the additional double bond present in 7DHC but absent in cholesterol.

tilt and condensation effects, this may explain why neither effect has been reproduced in any of the CG representations simulated in this chapter.

In order to further explore the effect of 7DHC on bilayers and therefore the impact of this specific double bond, dual resolution simulations of binary systems consisting of DSPC/7DHC should be undertaken.

In contrast to the condensation and tilt effects discussed above, the effect of increased cholesterol concentration on the total bilayer bending modulus is different across the forcefields. The bending moduli calculated from MARTINI simulations behave qualitatively similarly to those from dual-resolution and atomistic simulations as cholesterol concentration is increased for both calculation methods. No such effect on bending modulus is seen for the ELBA parametrisations regardless of whether the extra beads to reflect cholesterol's rough and smooth faces are included.

The results in this chapter show that the parameterisation and modelling of cholesterol in forcefields is a subtle and complex task. Cholesterol's effects on bilayers are various and whilst one physical effect may be reproduced accurately in a forcefield, others may not. In particular, using the dual-resolution capability of ELBA in order to atomistically represent cholesterol whilst leveraging the speed-up available by modelling lipids and water with a CG forcefield is a quantitatively physically accurate representation of such binary systems, whereas any of the completely CG representation tested, are not.

Chapter 6

Conclusions

6.1 Conclusions and future work

This thesis has presented the results of molecular dynamics simulations with respect to three aspects of bilayer physical properties. First, with respect to bilayer phase transitions, we have shown that previously published experimental evidence for lyotropic and thermotropic phase transitions are reproduced in the ELBA1.0 and that the phases and transitions observed using ELBA1.0 are genuinely thermodynamically stable phases rather than sampling artefacts. In contrast the use of such enhanced sampling techniques has allowed us to conclude that reproduction of the phase diagram is not possible using the ELBA1.1 forcefield despite the two versions seemingly being quite similar. Trials of metadynamics techniques on this system have once again shown that the choice of collective variable to drive enhanced sampling is an extremely difficult task. Having found that one obvious descriptor of the phase transition is unsuitable, further work will be required to explore other possible collective variables; possibly including those based upon the geometric nature of head group packing in the gel phase and/or other collective variables chosen to drive nucleation of tail ordering.

Second, we have presented an analysis and direct comparison of five different techniques to measure bilayer bending rigidity; the original Helfrich Canham method, a Helfrich Canham based method to include consideration of lipid tilt, polymer brush theory, SALO and RSF. We have confirmed and quantified their respective size dependencies and confirmed that they all, at least qualitatively, reproduce the published experimental data for symmetric bilayers. However, further work is required by experimentalists, theoreticians and simulators to explore the applicability of existing measurement and analysis methods to asymmetric bilayers where the analysis of simulation and the measurements from experiment are very different. Any adjustment to the existing methods to consider each leaflet separately in an asymmetric bilayer must, of course, continue to consider them in aggregate in a symmetric bilayer.

Finally, turning to the complex nature of cholesterol's interactions with the phospholipid bilayer, we have compared atomistic, dual-resolution and CG simulation to elucidate the various effects that occur as cholesterol concentration is increased. Both the dual-resolution and atomistic classes of simulations showed cholesterol condensation. They also showed similar levels of cholesterol-induced PC tail ordering. By contrast, none of the CG simulation data showed cholesterol condensation when analysed in the same way - though MARTINI did display non-ideal mixing and also showed the most tail ordering of the CG forcefields. Taken together, these results accord with previous reported results that the cholesterol condensation effect is due to the ordering (and phase transition) of the phospholipid component in such binary systems. The results from the cholesterol tilt analyses followed a similar pattern. In contrast, we found the effect of increased cholesterol concentration on the total bilayer bending modulus to be different across the forcefields.

Our results show that the modelling and parameterisation of cholesterol is a subtle and complex task. Cholesterol's effects on bilayers are various and whilst one physical effect may be reproduced accurately using a particular forcefield, others may not. In particular, we have found that the detailed parameterisation of the cholesterol molecules is arguably more important that the PC lipids for the physical properties of the resultant binary systems. This means that using the dual-resolution capability of ELBA to atomistically represent cholesterol whilst leveraging the speed-up available by modelling lipids and water with a CG forcefield is a quantitatively physically accurate representation for many effects, whereas all of the completely CG representation tested, are not.

Further work is required to explore which details of cholesterol's structure, particular individual double bonds positions, must be reproduced even at a coarse grained level of abstraction, to reproduce the 'macro' effects on the bilayer. It is hoped that such exploration will also definitively show the precise nature of cholesterol's intriguing interactions.

Appendix A

CHARMM Membrane Builder CHARMM36 Standard Equilibration (generated by the CHARMM Membrane builder)

The following protocol is the 'standard' steps provided to download by the CHARMM Membrane Builder[157,158] website in the form of a series of 6 '.mdp' files for use in GROMACS for the equilibration of CHARMM36 membranes constructed using the site. Parameters integral to the forcefield (such as LJ cut-offs, treatment of long-range electrostatics etc) are not included in the descriptions of the steps here, but discussed above in chapter 1.

1. Energy minimisation
    - Maximum of 5000 energy minimisation steps
    - Steepest descent method
    - Maximum step size = 0.01 nm
    - Tolerance is 1000.0 kJ.mol.nm^{-1}
2. MD - NVT
    - Timestep = 1 fs
    - Length of simulation = 0.25 ps
    - Berendsen thermostat, lipids and water thermostatted separately, $\tau_T = 1.0$
3. MD - NVT
    - timestep = 1 fs
    - length of simulation = 0.25 ps
    - Berendsen thermostat, lipids and water thermostatted separately, $\tau_T = 1.0$
4. MD - NPT
    - timestep = 1 fs
    - length of simulation = 0.25 ns
    - Berendsen thermostat, lipids and water thermostatted separately, $\tau_T = 1.0$
    - Berendsen barostat, semiisotropic coupling, $\tau_P = 5.0$, compressibility $= 4.5x10^{-5}$
5. MD - NPT

- timestep = 2 fs
- length of simulation = 1.5 ps
- Berendsen thermostat, lipids and water thermostatted separately, $\tau_T = 1.0$
- Berendsen barostat, semiisotropic coupling, $\tau_P = 5.0$, compressibility $= 4.5x10^{-5}$

Appendix B

CHARMM Membrane Builder MARTINI Standard Equilibration (generated by the MARTINI Maker)

The following protocol is the ‘standard’ steps provided to download by the CHARMM Membrane Builder[157,180] website in the form of a series of 6 ‘.mdp’ files for use in GROMACS for the equilibration of MARTINI membranes constructed using the site. Parameters integral to the forcefield (such as LJ cut-offs, treatment of long-range electrostatics etc) are not included in the descriptions of the steps here, but discussed above in chapter 1.

