UNIVERSITY OF SOUTHAMPTON
Faculty of Engineering, Science & Mathematics
School of Chemistry

Computer simulation of Lipids and DNA using a Coarse Grain Methodology

by

G D Chellapa

Thesis for the degree of Doctor of Philosophy

January 2010

Abstract

The nucleus of the eukaryotic cell contains a large pool of lipids together with structural proteins and genomic DNA. The project aim was to develop simple and robust lipid and DNA models that will allow for these complex molecules to be mixed together in order to elucidate the possible interactions. The large percentage of lipids found within the nucleus makes it likely that they exist in aggregates, although the actual role and structure in which they exist is unknown.

While there has not been substantial work done to model such interactions between lipids and DNA in order to better understand the interactions within the nucleus, a substantial body of work exists on lipid/DNA complexes in relation to gene therapy. These simulations in many cases however, are too simple and the structures formed are pre-imposed to a certain degree. Our model would attempt to simulate these interactions without such pre-imposed conditions relying solely on interactions between the particles to drive the structures being formed. A coarse graining approach in which several groups of atoms are subsumed into single interaction sites was deemed suitable given the complexity of modelling a mixture of DNA and lipids, together with the solvent and ion environment. In this regard new models of lipids, DNA, ion and solvent models were developed in a purpose built molecular dynamics package called *LANKA*-***L***ipid ***A***nd ***N***ucleic acid ***K***omputer ***A***lgorithm.

The lipids in the model are represented as polar ellipsoids and the solvent as spheres with dipoles embedded within them. The interactions between the lipids and solvent are modelled using the Gay Berne potential. The developed lipid model was able to self assemble into a stable bilayer phase and reproduce many bilayer properties of a liquid crystal phase. The model was then extended to capture some of the other lipid phases seen in nature, including lyotropic phase transitions.

A simple study of lipid mixtures has also been undertaken during this period. The importance of considering multicomponent lipid systems has increasingly been highlighted in the literature to make the lipid models more realistic. The developed lipid models are simple enough to extend and attempt to simulate the formation of lipid rafts and domain formation.

Simulation of DNA in the past has largely focused on atomistic studies. While these have proved valuable they do not consider the macroscopic length scales of the molecule. Simplified models trying to capture long length scales have had to compromise on the molecular level detail. Coarse grain models while trying to bridge the gap have also remained largely idealistic in nature.

To develop a simple DNA model, the same potentials used in the lipid model were extended to capture the DNA interactions. The developed DNA model was able to reproduce many of the structural parameters of the B form of DNA including its persistence length. The developed DNA model also included charge neutralising counter ions.

The work done in developing these highly coarse grain lipid, DNA, solvent and ion models have laid a firm foundation for the extension of this work to include DNA and lipid mixtures to better understand the role endonuclear lipids play in the cell.

Contents

List of Figures

List of Tables

Acknowledgement

For the LORD is good and his love endures forever; his faithfulness continues through all generations. Psalm 100:5 (NIV)

First and foremost I wish to thank my family, my parents in particular who have supported me through the last four years through their encouragement and constant prayers. I also wish to thank my sister and my grandmother for their love support and my extended family of uncles, aunts and cousins for their support.

Thank you to my supervisor Prof. J W. Essex and my advisor Prof. G S Attard for their support and advice throughout my PhD and all members of the Essex group especially to Ifat, Juan, Mario, Nadia and Patrick. I would like to especially than Mario for all his support during my PhD as well as all the system administrators for maintaining the computer systems to enable me to work effectively, especially Juan and Patrick.

I also wish to thank the Swathling Methodist Church for the many years of support and fellowship that has made my stay in Southampton truly a pleasant one.

Thank you to all my wonderful friends especially Wendy and Prabalini, who helped me tremendously throughout my time in Southampton. Thanks also to Gavin, Jilu, Ilker, Adrian, Luckni, Adela, Jin and many others too many to mention here who have made my stay in Southampton a memorable and enjoyable one.

I would also like to thank the Institute of Chemistry Ceylon, where I did my undergraduate studies for providing me with a good foundation that I was able to build upon during my PhD.

Funding for my research and living expenses received from the European Union through the Neonuclei project and from the British Government through the Overseas Research Students Awards Scheme (ORSAS) is gratefully acknowledged.

Finally thank you to my people, all Sri Lankans who have made me who I am today.

I lift up my eyes to the hills - where does my help come from? My help comes from the LORD, the Maker of heaven and earth.