1. Energy minimisation
    - Maximum of 5000 energy minimisation steps
    - Steepest descent method
    - Maximum step size = 0.01 nm
    - Tolerance is 10.0 kJ.mol.nm^{-1}
2. MD with restrained lipid headgroup
    - Phosphate beads restrained in the xy plane using force constant = 200 kJ.mol^{-1}.nm^{-2}
    - Timestep = 2 fs
    - Length of simulation = 1 ns
    - V-rescale thermostat, lipids and water thermostatted separately, $\tau_T = 1.0$
    - Berendsen barostat, semiisotropic coupling, $\tau_P = 5.0$, compressibility $= 3x10^{-4}$
3. MD with restrained lipid headgroup
    - Phosphate beads restrained in the xy plane using force constant = 100 kJ.mol^{-1}.nm^{-2}
    - timestep = 5 fs
    - length of simulation = 1 ns
    - v-rescale thermostat, lipids and water thermostatted separately, $\tau_T = 1.0$
    - berendsen barostat, semiisotropic coupling, $\tau_P = 5.0$, compressibility $= 3x10^{-4}$
4. MD with restrained lipid headgroup

- Phosphate beads restrained in the xy plane using force constant = 50 kJ.mol^{-1}.nm^{-2}
- timestep = 10 fs
- length of simulation = 1 ns
- v-rescale thermostat, lipids and water thermostatted separately, τ_T = 1.0
- berendsen barostat, semiisotropic coupling, τ_P = 5.0, compressibility = $3x10^{-4}$

5. MD with restrained lipid headgroup
    - Phosphate beads restrained in the xy plane using force constant = 20 kJ.mol^{-1}.nm^{-2}
    - timestep = 15 fs
    - length of simulation = 0.75 ns
    - v-rescale thermostat, lipids and water thermostatted separately, τ_T = 1.0
    - berendsen barostat, semiisotropic coupling, τ_P = 5.0, compressibility = $3x10^{-4}$
6. MD with restrained lipid headgroup
    - Phosphate beads restrained in the xy plane using force constant = 10 kJ.mol^{-1}.nm^{-2}
    - timestep = 20 fs
    - length of simulation = 1 ns
    - v-rescale thermostat, lipids and water thermostatted separately, τ_T = 1.0
    - berendsen barostat, semiisotropic coupling, τ_P = 5.0, compressibility = $3x10^{-4}$

Appendix C

Full parameterisation of bonded interactions for ELBA 1 cholesterol (no extra methyl beads)

The bead identifying letters are in accordance with figure 5.3.

C.1 Bonds

Beads	K_r (kJ/Å)	r_0 (Å)
OH A	677.7	1.826
A C	243.4	2.110
A B	228.8	3.322
C D	358.6	2.691
B C	534.4	2.302
B D	643.7	2.685
D F	310.8	2.328
D F	566.2	3.275
E F	488.7	2.386
F G	82.4	2.445
E G	333.8	3.117
G H	38.0	2.986
H J	33.0	3.001

C.2 Angles

Beads	K_θ (kJ/mol)	θ_0 (°)
OH A C	628.3	160.00
OH A B	554.0	140.00

Continued on next page

Table C.2 – continued from previous page

Beads	K_θ (kJ/mol)	θ_0 (°)
A C D	404.1	140.00
A B D	825.7	90.00
C D F	445.0	147.00
C D E	1661.4	155.00
B D F	1000.0	138.00
B D E	787.2	90.00
D F G	475.0	150.00
D E G	430.0	90.00
F G H	73.0	150.00
E G H	43.2	112.50
G H J	80.53	150.0
OH dipole	1.195	132.5

C.3 Dihedrals

Beads	K_{dihed} (kJ/mol)	ϕ_0 (°)
A B D C	52.504	-37
A B D E	7.224	-66
D C A B	67.592	33
C A B D	170.536	-18
A B C D	92.845	-156
A B D C	187.445	16

Continued on next page

Table C.3 – continued from previous page

Beads	K_{dihed} (kJ/mol)	ϕ_0 (°)
C D F E	14.600	149
C D F G	4.000	-145
F D C B	54.320	-143
C D E F	9.135	-137
E D C B	11.016	-29
C B D F	46.368	150
B D F E	65.556	26
B D F G	30.356	89
B D E F	131.357	-163
D F E G	97.069	156
G F D E	48.780	64
F D E G	146.061	-19
D E G F	191.556	17

Appendix D

Full parameterisation of bonded interactions for ELBA 2 cholesterol (with extra methyl beads)

The bead identifying letters are in accordance with figure 5.4.

D.1 Bonds

Beads	K_r (kJ/Å)	r_0 (Å)
OH A	677.7	1.826
A B	900.2	2.110
A C	228.8	3.322
B D	464.6	2.385
B C	251.4	2.726
C D	643.7	2.685
D F	935.8	2.181
D E	566.2	3.275
E F	657.7	2.514
F G	103.7	2.503
E G	333.8	3.117
G H	38.0	2.986
H J	33.0	3.001
B K	541.2	1.942
F L	644.4	1.957

D.2 Angles

Beads	K_θ (kJ/mol)	θ_0 (°)
OH A B	2785.3	160.00
OH A C	554.0	140.00
A B D	1251.1	140.00
A C D	825.7	90.00
B D F	1423.0	150.00
B D E	1539.4 155.00	
C D F	2089.0 144.00	
C D E	834.2 94.00	
D F G	3171.0 163.00	
D E G	443.0 93.00	
F G H	73.0 156.00	
E G H	43.2 112.50	
G H J	80.53 150.0	
A B K	160.21 100.0	
D B K	423.80 75.00	
C B K	135.80 80.00	
D F L	230.80 95.00	
E F L	417.80 71.00	
OH dipole	1.195	132.5

D.3 Dihedrals

Beads	K_{dihed} (kJ/mol)	ϕ_0 (°)
A B D C	24.259	3
A B D E	5.384	17
D B A C	30.362	-2
B A C D	129.678	1
K B A C	119.246	-79
A C D B	191.205	-1
OH A B K	4.917	102
L F D E	363.702	-71
A C B K	136.808	101
B D F K	17.722	122
E D B C	9.871	14
C D B K	104.104	86
C D F E	47.136	-7
C D F G	5.972	-20
C D F L	48.795	-78
D F B K	229.079	-78
G F D E	6.907	-12
F D E G	123.714	3
D E G F	176.991	-2
D E F L	199.751	96
F D B K	29.904	-107

Continued on next page

Table D.3 – continued from previous page

Beads	K_{dihed} (kJ/mol)	ϕ_0 (°)
E D B K	7.987	101
E G F L	287.991	73
G E F L	144.994	-81

Bibliography

[1] Luisi, P. L. (2003) Autopoesis: A review and reappraisal. *Naturwissenschaften 90*, 49–59.

[2] Tieleman, D. P., Marrink, S. J., and Berendsen, H. J. C. (1997) A computer perspective of membranes: molecular dynamics studies of lipid bilayer systems. *Biochimica et Biophysica Acta 1331*, 235–270.

[3] Cantor, R. S. (1998) The lateral pressure profile in membranes:a physical mechanism of general anesthesia. *Toxicology Letters 100*, 451–458.

[4] Lee, A. G. (2004) How lipids affect the activities of integral membrane proteins. *Biochimica et Biophysica Acta (BBA) - Biomembranes 1666*, 62–87.

[5] Gullingsrud, J., and Schulten, K. (2004) Lipid Bilayer Pressure Profiles and Mechanosensitive Channel Gating. *Biophysical Journal 86*, 3496–3509.

[6] Jensen, M. Ø., and Mouritsen, O. G. (2004) Lipids do influence protein function - the hydrophobic matching hypothesis revisited. *Biochimica et Biophysica Acta (BBA) - Biomembranes 1666*, 205–226.

[7] Sanders, C. R., and Hutchison, J. M. (2018) Membrane properties that shape the evolution of membrane enzymes. *Current Opinion in Structural Biology 51*, 80–91.

[8] Cordero-Morales, J. F., and Vásquez, V. (2018) How lipids contribute to ion channel function, a fat perspective on direct and indirect interactions. *Current Opinion in Structural Biology 51*, 92–98.

[9] Stillwell, W. *Introduction to Biological Membranes*, 2nd ed.; Elsevier B.V., 2016.

[10] Verkman, A. (1999) Membrane permeability: 100 years since Ernest Overton. *Current Topics in Membranes 48*, 97–121.

[11] Kleinzeller, A. (1997) Ernest Overton's Contribution to the Cell Membrane Concept: A Centennial Appreciation. *Physiology 12*, 49–53.

[12] Gorter, E., and Grendel, F. (1925) On Biomolecular Layers of Lipoids on the Chromocytes of the Blood. *Journal of Experimental Medicine 41*, 439–443.

[13] Leathes, J. (1925) Croonian Lectures: On the role of fats in vital phenomena. *The Lancet 205*, 853–856.

[14] Frederic, D. J., and Hugh, D. (2018) A contribution to the theory of permeability of thin films. *Journal of Cellular and Comparative Physiology 5*, 495–508.