Psalm 121:1-2 (NIV)

Chapter 1

Introduction and Overall Project Aims

The nucleus of the eukaryotic cell is the central controlling unit in the cell, controlling the growth and replication of the cell as well as responding to external stimuli. The nucleus is a highly heterogeneous system, consisting of genetic material in the form of DNA, proteins and substantial amounts of lipids [1]. The DNA wraps itself around the histone proteins and then is further coiled and supercoiled to form the chromosomes. The DNA structure within the nucleus is fairly well known but the role and the structure of the nuclear matrix especially the role of lipids within the eukaryotic nucleus is still unknown. The presence of a variety of lipids and in significant quantities (up to 10% by weight in water) certainly hints at the possible existence of large aggregates of lipids playing a vital role within the nucleus [1]. While these lipids have been identified, electron microscopic measurements have failed to show any membrane like structures in the naked (nuclear membrane removed) nucleus. Lipids such as phospholipids are not only synthesised in the endoplasmic reticulum but also in the nuclei; the endo nuclear lipids are synthesized by a separate mechanism within the nucleus and the concentrations of saturated and unsaturated lipids are tightly controlled. The tight control as well as the endonuclear synthesis certainly gives rise to the possibility of a biologically relevant role for lipids in the cell cycle [1][2][3][4]. The nuclear matrix was long ignored in the scientific community, instead focussing on the genetic material DNA and its structural hierarchy. However, increasingly it is becoming evident that changes in the nuclear matrix have a profound effect on the functioning of the (DNA) genetic material during the cell cycle [2][3] but one of the largely unanswered questions is the role that lipids play. Determining the role that they play is being attempted in various ways, including experimental and simulation methods. One of the possible methods of looking at these complex interactions is through the use of computer models. Several *in vitro* studies have indicated that lipids have a role to play in the expression of genes. It is also well known that the concentrations of endonuclear lipids varies with the cell cycle again indicating a possible regulatory role in the life cycle of the cell [1][2][3][4].

The work undertaken in this study was to develop very simple computer models of lipids and DNA, test their ability to reproduce certain basic properties of these macromolecules. It is hoped that these developed models then can then be used to simulate lipid-DNA mixtures in the hope that they may shed light on possible interactions taking place within the nucleus between lipids and DNA.

One of the methodologies used in computer simulation is known as Molecular Dynamics (MD) which is based on classical mechanical principles and is applicable to large systems of the kind being modelled in this project. A short introduction to the MD methodology is presented in Chapter 2.

The interaction between lipids and DNA have been studied in relation to gene transport using both theoretical/computational and experimental methods for several years; these attempt to better understand interactions between cationic lipids, neutral lipids and DNA and their possible therapeutic use [5][6][7][8][9].

Lipids have a variety of roles within a living cell, from fuel molecules, highly concentrated energy stores, signal molecules and components of membranes. All lipids, whatever role they play, possess a common structural theme; they are amphiphilic in nature. The wide variety of structures that lipids form is a direct consequence of this amphiphilic behaviour. The van der Waals attractive forces between the lipid tail and the hydrogen bonding and electrostatic interactions between the water and lipid head leads to many lipid phases.

DNA is primarily known as the carrier of genetic information and controls the overall functioning of the cell. The DNA molecules are made up of nucleotides; they serve as the building blocks of DNA [10]. Each nucleotide consists of a deoxyribose sugar, phosphate group and a nitrogen base. The sugar and phosphate group serve as the backbone of the DNA molecule while the nitrogen bases may vary between adenine (A), thymine (T), cytosine (C) and guanine (G). The nucleotides are connected together through the phosphate groups to form a long strand. Two such strands running anti-parallel to each other form the DNA molecule; the strands are held together by hydrogen bonds that form between the nitrogen bases. Two hydrogen bonds form between adenine and thymine and three between cytosine and guanine. The resulting flexible ladder-like structure wraps around an imaginary central axis to form the classical DNA double helix first described by Watson and Crick [11]. The spaces formed between the helical backbone and the imaginary cylinder that encloses the

DNA are termed the major and minor grooves; they have different dimensions because of the sugar-based asymmetry with respect to the base-pair plane [10].

There are three main canonical forms of DNA, although several others have also been identified. These canonical forms are defined based on several structural factors such as the direction of the helix turn (right or left handed turn), dimensions of the major and minor grooves, number of nucleotide pairs per helix turn and spacing between the nucleotide pairs (rise). The most dominant form under physiological conditions is the right handed B-DNA [12]. DNA has many levels of structural hierarchy from length scales of nanometres for several base pairs to micrometres for several thousand base pairs. The breadth of DNA is microscopic while its length is macroscopic; this ensures that DNA molecules needs to be studied using several different models in order to fully understand their functions. The study of DNA's rich hierarchy of organisation levels is particularly challenging to modellers [5]. DNA modellers will deal with large spatial scales (thousands of angstroms) as well as temporal scales (picoseconds to minutes and more) [5]. Understanding the different levels of structure is important for understanding fully the role DNA plays within a living cell. Hydration effects also play an important role in stabilising a particular helical conformation. Both hydrogen bonding and base stacking are considered intrinsic factors for helical stability [5]. DNA is highly charged and interacts strongly with its solvent environment and solutes over long distances [13].

Lipids and DNA have been studied by several models ranging from all atom MD models to continuum models [5][12][13][14][15][16][17][18][19]. There has been considerable progress in all atom molecular dynamics (MD) simulation techniques looking at small strands of DNA and small patches of lipid bilayers. However there is still a gap to be bridged between all atom and macroscopic representations of nucleic acids and lipids. Bridging this gap are the coarse grain models, making use of simplified representations of lipids and DNA, where the finer details of the molecules are lost although the underlying details are retained [9][14][15][16][20][21][22][23].

A brief overview of the biological background to lipids and DNA primarily focussing on experimental observations is given Chapter 3 while Chapter 4 reviews some of the published coarse grain lipid and DNA models.

Despite the existence of several DNA models, of which most are all atom models, there is still a lack of understanding of the large-scale DNA structure and how it relates to its function. In this project the aim is to develop a simplified representation of lipids and DNA, using coarse grain dynamics, and study the interactions that arise as a result of forces between the DNA and lipids. This can only be done if the underlying modelling principles in the developed models are based on the correct biological and physical properties. To access the behaviour of large-scale assemblies of lipids and DNA mixtures the model needs to be highly simplified.

The importance of solvent has been highlighted above and this has been one of the primary shortcomings of the existing models. In order to solve this issue a dipolar solvent with the correct dielectric behaviour was required. However, representing the solvent in a simulation is a double edged sword. While essential to capture the properties of lipids and DNA, it usually turns out to be a massive bottle neck in simulations since the computational effort is approximately proportional to the square of the number of particles in the system [24]. One way of overcoming this issue is the use of coarse grain water models. Such a model has been developed in this work, that correctly captures the density, dielectric constant, diffusion and liquid-like structure akin to real liquid water at room temperature.

To make the lipid model simple and fast the lipid molecules need to be simplified. The lipid model developed was simplified to the very essential features of a lipid which is its hydrophilic head and hydrophobic tail. In order to capture the hydrophobic tail using a single site an ellipsoidal tail was deemed appropriate. The interactions between the tail sites are captured using the Gay Berne potential [25], a potential that has seen wide spread use in the liquid crystal community [26][27][28][29][30][31][32]. Lipids are nature's own liquid crystals, thereby justifying the use of this potential to capture the many different phases of lipids. The lipid head groups are represented by dipolar spheres. The dipole embedded in the head group allows us to capture the electrostatic nature of the hydrophilic head group. The interactions between the dipolar head were captured using the Lennard Jones and dipolar potentials. The developed lipid model forms bilayers spontaneously and remains in a liquid crystalline state. The structural and dynamic properties of the bilayer are in good agreement with experimental measurements of phosphatidylcholine lipids.

The rich mesomorphism of lipids is driven by temperature and concentration but the effective shape of the lipids also has a role to play in this behaviour. Using the packing theory as a guide the developed bilayer forming lipid has been used to capture other phases such as micelles, hexagonal, inverse micelles, inverse hexagonal and cubic phases simply by adjustment of relative sizes of the heads and tails [33][34][35]. Not only are the different phases captured using this method, using the same set of parameters the model is able to show lyotropic behaviour based on the concentration of water/lipid (variation of lipid phases with varying solvent concentration) [36].

To develop a new DNA model, but keeping the development process simple as well as to ensure seamless integration with the developed models of lipids and water, a DNA model was developed based on the Gay Berne potential [25][28]. However, instead of using prolate like objects that were used in the lipids, the potential is used to represent oblate like objects. These oblate objects represent the two nitrogen base pairs in double stranded DNA. The natural tendency of these oblate objects to stack on top of each other is uniquely used in the developed DNA model to provide stability. Nature uses stacking interactions between bases pairs for stabilising the DNA and in fact it is deemed more important than hydrogen bonding for providing stability to the DNA [21][37]. The developed DNA model is able to capture many of the structural properties of DNA including the persistence length in good agreement with experimental measurements.

The development and parameterisation of the lipid and DNA models as well as the challenges faced and methods used to overcome them are highlighted and discussed in depth in Chapter 5.

In living systems biological molecules such as lipids and DNA exist in mixtures. Membranes are collection of different types of lipids each with its own physical, chemical and mechanical properties. By varying the ratios of these different lipids, nature is able to fine-tune the functions of each membrane. The developed lipid model is simple and efficient enough to look at the effect of different lipids on the properties of the bilayer. Chapter 6 focuses on the effect of lipid mixtures on bilayer structural and mechanical properties.

Chapter 7 highlights the possible future applications of the developed models as well as areas in which modifications and improvements can be made to better represent the underlying molecules and improve the agreement with experimental observations. These include making the dipole vector in the head group, currently fixed perpendicular to the long axis of the ellipsoid, an additional degree of freedom to better capture the reality of the fluctuating head group in lipids. Representing the hydrocarbon tail by two ellipsoids will allow for the splaying of the tails with increasing temperature and capture the variation of different structural properties with temperature that arise as a result.

Currently the DNA model is incapable of being utilised to study the melting dynamics of DNA; however by representing the two nitrogen base pairs separately as two disks and capturing the hydrogen bonding by use of Morse type potentials the melting and complimentary base pairing may be captured.