[15] Singer, S. J., and Nicholson, G. L. (1972) The fluid mosaic of the structure of cell membranes. *Science 175*, 720.

[16] Simons, K., and Ikonen, E. (1997) Functional rafts in cell membranes. *Nature 387*, 569–572.

[17] Berg, J. M., Tymoczko, J. L., and Stryer, L. In *Biochemistry, Fifth Edition*, 5th ed.; Berg, J., Tymoczko, J., and Stryer, L., Eds.; W.H. Freeman, 2002.

[18] Mouritsen, O. G. *Life as a Matter of Fat. The Emerging Science of Lipidomics; Springer, first ed., 2005.*, 1st ed.; Springer: New York, 2005.

[19] Silverstein, T. P. (1998) The Real Reason Why Oil and Water Don't Mix. *Journal of Chemical Education 75*, 116.

[20] Tanford, C. (1972) Micelle shape and size. *The Journal of Physical Chemistry 76*, 3020–3024.

[21] Israelachvili, J. N., Mitchell, D. J., and Ninham, B. W. (1976) Theory of self-assembly of hydrocarbon amphiphiles into micelles and bilayers. *Journal of the Chemical Society, Faraday Transactions 2 72*, 1525.

[22] Boal, D. *Mechanics of the Cell*; Cambridge University Press, 2002.

[23] Nagarajan, R. (2002) Molecular Packing Parameter and Surfactant Self-Assembly: The Neglected Role of the Surfactant Tail. *Langmuir 18*, 31–38.

[24] Loudet, C., Diller, A., Grélard, A., Oda, R., and Dufourc, E. J. (2010) Biphenyl phosphatidylcholine: a promoter of liposome deformation and bicelle collective orientation by magnetic fields. *Progress in Lipid Research 49*, 289–297.

[25] Ipsen, J. H., Karlstrom, G., Mouritsen, O. G., Wennerstrom, H., and Zuckermann, M. J. (1987) Phase equilibria in the phosphatidylcholine-cholesterol system. *Biochim Biophys Acta 905*, 162–172.

[26] Owicki, J. C., and McConnell, H. M. (1980) Lateral Diffusion in inhomogeneous membranes including cholesterol. *Biophys J. 30*, 383–397.

[27] Lange, Y., Swaisgood, M. H., Ramos, B. V., and Steck, T. L. (1989) Plasma membranes contain half the phospholipid and 90% of the cholesterol and sphingomyelin in cultured human fibroblasts. *J. Biol. Chem. 264*, 3786–3793.

[28] Brown, D., and London, E. (1998) Functions of lipid rafts in biological membranes. *Annual Review of Cell and Developmental Biology 14*, 111–36.

[29] Munro, S. (2003) Lipid Rafts: Elusive or Illusive? *Cell 115*, 377–388.

[30] Edidin, M. (2003) The state of lipid rafts: from model membranes to cells. *Annu. Rev. Biophys. Biomol. Struct. 32*, 257–283.

[31] Veatch, S. L., and Keller, S. L. (2002) Lateral organization in lipid membranes containing cholesterol. *Phys. Rev. Lett. 89*, 268101.

[32] Silvius, J. R. (2003) Role of Cholesterol in Lipid Raft formation: lessons from lipid model systems. *Biochim. Biophys Act. 1610*, 174–183.

[33] Baumgart, T., Hess, S. T., and Webb, W. W. (2003) Imaging coexisting fluid domains in biomembrane models coupling curvature and line tension. *Biophysical Journal 425*, 821–824.

[34] Sezgin, E., Levental, I., Mayor, S., and Eggeling, C. (2017) The mystery of membrane organization: composition, regulation and roles of lipid rafts. *Nature Reviews Molecular Cell Biology 18*, 361.

[35] Niemela, P. S., Ollila, S., Hyvonen, M. T., Karttunen, M., and Vattullinen, I. (2007) Assessing the nature of lipid raft membranes. *PLoS Comput. Bio. 3*, 334.

[36] Risselada, J., and Marrink, S. J. (2008) The molecular face of lipid rafts in model membranes. *Proceedings of the National Academy of Sciences of the United States of America 105*, 17367–17372.

[37] Lingwood, D., and Simons, K. (2010) Lipid rafts as a membrane organising principle. *Science 327*, 46–50.

[38] Sevcsik, E., Brameshuber, M., Fölser, M., Weghuber, J., Honigmann, A., and Schütz, G. J. (2015) GPI-anchored proteins do not reside in ordered domains in the live cell plasma membrane. *Nature Communications 6*, 6969.

[39] Lyman, E., Chia-Lung, H., and Eggeling, C. (2018) From Dynamics to Membrane Organization: Experimental Breakthroughs Occasion a "Modelling Manifesto". *Biophysical Journal 115*, 595–604.

[40] Jämbeck, J. P. M., and Lyubartsev, A. P. (2012) Derivation and Systematic Validation of a Refined All-Atom Force Field for Phosphatidylcholine Lipids. *The Journal of Physical Chemistry B 116*, 3164–3179.

[41] Markvoort, B., Spijker, P., and Hilbers, P. PumMa Image Gallery. `http://cbio.bmt.tue.nl/pumma/index.php/Main/Gallery`.

[42] McQuarrie, D. *Statistical Mechanics*; Harper and Row: New York, 1976.

[43] Maitland, G. C., Rigby, M., Smith, E. B., and Wakeham, W. *Intermolecular Forces*; Clarendon Press: Oxford, 1981.

[44] Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An NlogN method for Ewald sums in large systems. *The Journal of Chemical Physics 98*, 10089–10092.

[45] Orsi, M., Michel, J., and Essex, J. W. (2010) Coarse-grain modelling of DMPC and DOPC lipid bilayers. *Journal of Physics. Condensed matter : an Institute of Physics Journal 22*, 155106.

[46] Hill, R. (1963) Elastic properties of reinforced solids: some theoretical principles. *Journal of the Mechanics and Physics of Solids 11*, 357–372.

[47] Klauda, J. B., Venable, R. M., Freites, J. A., O'Connor, J. W., Tobias, D. J., Mondragon-Ramirez, C., Vorobyov, I., MacKerell, A. D., and Pastor, R. W. (2010) Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *The Journal of Physical Chemistry B 114*, 7830–7843.

[48] Pastor, R. W., and MacKerell, A. D. (2011) Development of the CHARMM Force Field for Lipids. *The Journal of Physical Chemistry Letters 2*, 1526–1532.

[49] Jämbeck, J. P. M., and Lyubartsev, A. P. (2012) An Extension and Further Validation of an All-Atomistic Force Field for Biological Membranes. *Journal of Chemical Theory and Computation 8*, 2938–2948.

[50] Marrink, S. J., Vries, A. H. D., and Mark, A. E. (2004) Coarse Grained Model for Semiquantitative Lipid Simulations. *PoLAR* 750–760.

[51] Marrink, S., Risselada, H., Yefimov, S., Tieleman, D., and de Vries, A. (2007) The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *Journal of Physical Chemistry B 111*, 7812–7824.

[52] Páll, S., and Hess, B. (2013) A flexible algorithm for calculating pair interactions on SIMD architectures. *Computer Physics Communications 184*, 2641–2650.

[53] Orsi, M., and Essex, J. W. Permeability of drugs and hormones through a lipid bilayer: insights from dual-resolution molecular dynamics. *Soft Matter 6*, 3797–3808.

[54] Orsi, M., Haubertin, D. Y., Sanderson, W. E., and Essex, J. W. (2008) A quantitative coarse-grain model for lipid bilayers . *Journal of Physical Chemistry B 112*, 802.

[55] Orsi, M., Noro, M. G., and Essex, J. W. Dual-resolution molecular dynamics simulation of antimicrobials in biomembranes. *Journal of the Royal Society: Interface 8*, 826–841.

[56] Orsi, M., and Essex, J. W. (2011) The ELBA Force Field for Coarse-Grain Modelling of Lipid Membranes. *PloS One 6*, 1–36.