Chapter 2

Introduction to Molecular Dynamics Simulation

Molecular Dynamics(MD) simulations compute the motions of individual molecules in models of solids, liquids and gases. The key idea here is *motion*, which describes how positions, velocities and orientations change with time. In effect MD constitutes a motion picture of the way molecules behave as they move [38]. This chapter gives a brief overview of Molecular Dynamics methods in relation to the project being undertaken.

2.1 Describing the motion of molecules within the system

Molecular dynamics assumes that the motion of molecules may be described in terms of *Classical Mechanics*, where the total energy of a system of N molecules is given as the sum of kinetic and potential energy functions of the set of coordinates $\mathbf{q_i}$ and momenta $\mathbf{p_i}$ of each molecule i [24]. The total energy E may be written as the sum of the potential U and kinetic K energies as given in Equation 2.1.

$$E(\mathbf{q}, \mathbf{p}) = K(\mathbf{p}) + U(\mathbf{q}) \tag{2.1}$$

2.2 Potential Energy

Differentiating the potential energy U_i on particle i gives the force on molecule i. If the force acting on the molecule at each instance is known, then the motion of the molecule may conveniently be followed giving the trajectory of that particle, using Equation 2.2, where $\mathbf{f}_i(t)$ represents the force on particle i and m_i and $\ddot{\mathbf{r_i}}(t)$ its mass and acceleration respectively[39].

$$\mathbf{f}_i(t) = m_i \ddot{\mathbf{r_i}}(t) = -\nabla_{\mathbf{r}} U \tag{2.2}$$

A realistic interaction between the molecules is known to possess both an attractive and repulsive component due to the fact that on the one hand solids and liquids have the property of cohesion, but at the same time do not collapse indefinitely to a point of singularity under the action of these forces. Soft intermolecular potentials attempt to model these interactions by comprising an attractive outer and repulsive inner region.

2.3 Interaction Potentials

2.3.1 The Lennard-Jones Potential

The Lennard Jones (LJ) pair potential is commonly used to model interactions between spherical particles. It has both an attractive element and a repulsive element. The attractive part comes into play at large intermolecular separations and is used to model dispersive interactions arising from correlation between electron clouds surrounding the molecules. At shorter distances a negative well represents the cohesion in condensed phases and finally a sharply rising repulsive element at shorter distances corresponds to non-bonded overlap between electron clouds. The LJ pair-potential is characterised by a long-range attractive tail of the form $-1/r^6$, a negative well of depth ϵ, and a steeply rising repulsive wall at distances less than $r \sim \sigma$.

For a pair of atoms i and j located at $\mathbf{r_i}$ and $\mathbf{r_j}$ the potential energy is given by Equation 2.3 where $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ The parameter ϵ governs the strength of the interactions while σ defines a length scale.

$$\mathcal{U}(r_{ij}) = \begin{cases} 4\epsilon[(\sigma/r_{ij})^{12} - (\sigma/r_{ij})^6] & \text{if } r_{ij} < r_c \\ 0 & \text{if } r_{ij} \geq r_c \end{cases} \tag{2.3}$$

The LJ potential is an idealised potential suitable to model spherically symmetric molecules such as low density gases given the appropriate choice of parameters for σ and ϵ [40].

2.3.2 The Gay-Berne Potential

There have been several attempts to effectively capture interactions between non spherical (anisotropic) sites using a single interaction potential. Anisotropic potentials depend not only on the separation between sites but also on the orientation of the two interacting particles. One of the best known potentials is the Gay-Berne potential [25] and its predecessor the Berne-Pechukas potential [41].

The Gay-Berne potential is an anisotropic potential of the form:

$$\mathcal{U}(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \mathbf{r}) = 4\epsilon(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}}) \left[\left(\frac{\sigma_0}{r - \sigma(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}}) + \sigma_0} \right)^{12} - \left(\frac{\sigma_0}{r - \sigma(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}}) + \sigma_0} \right)^{6} \right] \tag{2.4}$$

The potential energy is given by $\mathcal{U}(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \mathbf{r})$ and σ_0 represents the diameter of the particle. The orientations of the two particles are given by the unit vectors $\hat{\mathbf{u}}_i$ and $\hat{\mathbf{u}}_j$, and $\mathbf{r} \equiv r\hat{\mathbf{r}}$ being the vector joining the centres of mass. The Gay-Berne potential can simulate ellipsoids with varying lengths and breadths and when the length and breadth are made equal collapses to the standard LJ potential; hence the Gay-Berne potential is a generalised LJ potential. The energy well depth ϵ and interaction parameter σ are dependent not only on the separation of the two particles but also on their orientation. Derivation of the forces and torques as well as definitions for σ and ϵ for the Gay-Berne potential are given in Appendix B.