[57] Orsi, M., and Essex, J. W. (2013) Physical properties of mixed bilayers containing lamellar and nonlamellar lipids: insights from coarse grained molecular dynamics simulations. *Faraday Discussions 161*, 249–272.

[58] Steinbach, P. J., and Brooks, B. R. (1994) New spherical cut-off methods for long-range forces in macromolecular simulation. *J. Comput. Chem. 15*, 667–683.

[59] Yamashita, K., and Daiguji, H. (2016) Coarse-grained molecular dynamics simulations of capillary evaporation of water confined in hydrophilic mesopores. *Molecular Physics 114*, 884–894.

[60] Siani, P., de Souza, R. M., Dias, L. G., Itri, R., and Khandelia, H. (2016) An overview of molecular dynamics simulations of oxidized lipid systems, with a comparison of ELBA and MARTINI force fields for coarse grained lipid simulations. *Biochimica et Biophysica Acta - Biomembranes 1858*, 2498–2511.

[61] Hockney, R., Goel, S., and Eastwood, J. (1974) Quiet high-resolution computer models of a plasma. *Journal of Computational Physics 14*, 148–158.

[62] Swope, W. C., Anderson, H. C., Berens, P. H., and Wilson, K. R. (1982) A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: application to small water clusters. *J Chem Phys 76*, 637–649.

[63] Dullweber, A., Leimkuhler, B., and McLachlan, R. (1997) Symplectic splitting methods for rigid body molecular dynamics. *Journal of Chemical Physics 107*, 5840–5851.

[64] Allen, M. P., and Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press, USA, 1989.

[65] Tuckerman, M., Berne, B. J., and Martyna, G. J. (1992) Reversible multiple time scale molecular dynamics. *The Journal of Chemical Physics 97*, 1990–2001.

[66] Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics 81*, 3684.

[67] Morishita, T. (2000) Fluctuation formulas in molecular dynamics simulations with the weak coupling heat bath. *Journal of Chemical Physics 113*, 2976–2982.

[68] Nosé, S. (1984) A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics 52*, 255–268.

[69] Hoover, W. G. (1985) Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A 31*, 1695–1697.

[70] Abraham, M., van der Spoel, D., Lindahl, E., and Hess, B. GROMACS Groningen Machine for Chemical Simulations USER MANUAL v 5.0.7. 2015; `ftp://ftp.gromacs.org/pub/manual/manual-5.0.7.pdf`.

[71] Frenkel, D., Smit, B., Tobochnik, J., McKay, S. R., and Christian, W. *Computers in Physics*; Academic Press, Inc. Orlando, FL, USA, 1997; Vol. 11; p 351.

[72] Parrinello, M., and Rahman, A. (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics 52*, 7182–7190.

[73] Rand, P. R., and Parsegian, V. A. In *Current Topics in Membranes 44*; Epand, R. M., Ed.; Academic Press: San Diego, 1997; pp 167–189.

[74] Bryant, G., and Wolfe, J. (1992) Interfacial Forces in Cryobiology and Anhydrobiology. *Cryo Letters 13*, 23–28.

[75] Koynova, R., and Caffrey, M. (1998) Phases and phase transitions of the phosphatidylcholines. *Biochimica Biophysica Acta 1376*, 91–145.

[76] Nagle, J. F., and Tristram-Nagle, S. (2000) Structure of lipid bilayers. *Biochimica Biophysica Acta 1469*, 159–195.

[77] Marsh, D. (2010) Structural and thermodynamic determinants of chain-melting transition temperatures for phospholipid and glycolipids membranes. *Biochimica et Biophysica Acta (BBA) - Biomembranes 1798*, 40–51.

[78] Marsh, D. (1991) Analysis of the chainlength dependence of lipid phase transition temperatures: Main and pretransitions of phosphatidylcholines; main and non-lamellar transitions of phosphatidylethanolamines. *Biochimica et Biophysica Acta (BBA) - Biomembranes 1062*, 1–6.

[79] Seddon, J. M., Cevc, G., and Marsh, D. (1983) Calorimetric studies of the gel-fluid (L.beta.-L.alpha.) and lamellar-inverted hexagonal (L.alpha.-HII) phase transitions in dialkyl- and diacylphosphatidylethanolamines. *Biochemistry 22*, 1280–1289.

[80] Lipka, G., Chowdhry, B. Z., and Sturtevant, J. M. (1984) A comparison of the phase transition properties of 1,2-diacylphosphatidylcholines and 1,2-diacylphosphatidylethanolamines in water and deuterium oxide. *The Journal of Physical Chemistry 88*, 5401–5406.

[81] Burke, T. G., Rudolph, A. S., Price, R. R., Sheridan, J. P., Dalziel, A. W., Singh, A., and Schoen, P. E. (1988) Differential scanning calorimetric study of the thermotropic phase behavior of a polymerizable, tubule-forming lipid. *Chemistry and Physics of Lipids 48*, 215–230.

[82] Lewis, R. N. A. H., Mak, N., and McElhaney, R. N. (1987) A differential scanning calorimetric study of the thermotropic phase behavior of model membranes composed of phosphatidylcholines containing linear saturated fatty acyl chains. *Biochemistry 26*, 6118–6126.

[83] Lewis, R., Mak, N., and McElhaney, R. N. (1987) A differential scannning calorimetric study of the thermotropic phase behaviour of model membranes composed of phosphatidylcholines containing linear saturated fatty acyl chains. *Biochemistry 26*, 6118–6129.

[84] Lewis, R. N. A. H., Sykes, B. D., and McElhaney, R. N. (1988) Thermotropic phase behavior of model membranes composed of phosphatidylcholines containing cis-monounsaturated acyl chain homologs of oleic acid: differential scanning calorimetric and phosphorus-31 NMR spectroscopic studies. *Biochemistry 27*, 880–887.

[85] Böckmann, R. A., Hac, A., Heimburg, T., and Grubmüller, H. (2003) Effect of Sodium Chloride on a Lipid Bilayer. *Biophysical Journal 85*, 1647–1655.

[86] Janiak, M., Small, D., and Shipley, G. (1978) Temperature and Compositional Dependence of the Structure of Hydrated Dimyristoyl Lecithin. *The Journal of Biological Chemistry 254*, 6068–6078.

[87] Qin, S.-S., Yu, Z.-W., and Yu, Y.-X. (2009) Structural characterization on the gel to liquid-crystal phase transition of fully hydrated DSPC and DSPE bilayers. *The journal of physical chemistry. B 113*, 8114–8123.

[88] Marrink, S. J., Risselada, J., and Mark, A. E. (2005) Simulation of gel phase formation and melting in lipid bilayers using a coarse grained model. *Chemistry and Physics of Lipids 135*, 223–244.

[89] Stevens, M. J. (2004) Coarse-grained simulations of lipid bilayers. *The Journal of Chemical Physics 121*, 11942–11948.

[90] Faller, R., and Marrink, S. J. (2004) Simulation of domain formation in DLPC-DSPC mixed bilayers. *Langmuir 20*, 7686–7693.

[91] Bennun, S. V., Longo, M., and Faller, R. (2007) Phase and Mixing Behavior in Two-Component Lipid Bilayers: A Molecular Dynamics Study in DLPC/DSPC Mixtures. *The Journal of Physical Chemistry B 111*, 9504–9512.

[92] Marsh, D. *CRC Handbook of Lipid Bilayers*, 2nd ed.; CRC Press: New York, 1990.

[93] Orsi, M. brahms-md. 2008; `http://code.google.com/p/brahms-md/`.

[94] Tristram-Nagle, S., Zhang, R., Suter, R. M., Worthington, C. R., and Sun, W. J. (1993) Measurement of chain tilt angle in fully hydrated bilayers of gel phase lecithins. *Biophysical Journal 64*, 1097–1109.

[95] Balgavý, P., Dubničková, M., Kučerka, N., Kiselev, M. A., Yaradaikin, S. P., and Uhriková, D. (2001) Bilayer thickness and lipid interface area in unilamellar extruded 1,2-diacylphosphatidylcholine liposomes: a small-angle neutron scattering study. *Biochimica et Biophysica Acta (BBA) - Biomembranes 1512*, 40–52.