2.3.3 Generalisation of the Gay-Berne Potential

The Gay-Berne potential was generalised to capture the interactions between non-equivalent uniaxial particles by modifying the range parameter σ in the GB potential [42][43]. The generalised version of the range and well depth parameters are given below:

$$\sigma(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}}) = \sigma_{\mathrm{GB}} \left\{ 1 - \left[\frac{\chi\alpha^2(\hat{\mathbf{u}}_i \cdot \widehat{\mathbf{r}_{ij}})^2 + \chi\alpha^{-2}(\hat{\mathbf{u}}_j \cdot \widehat{\mathbf{r}_{ij}})^2 - 2\chi^2(\hat{\mathbf{u}}_i \cdot \widehat{\mathbf{r}_{ij}})(\hat{\mathbf{u}}_j \cdot \widehat{\mathbf{r}_{ij}})(\hat{\mathbf{u}}_i \cdot \hat{\mathbf{u}}_j)}{1 - \chi^2(\hat{\mathbf{u}}_i \cdot \hat{\mathbf{u}}_j)^2} \right] \right\}^{-\frac{1}{2}} \tag{2.5}$$

$$\epsilon(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}}_{ij}) = \epsilon_{GB}\epsilon_1^{\nu}(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j)\epsilon_2^{\mu}(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}}_{ij}) \tag{2.6}$$

$$\epsilon_1(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j) = \left[1 - \chi^2(\hat{\mathbf{u}}_i \cdot \hat{\mathbf{u}}_j)^2\right]^{-\frac{1}{2}} \tag{2.7}$$

$$\epsilon_2(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}}_{ij}) = 1 - \left\{ \frac{\chi^{'}\alpha^{'2}(\hat{\mathbf{u}}_i \cdot \hat{\mathbf{r}}_{ij})^2 + \chi^{'}\alpha^{'-2}(\hat{\mathbf{u}}_j \cdot \widehat{\mathbf{r}_{ij}})^2 - 2\chi^{'2}(\hat{\mathbf{u}}_i \cdot \widehat{\mathbf{r}_{ij}})(\hat{\mathbf{u}}_j \cdot \widehat{\mathbf{r}_{ij}})(\hat{\mathbf{u}}_i \cdot \hat{\mathbf{u}}_j)}{1 - \chi^{'2}(\hat{\mathbf{u}}_i \cdot \hat{\mathbf{u}}_j)^2} \right\} \tag{2.8}$$

Where the definitions for σ_{GB}, χ, α^2, $\chi^{'}$ and $\alpha^{'2}$ are given by Equations 2.9-2.13 where l and d represent the length and breadth of the particles. The notation ϵ_{GB} is the well depth of the cross configuration of two Gay-Berne particles and ϵ_E and ϵ_S the end-to-end and side-by-side well depths. For unlike particles Lorentz-Berthelot mixing rules were used [24]. The values for μ and ν were set to 2 and 1 respectively as in the original GB model [25].

$$\sigma_{GB} = \sqrt{d_i^2 + d_j^2} \tag{2.9}$$

$$\chi = \left[\frac{(l_i^2 - d_i^2)(l_j^2 - d_j^2)}{(l_j^2 + d_i^2)(l_i^2 + d_j^2)}\right]^{\frac{1}{2}} \tag{2.10}$$

$$\alpha^2 = \left[\frac{(l_i^2 - d_i^2)(l_j^2 + d_i^2)}{(l_j^2 + d_j^2)(l_i^2 + d_j^2)}\right]^{\frac{1}{2}} \tag{2.11}$$

$$\chi' = \frac{1 - \left(\frac{\epsilon_E}{\epsilon_S}\right)^{\frac{1}{\mu}}}{1 + \left(\frac{\epsilon_E}{\epsilon_S}\right)^{\frac{1}{\mu}}} \tag{2.12}$$

$$\alpha'^2 = \left[1 + \left(\frac{\epsilon_E}{\epsilon_S}\right)^{\frac{1}{\mu}}\right]^{-1} \tag{2.13}$$

If one of the interacting particles is made spherical, for example $l_i = d_i = d$, then both χ and α tend to zero, the range parameter, nevertheless, remains finite in this limit and tends smoothly to equation 2.14 enabling the use of this potential to calculate interactions between spheres and ellipsoids.

$$\sigma(\hat{\mathbf{r}}_{ij}, \hat{\mathbf{u}}_j) = \sigma_{\mathrm{GBLJ}}[1 - \chi_{\mathrm{GBLJ}}(\hat{\mathbf{r}}_{ij} \cdot \hat{\mathbf{u}}_j)^2]^{-1/2}\,, \tag{2.14}$$

where the ratio χ_{GBLJ} is given by equation by equation 2.15 [42].

$$\chi_{\mathrm{GBLJ}} = \frac{\sigma_{\mathrm{GB}}^2(\kappa_{\mathrm{GB}}^2 - 1)}{(\kappa_{\mathrm{GB}}\sigma_{\mathrm{GB}})^2 + \sigma_{\mathrm{LJ}}^2}\,. \tag{2.15}$$

The mixed interactions were considered for interactions between the lipid and solvent particles, inter-lipid interactions, as well as interactions of the DNA with lipids, DNA with solvent, and inter DNA interactions. The values for σ_0 and ϵ_0 for the Gay-Berne-Lennard Jones interactions were calculated using the Lorentz-Berthelot mixing rules in some instances [24] and were scaled in other instances during parameterisation.

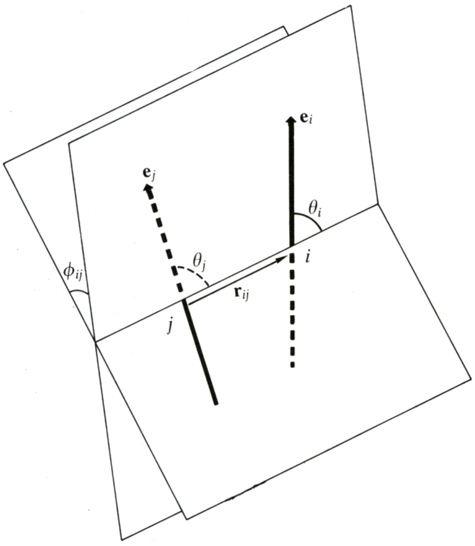

Figure 2.1: The relative orientation of two linear rigid bodies [24].

2.3.4 Dipole-Dipole Potential

The configurations of a pair of molecules is shown in 2.1, [24]. The centres of the molecule are separated by the vector $\mathbf{r}_{ij}$. θ_i and θ_j are the angles between $\mathbf{r}_{ij}$ and the unit vectors directed along the molecular axes, $\mathbf{e}_i$ and $\mathbf{e}_j$. ϕ_{ij} is the angle between the plane containing $\mathbf{e}_i$ and $\mathbf{r}_{ij}$. The angle γ_{ij} is defined as given in equation 2.16. If the molecules i and j have a dipole moment of magnitude μ, then the interaction potential is given by equation 2.17

$$\cos\gamma_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j = \cos\theta_i\cos\theta_j + \sin\theta_i\sin\theta_j\cos\phi_{ij} \tag{2.16}$$

$$u_{ij}^{\mu\mu}(\hat{\mathbf{e}}_i, \hat{\mathbf{e}}_j, \hat{\mathbf{r}}) = \frac{\mu^2}{r^3}(\cos\gamma_{ij} - 3\cos\theta_i\cos\theta_j) \tag{2.17}$$

2.3.5 Monopole-Monopole Potential

If the particles bear charges of $Q1$ and $Q2$ and are a distance r apart their potential energy is given by Equation 2.18 where ϵ_0 is the electric permitvity of vacuum.

$$u_{ij}^{Q_1Q_2}(r) = \frac{Q_1Q_2}{4\pi\epsilon_0 r} \tag{2.18}$$

2.3.6 Monopole-Dipole Potential

If one of particles bears a charges of Q and the other a dipole moment of μ and they are a distance r apart their potential energy is given by Equation 2.19, where $\hat{\mathbf{e}_\mu}$ is the orientation vector of the dipole and $\mathbf{r}_{ij}$ the inter-particle distance vector.

$$u_{ij}^{Q\mu}(\hat{\mathbf{e}}_\mu, \hat{\mathbf{r}}) = \frac{Q\mu}{4\pi\epsilon_0 r^3}(\hat{\mathbf{e}}_\mu \cdot \hat{\mathbf{r}}) \tag{2.19}$$

2.3.7 Bond Potential

The bonding between two sites in the same molecule is established using the Hooke potential as given by Equation 2.20 where l and l_0 are the actual and reference bond lengths while k is the Hooke's force constant that determines the bonds resistance to displacement. The bond potential has been used in the development of the DNA model. It links the DNA backbone sites together.

$$u_{ij}(l) = \frac{1}{2}k\,[l - l_0]^2 \tag{2.20}$$

2.3.8 Angle Potential

A cosine angle potential is given by Equation 2.21 where $\cos\theta$ and $\cos\theta_0$ are the actual and reference angles and $k^{'}$ the force constant. The angle potential is implemented in the DNA model between adjacent three backbone sites, helping maintain the correct helical twist.

$$u_{ijk}(\cos\theta) = \frac{1}{2}k^{'}\,[\cos\theta - \cos\theta_0]^2 \tag{2.21}$$

2.3.9 Torsion Potential

A torsion angle potential was utilised in this work as given by Equation 2.22 based on the work by Ryckaert and Bellemans [44]. The coefficients $k_w^{''}$ define the potential and in this case it is truncated at $w = 6$. The coefficients were obtained from the original work by Ryckaert and Bellemans [44], but the signs changed in $LANKA$ to favour the *cis* configuration instead of the *trans* as in the original. The coefficients used are as follows -1.000, -1.310, 1.414, 0.330, -2.828, 3.394 [24][40]. The torsion potential is found in the DNA model, helping to maintain the twist in the DNA as uni-directional .

$$u_{ijkl}(\cos\psi) = \sum_{w\geq 0} k_w^{''}\cos^w\psi \tag{2.22}$$

2.4 Kinetic Energy

The kinetic energy of a point particle takes the form

$$K = \sum_{i=1}^{N}\sum_{\alpha}\frac{p_{i\alpha}^2}{2m_i} \tag{2.23}$$

where m_i is the mass of the particle and the index α runs over the different (x, y, z) components of the momentum $p_{i\alpha}$ [24]. It is possible to obtain the temperature of the system from the kinetic energy based on the equipartition theorem [24][38][39][45].