[96] Matsumoto, M., Saito, S., and Ohmine, I. (2002) Molecular dynamics simulation of the ice nucleation and growth process leading to water freezing. *Nature 416*, 409.

[97] Wales, D. J., Doye, J. P. K., Miller, M. A., Mortenson, P. N., and Walsh, T. R. (2000) Energy landscapes: From clusters to biomolecules. *Advances in Chemical Physics 115*, 1.

[98] Mucha, M., and Jungwirth, P. (2003) Salt crystallization from an Evaporating Aqueous Solution by Molecular Dynamics Simulations. *Journal of Physical Chemistry B 107*, 8271.

[99] Darve, E., and Pohorille, A. (2001) Calculating free energies using average force. *The Journal of Chemical Physics 115*, 9169–9183.

[100] Laio, A., and Parrinello, M. (2002) Escaping free-energy minima. *Proceedings of the National Academy of Sciences of the United States of America 99*, 12562–6.

[101] Barducci, A., Bussi, G., and Parrinello, M. (2008) Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Physical Review Letters 100*, 020603.

[102] Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C., and Bussi, G. (2014) PLUMED 2: New feathers for an old bird. *Computer Physics Communications 185*, 604–613.

[103] Torrie, G., and Valleau, J. (1977) Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics 23*, 187–199.

[104] Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., and Kollman, P. A. (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. *Journal of Computational Chemistry 13*, 1011–1021.

[105] Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., and Kollman, P. A. (1995) Multidimensional free-energy calculations using the weighted histogram analysis method. *Journal of Computational Chemistry 16*, 1339–1350.

[106] Grossfield, A. WHAM: the weighted histogram analysis method.

[107] Lemkul, J. GROMACS Tutorial - Step Five: Generating Configurations. `http://www.bevanlab.biochem.vt.edu/Pages/Personal/justin/gmx-tutorials/umbrella/05{_}pull.html`.

[108] Van Meer, G., Voelker, D. R., and Feigenson, G. W. (2008) Membrane lipids: Where they are and how they behave. *Nature Reviews Molecular Cell Biology 9*, 112–124.

[109] Bretscher, M. S. (1972) Phosphatidyl-ethanolamine: differential labelling in intact cells and cell ghosts of human erythrocytes by a membrane-impermeable reagent. *Journal of molecular biology 71*, 523–528.

[110] Zachowski, A. (1993) Phospholipids in animal eukaryotic membranes: transverse asymmetry and movement. *Biochemical Journal 294*, 1 – 14.

[111] Kamio, Y., and Nikaido, H. (1976) Outer membrane of Salmonella typhimurium: accessibility of phospholipid head groups to phospholipase C and cyanogen bromide activated dextran in the external medium. *Biochemistry 15*, 2561–2570.

[112] Charalambous, K., Booth, P. J., Woscholski, R., Seddon, J. M., Templer, R. H., Law, R. V., Barter, L. M. C., and Ces, O. (2012) Engineering de novo membrane-mediated protein-protein communication networks. *Journal of the American Chemical Society 134*, 5746–5749.

[113] Middleton Boon, J., and Smith, B. D. (2002) Chemical control of phospholipid distribution across bilayer membranes. *Medicinal Research Reviews 22*, 251–281.

[114] Daleke, D. (2007) Phospholipid Flippases. *The Journal of Biological Chemistry 282*, 821–825.

[115] Perozo, E., Kloda, A., Cortes, D. M., and Martinac, B. (2002) Physical principles underlying the transduction of bilayer deformation forces during mechanosensitive channel gating. *Nature Structural Biology 9*, 696–703.

[116] Andersen, O. S., and Koeppe, R. E. (2007) Bilayer Thickness and Membrane Protein Function: An Energetic Perspective. *Annual Review of Biophysics and Biomolecular Structure 36*, 107–130.

[117] Rappolt, M., and Pabst, G. *Structure and Dynamics of Membranous Interfaces*; 2008; pp 45–81.

[118] Evans, E., and Needham, D. (1987) Physical properties of surfactant bilayer membranes: thermal transitions, elasticity, rigidity, cohesion and colloidal interactions. *The Journal of Physical Chemistry 91*, 4219–4228.

[119] Rawicz, W., Olbrich, K. C., McIntosh, T., Needham, D., and Evans, E. (2000) Effect of Chain Length and Unsaturation on Elasticity of Lipid Bilayers. *Biophysical Journal 79*, 328–339.

[120] Needham, D., and Nunn, R. S. (1990) Elastic deformation and failure of lipid bilayer membranes containing cholesterol. *Biophysical Journal 58*, 997–1009.

[121] Helfrich, W. (1973) Elastic properties of lipid bilayers: theory and possible experiments. *Z Naturforsch [C] 28*, 693–703.

[122] Dimova, R. (2014) Recent developments in the field of bending rigidity measurements on membranes. *Advances in Colloid and Interface Science 208*, 225–234.

[123] Bochicchio, D., and Monticelli, L. In *Chapter Five - The Membrane Bending Modulus in Experiments and Simulations: A Puzzling Picture*; Iglič, A., Kulkarni, C. V., Rappolt, M. B. T. A. i. B., and Self-Assembly, L., Eds.; Academic Press, 2016; Vol. 23; pp 117–143.

[124] Reeves, J., and Dowben, R. (1969) Formation and properties of thin-walled phospholipid vesicles. *J. Cell. Physiol 73*, 49–60.

[125] Angelova, M. I., and Dimitrov, D. S. (1986) Liposome electroformation. *Faraday Discussions Chem Soc 81*.

[126] Elani, Y., Purushothaman, S., Booth, P. J., Seddon, J. M., Brooks, N. J., Law, R. V., and Ces, O. (2015) Measurements of the effect of membrane asymmetry on the mechanical properties of lipid bilayers. *Chem. Commun. 51*, 6976–6979.

[127] Manley, S., and Gordon, V. D. (2008) Making Giant Unilamellar Vesicles via Hydration of a Lipid Film. *Current Protocols in Cell Biology 40*, 24.3.1–24.3.13.

[128] Pautot, S., Frisken, B. J., and Weitz, D. (2003) Production of Unilamellar Vesicles Using an Inverted Emulsion. *Langmuir 19*, 2870–2879.

[129] Pautot, S., Frisken, B. J., and Weitz, D. A. (2003) Engineering asymmetric vesicles. *Proceedings of the National Academy of Sciences of the United States of America 100*, 10718–21.

[130] Helfrich, W. (1978) Steric Interaction of Fluid Membranes in Multilayer Systems. *Z Naturforsch [C] 33*, 305–313.

[131] Brochard, F., and Lennon, J. F. (1975) Frequency spectrum of the flicker phenomenon in erythrocytes. *J. Phys.(Paris) 36*, 1035–1047.

[132] Schneider, M. B., Jenkins, J. T., and Webb, W. W. (1984) Thermal fluctuations of large quasi-spherical bimolecular phospholipid vesicles. *J. Phys. France 45*, 1457–1472.

[133] Faucon, J., Mitov, M. D., Méléard, P., Bivas, I., and Bothorel, P. (1989) Bending elasticity and thermal fluctuations of lipid membranes. Theoretical and experimental requirements. *Journal de Physique 50*, 2389–2414.

[134] Nagle, J. F., Jablin, M. S., Tristram-Nagle, S., and Akabori, K. (2015) What are the true values of the bending modulus of simple lipid bilayers? *Chemistry and Physics of Lipids 185*, 3–10.

[135] Lei, N., Safinya, C. R., and Bruinsma, R. F. (1995) Discrete Harmonic Model for Stacked Membranes: Theory and Experiment. *J. Phys. II France 5*, 1155–1163.

[136] Zhang, R., Tristram-Nagle, S., Sun, W., Headrick, R., Irving, T., Suter, R., and Nagle, J. (1996) Small-angle x-ray scattering from lipid bilayers is well described by modified Caillé theory but not by paracrystalline theory. *Biophysical Journal 70*, 349–357.

[137] Nagle, J. F. (2013) Introductory Lecture: Basic quantities in model biomembranes. *Faraday Discussions 161*, 11–29.

[138] Gracià, R. S., Bezlyepkina, N., Knorr, R. L., Lipowsky, R., and Dimova, R. (2010) Effect of cholesterol on the rigidity of saturated and unsaturated membranes: fluctuation and electrodeformation analysis of giant vesicles. *Soft Matter 6*, 1472.

[139] Lindahl, E., and Edholm, O. (2000) Mesoscopic undulations and thickness fluctuations in lipid bilayers from molecular dynamics simulations. *Biophysical Journal 79*, 426–433.

[140] Waheed, Q., and Edholm, O. (2009) Undulation Contributions to the Area Compressibility in Lipid Bilayer Simulations. *Biophysical Journal 97*, 2754–2760.

[141] Hu, M., Diggins, P., and Deserno, M. (2013) Determining the bending modulus of a lipid membrane by simulating buckling. *The Journal of Chemical Physics 138*, 214110.

[142] Watson, M. C., Brandt, E. G., Welch, P. M., and Brown, F. L. H. (2012) Determining Biomembrane Bending Rigidities from Simulations of Modest Size. *Physical Review Letters 109*, 28102.

[143] Hofsäß, C., Lindahl, E., and Edholm, O. (2003) Molecular dynamics simulations of phospholipid bilayers with cholesterol. *Biophysical Journal 84*, 2192–2206.

[144] Levine, Z. A., Venable, R. M., Watson, M. C., Lerner, M. G., Shea, J.-E., Pastor, R. W., and Brown, F. L. H. (2014) Determination of Biomembrane Bending Moduli in Fully Atomistic Simulations. *Journal of the American Chemical Society 136*, 13582–13585.

[145] Lu, L., Doak, W. J., Schertzer, J. W., and Chiarot, P. R. (2016) Membrane mechanical properties of synthetic asymmetric phospholipid vesicles. *Soft Matter 12*, 7521–7528.

[146] Flory, P. J., and Volkenstein, M. (1969) Statistical mechanics of chain molecules. *Biopolymers 8*, 699–700.

[147] Feller, S. E., and Pastor, R. W. (1999) Constant surface tension simulations of lipid bilayers: The sensitivity of surface areas and compressibilities. *The Journal of Chemical Physics 111*, 1281.

[148] Piggot, T. J., Piñeiro, Á., and Khalid, S. (2012) Molecular Dynamics Simulations of Phosphatidylcholine Membranes: A Comparative Force Field Study. *Journal of Chemical Theory and Computation 8*, 4593–4609.

[149] Canham, P. B. (1970) The minimum energy of bending as a possible explanation of the biconcave shape of the human red blood cell. *J Theor Biol 26*, 61–81.

[150] Safan, S. *Statistical Thermodynamics of Surfaces, Interfaces and Membranes*; Addison-Wesley Reading, MA, 1994.

[151] Venable, R. M., Brown, F. L. H., and Pastor, R. W. (2015) Mechanical properties of lipid bilayers from molecular dynamics simulation. *Chemistry and Physics of Lipids 192*, 60–74.

[152] May, E. R., Narang, A., and Kopelevich, D. I. (2007) Role of molecular tilt in thermal fluctuations of lipid membranes. *Physical Review E 76*, 21913.

[153] Khelashvili, G., Kollmitzer, B., Heftberger, P., Pabst, G., and Harries, D. (2013) Calculating the Bending Modulus for Multicomponent Lipid Membranes in Different Thermodynamic Phases. *Journal of Chemical Theory and Computation 9*, 3866–3871.

[154] Johner, N., Harries, D., and Khelashvili, G. (2014) Curvature and Lipid Packing Modulate the Elastic Properties of Lipid Assemblies: Comparing HII and Lamellar Phases. *The Journal of Physical Chemistry Letters 5*, 4201–4206.

[155] Johner, N., Harries, D., and Khelashvilli, G. (2016) Implementation of a methodology for determining elastic properties of lipid assemblies from molecular dynamics simulations. *BMC Bioinformatics 17*, 161.

[156] Doktorova, M., Harries, D., and Khelashvili, G. (2017) Determination of bending rigidity and tilt modulus of lipid membranes from real-space fluctuation analysis of molecular dynamics simulations. *Physical Chemistry Chemical Physics 19*, 16806–16818.

[157] Jo, S., Kim, T., Iyer, V. G., and Im, W. (2008) CHARMM-GUI: A web-based graphical user interface for CHARMM. *Journal of Computational Chemistry 29*, 1859–1865.

[158] Lee, J. et al. (2016) CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *Journal of Chemical Theory and Computation 12*, 405–413.

[159] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics 79*, 926–935.

[160] Berendsen, H. J. C., van der Spoel, D., and van Drunen, R. (1995) GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications 91*, 43–56.

[161] Lindahl, E., Hess, B., and van der Spoel, D. (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis. *Molecular Modeling Annual 7*, 306–317.

[162] Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. C. (2005) GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry 26*, 1701–1718.

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

[164] Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, M. R., Smith, J. C., Kasson, P. M., van der Spoel, D., Hess, B., and Lindahl, E. (2013) GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics 29*, 845–854.

[165] Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., and Lindahl, E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX 1-2*, 19–25.

[166] Nosé, S., and Klein, M. L. (1983) Constant pressure molecular dynamics for molecular systems. *Molecular Physics 50*, 1055–1076.

[167] Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *The Journal of Chemical Physics 98*, 10089–10092.

[168] Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh Ewald method. *The Journal of Chemical Physics 103*, 8577–8593.

[169] Steinbach, P. J., and Brooks, B. R. (1994) New spherical-cutoff methods for long-range forces in macromolecular simulation. *Journal of Computational Chemistry 15*, 667–683.

[170] Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997) LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry 18*, 1463–1472.

[171] Hess, B. (2008) P-LINCS:A Parallel Linear Constraint Solver for Molecular Simulation. *Journal of Chemical Theory and Computation 4*, 116–122.

[172] Shchelokovskyy, P., Tristram-Nagle, S., and Dimova, R. (2011) Effect of the HIV-1 fusion peptide on the mechanical properties and leaflet coupling of lipid bilayers. *New journal of physics 13*, 25004.

[173] Méléard, L. F.-P., Bivas, I., Mitov, M. D., and P, (1994) Temperature and Chain Length Effects on Bending Elasticity of Phosphatidylcholine Bilayers. *EPL (Europhysics Letters) 28*, 181.

[174] Nagle, J. F. (2017) Experimentally determined tilt and bending moduli of single-component lipid bilayers. *Chemistry and Physics of Lipids 205*, 18–24.

[175] Doktorova, M., Heberle, F., Dzikovski, B., Chandrasekaran, S., Katsaras, J., Feigensen, G., and Weinstein, H. Interleaflet Coupling in assymetric membranes: Protocols and Revelations (poster). 62nd Annual Meeting of the Biophysical Society. San Francisco, 2018.

[176] Weiner, M. Molecular Dynamics simulations reveal the impact of compositional asymmetry in phase separated lipid membranes on phospholipid physical properties. 62nd Annual Meeting of the Biophysical Society. San Francisco, 2018.

[177] Kucěrka, N., Tristram-Nagle, S., and Nagle, J. F. (2006) Structure of Fully Hydrated Fluid Phase Lipid Bilayers with Monounsaturated Chains. *Journal of Membrane Biology 208*, 193–202.

[178] Kučerka, N., Nieh, M. P., and Katsaras, J. (2011) Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature. *Biochimica et Biophysica Acta - Biomembranes 1808*, 2761–2771.

[179] Heberle, F. A., Pan, J., Standaert, R. F., Drazba, P., Kucerka, N., and Katsaras, J. (2012) Model-based approaches for the determination of lipid

bilayer structure from small-angle neutron and X-ray scattering data. *European Biophysics Journal 41*, 875–890.