2.5 Computing Properties from Molecular Dynamics

Simulations allow us to look at a system in terms of individual atoms/particles/sites and specify each site by its position and momentum. If we specify the position and momentum of particle i by χ_i then the instantaneous property such as A is specified $A(\chi_1, \chi_2....\chi_N)$ where N is the number of particles/atoms in the system. However, the thermodynamic state of the system is not based on the microscopic state of individual sites but on macroscopic properties of temperature, pressure and number of particles. Statistical mechanics helps to connect the microscopic state of individual sites to the experimentally measured properties. If we assume that the position and momentum of particle i is a single coordinate in a multidimensional phase space where if N is the number of sites in the system then the phase space is said to be $6N$ dimensional, with each particle associated with three momentum variables and three position variables. If a single state in the phase space is represented by Γ then the property of A at this particular point at a particular instance may be represented by $A(\Gamma)$. With time as the system evolve so does the value of $A(\Gamma)$. What is observed in experiments is the time average of the instantaneous values of $A(\Gamma)$ as given in Equation 2.24 [24][38][46].

$$A_{obs} =< A >_{time}= \lim_{x \to \infty} \frac{1}{t_{obs}} \int_0^{t_{obs}} A(\Gamma(t))dt \tag{2.24}$$

As seen in Equation 2.24, the calculation of the average property of A requires the dynamic change in A with time. This is measured in molecular dynamics by following the trajectories of the particles. The trajectories are obtained by solving Newton's equations of motion. However, there are several limitations in the use of simulations to obtain average properties; primarily the system sizes are limited to a few hundreds to thousands of particles, far fewer than are normally measured in experiments, which are in the range of 10^{23} sites. This deficiency can be overcome by the use of periodic boundary conditions as described below. The other drawback is that these measurements are made over a finite time which is generally found to be adequate for most measurables [24].

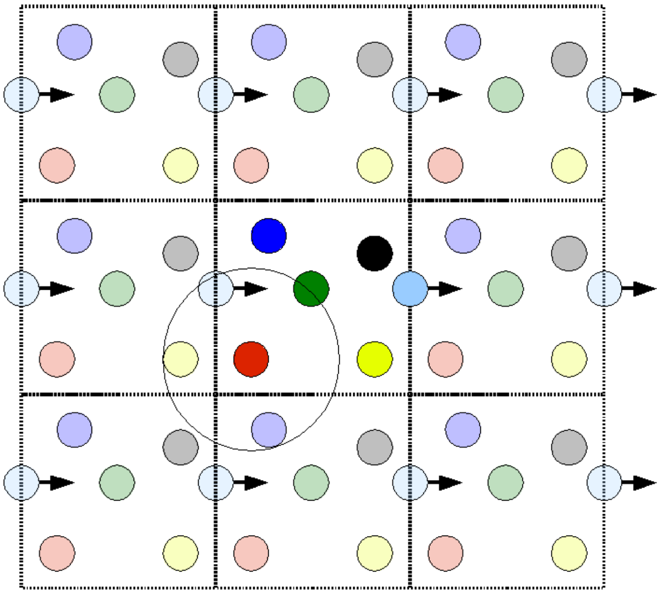

Figure 2.2: The periodic box with the central simulation box in dark and the minimum image convention (MIC) [5][38].

2.6 Periodic Boundary Conditions

Simulation of system sizes often encountered in real experiments is intractable; the system sizes are often limited to several hundreds or thousands of particles. For a lattice system of 10 × 10 × 10 lattice sites, approximately 50% of the system is on the surface and what is being simulated is a interfacial system rather than a bulk medium. This constraint placed on molecular dynamics is solved by the use of periodic boundary conditions where the central simulation box (or other space filling shape) is surrounded by exact copies of itself which in turn are surrounded by more duplicate copies. When a particle moves out of the central box its identical image replaces it through the opposite face of the box. A two dimensional version of the periodic boundary condition is depicted in Figure 2.2 [5][38].

One possible problem with periodic boundary conditions is that a molecule may ”see” its own image and will interact with itself. However, this is usually not the case as most interactions are short ranged and are cut at distances less than half the shortest box length of the simulation box. By using a cutoff and application of the minimum image convention(MIC) it can be ensured that the system interacts with only one of the sites. While the use of a cutoff may at first glance seem inappropriate, it is borne out by the fact that usually the largest contribution to the interaction energy is made by close molecules.