[180] Qi, Y., Ingólfsson, H. I., Cheng, X., Lee, J., Marrink, S. J., and Im, W. (2015) CHARMM-GUI Martini Maker for Coarse-Grained Simulations with the Martini Force Field. *Journal of Chemical Theory and Computation 11*, 4486–4494.

[181] Lee, J. et al. (2016) CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *Journal of Chemical Theory and Computation 12*, 405–413.

[182] Fowler, P., Helie, J., Duncan, A., Chavent, M., Koldso, H., and Sansom, M. (2016) Membrane stiffness is modified by integral membrane proteins. *Soft Matter 12*, 7792–7803.

[183] Karamdad, K., Law, R. V., Seddon, J. M., Brooks, N. J., and Ces, O. (2016) Studying the effects of asymmetry on the bending rigidity of lipid membranes formed by microfluidics. *Chemical Communications 52*, 5277–5280.

[184] Sreekumari, A., and Lipowsky, L. (2018) Lipids with bulky head groups generate large membrane curvatures by small compositional asymmetries. *The Journal of Chemical Physics 149*, 084901.

[185] Liscum, L., and Underwood, K. W. (1995) Intracellular cholesterol transport and compartmentation. *The Journal of Biological Chemistry 270*, 15443–6.

[186] Liscum, L., and Munn, N. J. (1999) Intracellular cholesterol transport. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids 1438*, 19–37.

[187] Simons, K., and Ikonen, E. (2000) How cells handle cholesterol. *Science (New York, N.Y.) 290*, 1721–6.

[188] Mouritsen, O. G., and Zuckerman, M. J. (2004) What's so special about Cholesterol? *Lipids 39*, 1101–1113.

[189] Ikonen, E. (2008) Cellular cholesterol trafficking and compartmentalization. *Nature Reviews Molecular Cell Biology 9*, 125–138.

[190] Pan, J., Mills, T. T., Tristram-Nagle, S., and Nagle, J. F. (2008) Cholesterol Perturbs Lipid Bilayers Nonuniversally. *Physical Review Letters 100*, 198103.

[191] Pan, J., Tristram-Nagle, S., and Nagle, J. F. (2009) Effect of cholesterol on structural and mechanical properties of membranes depends on lipid chain saturation. *Physical Review E 80*, 021931.

[192] Almeida, P. F. F., Vaz, W. L. C., and Thompson, T. E. (1992) Lateral diffusion in the liquid phases of dimyristoylphosphatidylcholine/cholesterol lipid bilayers: a free volume analysis. *Biochemistry 31*, 6739–6747.

[193] McMullen, T. P. W., Lewis, R., and McElhaney, R. N. (2004) Cholesterol–phospholipid interactions, the liquid-ordered phase and lipid rafts in model and biological membranes. *Current Opinion in Colloid & Interface Science 8*, 459–468.

[194] McMullen, T. P., Vilchèze, C., McElhaney, R. N., and Bittman, R. (1995) Differential scanning calorimetric study of the effect of sterol side chain length and structure on dipalmitoylphosphatidylcholine thermotropic phase behavior. *Biophysical Journal 69*, 169–76.

[195] Vist, M. R., and Davis, J. H. (1990) Phase equilibria of cholesterol/dipalmitoylphosphatidylcholine mixtures: deuterium nuclear magnetic resonance and differential scanning calorimetry. *Biochemistry 29*, 451–464.

[196] Huang, T.-H., B Lee, C. W., Das Gupta, S. K., Blume, A., and Griffin, R. G. (1993) A 13C and 2H Nuclear Magnetic Resonance Study of Phosphatidylcholine/ Cholesterol Interactions: Characterization of Liquid-Gel Phases. *Biochemistry 32*, 13277–13287.

[197] Thewalt, J. L., and Bloom, M. (1992) Phosphatidylcholine: cholesterol phase diagrams. *Biophysical Journal 63*, 1176–1181.

[198] Shimshick, E. J., and McConnell, H. M. (1973) Lateral phase separation in phospholipid membranes. *Biochemistry 12*, 2351–60.

[199] Marsh, D. (2010) Liquid-ordered phases induced by cholesterol: A compendium of binary phase diagrams. *Biochimica et Biophysica Acta - Biomembranes 1798*, 688–699.

[200] Veatch, S. L., and Keller, S. L. (2005) Seeing spots: complex phase behavior in simple membranes. *BBA-Molecular Cell Research 1746*, 172–185.

[201] Almeida, P. F. (2011) A Simple Thermodynamic Model of the Liquid-Ordered State and the Interactions between Phospholipids and Cholesterol. *Biophysical Journal 100*, 420–429.

[202] Simons, K., and Vaz, W. L. C. (2004) Model Systems, Lipid Rafts, and Cell Membranes. *Annual Review of Biophysics and Biomolecular Structure 33*, 269–295.

[203] Simons, K., and Toomre, D. (2000) Lipid rafts and signal transduction. *Nature Reviews Molecular Cell Biology 1*, 31–39.

[204] Staubach, S., and Hanisch, F.-G. (2011) Lipid rafts: signaling and sorting platforms of cells and their roles in cancer. *Expert Review of Proteomics 8*, 263–277.

[205] Dart, C. (2010) Symposium Review: Lipid microdomains and the regulation of ion channel function. *The Journal of Physiology 588*, 3169–3178.

[206] Lajoie, P., Goetz, J. G., Dennis, J. W., and Nabi, I. R. (2009) Lattices, rafts, and scaffolds: domain regulation of receptor signaling at the plasma membrane. *The Journal of Cell Biology 185*, 381–5.

[207] George, K. S., and Wu, S. (2012) Lipid raft: A floating island of death or survival. *Toxicology and Applied Pharmacology 259*, 311–9.

[208] McConnell, H. M., and Radhakrishnan, A. (2003) Condensed complexes of cholesterol and phospholipids. *Biochimica et Biophysica Acta (BBA) - Biomembranes 1610*, 159–173.

[209] Rog, T., and Pasenkiewicz-Gierula, M. (2001) Cholesterol effects on the phospholipid condensation and packing in the bilayer: a molecular simulation study. *FEBS Letters 502*, 68–71.

[210] Edholm, O., and Nagle, J. F. (2005) Areas of Molecules in Membranes Consisting of Mixtures. *Biophysical Journal 89*, 1827–1832.

[211] Alwarawrah, M., Dai, J., and Huang, J. (2010) A molecular view of the cholesterol condensing effect in DOPC lipid bilayers. *Journal of Physical Chemistry B 114*, 7516–7523.

[212] Radhakrishnan, A., and McConnell, H. M. (1999) Condensed complexes of cholesterol and phospholipids. *Biophysical Journal 77*, 1507–17.

[213] de Meyer, F., and Smit, B. (2009) Effect of cholesterol on the structure of a phospholipid bilayer. *Proceedings of the National Academy of Sciences of the United States of America 106*, 3654–8.

[214] Vist, M. R., and Davis, J. H. (1990) Phase equilibria of cholesterol/dipalmitoylphosphatidylcholine mixtures: deuterium nuclear magnetic resonance and differential scanning calorimetry. *Biochemistry 29*, 451–464.

[215] Blume, A., and Griffin, R. G. (1982) Carbon-13 and Deuterium Nuclear Magnetic Resonance Study of the Interaction of Cholesterol with Phosphatidylethanolamine1". *Biochemistry Biochemistry Phys. Rev. J. Magn. Resort. Biophys. J. Ann. N.Y. Acad. Sci. FEES Lett. J. Biol. Chem. Biochemistry Biochemistry Biochemistry 21*, 6230–6242.

[216] Radhakrishnan, A., Anderson, T. G., and McConnell, H. M. (2000) Condensed complexes, rafts, and the chemical activity of cholesterol in membranes. *Proceedings of the National Academy of Sciences 97*, 12422–12427.

[217] Ege, C., Ratajczak, M. K., Majewski, J., Kjaer, K., and Lee, K. Y. C. (2006) Evidence for lipid/cholesterol ordering in model lipid membranes. *Biophysical Journal 91*, L01–3.

[218] Huang, J. (2002) Exploration of molecular interactions in cholesterol superlattices: effect of multibody interactions. *Biophysical Journal 83*, 1014–25.

[219] Ali, M. R., Cheng, K. H., and Huang, J. (2007) Assess the nature of cholesterol-lipid interactions through the chemical potential of cholesterol in phosphatidylcholine bilayers. *Proceedings of the National Academy of Sciences 104*, 5372–5377.

[220] Khelashvili, G., Pabst, G., and Harries, D. (2010) Cholesterol Orientation and Tilt Modulus in DMPC Bilayers. *The Journal of Physical Chemistry B 114*, 7524–7534.

[221] Marsan, M., Muller, I., Ramos, C., Rodriguez, F., Dufourc, E., Czaplicki, J., and Milon, A. (1999) Cholesterol Orientation and Dynamics in Dimyristoylphosphatidylcholine Bilayers: A Solid State Deuterium NMR Analysis. *Biophysical Journal 76*, 351–359.

[222] Léonard, A., Escrive, C., Laguerre, M., Pebay-Peyroula, E., Néri, W., Pott, T., Katsaras, J., and Dufourc, E. J. (2001) Location of cholesterol in DMPC membranes. A comparative study by neutron diffraction and molecular mechanics simulation. *Langmuir 17*, 2019–2030.

[223] Harroun, T. A., Katsaras, J., and Wassall, S. R. (2008) Cholesterol Is Found To Reside in the Center of a Polyunsaturated Lipid Membrane. *Biochemistry 47*, 7090–7096.

[224] Kučerka, N., Perlmutter, J. D., Pan, J., Tristram-Nagle, S., Katsaras, J., and Sachs, J. N. (2008) The Effect of Cholesterol on Short- and Long-Chain Monounsaturated Lipid Bilayers as Determined by Molecular Dynamics Simulations and X-Ray Scattering. *Biophysical Journal 95*, 2792–2805.

[225] Xu, W., Wei, G., Su, H., Nordenskiöld, L., and Mu, Y. (2011) Effects of cholesterol on pore formation in lipid bilayers induced by human islet amyloid polypeptide fragments: A coarse-grained molecular dynamics study. *Physical Review E 84*, 51922.

[226] Ferraro, M., Masetti, M., Recanatini, M., Cavalli, A., and Bottegoni, G. (2016) Mapping Cholesterol Interaction Sites on Serotonin Transporter through Coarse-Grained Molecular Dynamics. *PLOS ONE 11*, e0166196.

[227] Prasanna, X., Sengupta, D., and Chattopadhyay, A. (2016) Cholesterol-dependent Conformational Plasticity in GPCR Dimers. *Scientific Reports 6*, 31858.

[228] Melo, M. N., Ingólfsson, H. I., and Marrink, S. J. (2015) Parameters for Martini sterols and hopanoids based on a virtual-site description. *The Journal of Chemical Physics 143*, 243152.

[229] Duncan, S. L., Dalal, I. S., and Larson, R. G. (2011) Molecular dynamics simulation of phase transitions in model lung surfactant monolayers. *Biochimica et Biophysica Acta (BBA) - Biomembranes 1808*, 2450–2465.

[230] Waheed, Q., Tjörnhammar, R., and Edholm, O. (2012) Phase Transitions in Coarse-Grained Lipid Bilayers Containing Cholesterol by Molecular Dynamics Simulations. *Biophysical Journal 103*, 2125–2133.

[231] Wang, Y., Gkeka, P., Fuchs, J. E., Liedl, K. R., and Cournia, Z. (2016) DPPC-cholesterol phase diagram using coarse-grained Molecular Dynamics simulations. *Biochimica et Biophysica Acta (BBA) - Biomembranes 1858*, 2846–2857.

[232] Daily, M. D., Olsen, B. N., Schlesinger, P. H., Ory, D. S., and Baker, N. A. (2014) Improved Coarse-Grained Modeling of Cholesterol-Containing Lipid Bilayers. *Journal of Chemical Theory and Computation 10*, 2137–2150.

[233] Hadley, K. R., and McCabe, C. (2010) A structurally relevant coarse-grained model for cholesterol. *Biophysical Journal*

[234] MacDermaid, C. M., Kashyap, H. K., DeVane, R. H., Shinoda, W., Klauda, J. B., Klein, M. L., and Fiorin, G. (2015) Molecular dynamics simulations of cholesterol-rich membranes using a coarse-grained force field for cyclic alkanes. *The Journal of Chemical Physics 143*, 12B625_1.

[235] Murtola, T., Falck, E., Karttunen, M., and Vattulainen, I. (2007) Coarse-grained model for phospholipid/cholesterol bilayer employing inverse Monte Carlo with thermodynamic constraints. *The Journal of chemical physics 126*, 075101.

[236] Khelashvili, G. A., and Scott, H. L. (2004) Combined Monte Carlo and molecular dynamics simulation of hydrated 18:0 sphingomyelin-cholesterol lipid bilayers. *The Journal of Chemical Physics 120*, 9841–9847.

[237] Murtola, T., Falck, E., Patra, M., Karttunen, M., and Vattulainen, I. (2004) Coarse-grained model for phospholipid/cholesterol bilayer. *The Journal of Chemical Physics 121*, 9156–9165.

[238] Huang, J., Buboltz, J. T., and Feigenson, G. W. (1999) Maximum solubility of cholesterol in phosphatidylcholine and phosphatidylethanolamine bilayers. *Biochimica et biophysica acta 1417*, 89–100.

[239] Hoover, W. G., and Holian, B. L. (1996) Kinetic moments method for the canonical ensemble distribution. *Physics Letters A 211*, 253–257.

[240] Parrinello, M., and Rahman, A. (1980) Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. *Physical Review Letters 45*, 1196–1199.

[241] Graham, J. A., Essex, J. W., and Khalid, S. (2017) PyCGTOOL: Automated Generation of Coarse-Grained Molecular Dynamics Models from Atomistic Trajectories. *Journal of Chemical Information and Modeling 57*, 650–656.

[242] Plimpton, S. (1995) Fast Parallel Algorithms for Short-Range Molecular Dynamics. *Journal of Computational Physics 117*, 1–19.

[243] Róg, T., Pasenkiewicz-Gierula, M., Vattulainen, I., and Karttunen, M. (2007) What Happens if Cholesterol Is Made Smoother: Importance of Methyl Substituents in Cholesterol Ring Structure on Phosphatidylcholine - Sterol Interaction. *Biophysical Journal 92*, 3346–3357.

[244] Pöyry, S., Róg, T., Karttunen, M., and Vattulainen, I. (2008) Significance of Cholesterol Methyl Groups. *Journal of Physical Chemistry B 112*, 2922–2929.

[245] Krause, M. R., Wang, M., Mydock-McGrane, L., Covey, D. F., Tejada, E., Almeida, P. F., and Regen, S. L. (2014) Eliminating the roughness in cholesterol's β-face: does it matter? *Langmuir : the ACS journal of surfaces and colloids 30*, 12114–8.

[246] Younglove, B. A., and Ely, J. F. (1987) Thermophysical Properties of Fluids. II. Methane, Ethane, Propane, Isobutane, and Normal Butane. *Journal of Physical and Chemical Reference Data 16*, 577–798.

[247] Disalvo, E., and Simon, S. *Permeability and Stability of Lipid Bilayers*; CRC Press, Boca Raton, FL, 1995.

[248] Craven, B. M. (1976) Crystal structure of cholesterol monohydrate. *Nature 260*, 727–729.

[249] Douliez, J.-P., Leonard, A., and Dufourc, E. J. (1995) Restatement of Order Parameters in Biomembranes: Calculation of C-C Bond Order

Parameters from C-D Quadrupolar Splittings. *Biophysical Journal 68*, 1727–1739.