One of the possible problems with the use of periodic boundary conditions is that density fluctuations of the order of the box dimensions and larger are suppressed, this is a problem in weak transitions as well as situations where the box dimensions may force phases, especially for liquid crystal systems. The problem is exacerbated in small systems [24][38][39].

2.7 Calculation of Trajectories

The main objective of MD simulations is to obtain the trajectories of the molecule allowing its motion to be followed. Any non-spherical particle contains both translational and rotational motion and both need to be calculated to obtain the trajectory of the molecule.

2.7.1 Translational Motion

The translation of the centre of mass of the molecule is governed by Newton's equations of motion and is obtained by differentiating the potential of the molecule at a given instance [45].

Calculation of Force

If the interaction between two particles i, j is considered the potential energy between the two is determined by the inter particle distance. It is possible to calculate the potential acting on i as a result of j using one of the many potential energy equations and from it the force acting on i due to j. Applying Newton's third law the force acting on j due to i is the equal and opposite to the force acting on i due to j. The force on the molecule i due to j is

$$\mathbf{f}_{ij} = -\nabla_{r_{ij}} U_{ij} \tag{2.25}$$

Integration of the linear motion

Several integration algorithms exist, the Verlet algorithm being by far the most popular and widely used. This is due to its ease of implementation and symplectic and time reversible nature. The Verlet algorithm is excellent for the use of long time steps and since it conserves energy well and requires $3N$storage, N being the number of particles in the system. Alternate integration methods include the Leap Frog algorithm and the Velocity Verlet both of which are modifications to the original Verlet scheme. The Velocity Verlet take the following form, where, δt is the time step $\mathbf{r}$ and $\mathbf{v}$ represent the position and velocity of the particle at different time intervals[24][39].

$$\mathbf{v}(t + \delta t/2) = \mathbf{v}(t) + \delta t \cdot \mathbf{f}(t)/2m \tag{2.26}$$

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \delta t \mathbf{v}(t + \delta t/2) \tag{2.27}$$

$$\mathbf{v}(t+\delta t) = \mathbf{v}(t+\delta t/2) + \delta t \cdot \mathbf{f}(t+\delta t)/2m \tag{2.28}$$

2.7.2 Rotational Motion

Rotation of a molecule is handled by considering the torque τ about its centre of mass. The torque enters the rotational equations of motion similar to the way that forces enter the translational equations of motion[40][24].

Calculation of Torques

The torque on molecule i due to molecule j is given by 2.29 and the rotational motion is handled by Euler's equations 2.30, where $\mathbf{T}_{ij}$ represents the torque and $\mathbf{e_i}$ the unit vector along the principal axes of the molecule in equation 2.29. $\mathbf{I}$ and ω represent the inertia tensor and angular velocity respectively in Equation 2.30 [24][40][45][47].

$$\mathbf{T}_{ij} = -\mathbf{e_i}\nabla_{e_i}U_{ij} \tag{2.29}$$

$$\mathbf{I} \cdot \dot{\omega} - \omega \times \mathbf{I} \cdot \omega = \mathbf{T} \tag{2.30}$$

Integration of the rotational motion

The handling of rotational motion is more complicated and there have been several schemes proposed, the most popular being the use of quaternions [48]. Although widely used, they perform poorly in terms of energy conservation. An alternate method proposed by Dullweber et al. [47] had been already implemented successfully in several studies [43][49][50][51] and this was utilised within this project due to its stability and since it allows for a large time step [47]. It is based on propagating a full rotation matrix from one step to the other, unlike the quaternion based method.

As a first step the torques are calculated, as a function of the rotation matrix $\mathbf{Q}$.The orientation of the molecule is described by an orientation matrix $\mathbf{Q}$ consisting of direction cosines between the body fixed frame and the laboratory global axis [40][47]. Once the torques are calculated for all the molecules in the system at time (t) the angular momentum L is propagated from time (t) to ($t + \frac{\delta}{2}t$) as given by Equation2.31.

$$\mathbf{L}^{n+\frac{1}{2}} = \mathbf{L}^{n} + \frac{1}{2}\delta t \mathbf{T}^{n} \tag{2.31}$$

Then consecutive rotations $\mathbf{R}_1$ to $\mathbf{R}_5$ given by Equations 2.32-2.36 are applied to all angular momenta and used to update the orientation matrices for the full time ($\mathbf{Q}^{n+1}$) step from t_n to t_{n+1} where I_1 I_2 and I_3 are the elements of the diagonal inertia tensor of a molecule and L_1 L_2 and L_3 the components of $\mathbf{L}$ in the body axis system[47].

$$\mathbf{R}_1 := \mathbf{R}_x\left(\frac{1}{2}\delta t\frac{L_1}{I_1}\right), \qquad \mathbf{L} = \mathbf{R}_1\mathbf{L}, \qquad \mathbf{Q} = \mathbf{Q}\mathbf{R}_1^T; \tag{2.32}$$

$$\mathbf{R}_2 := \mathbf{R}_y\left(\frac{1}{2}\delta t\frac{L_2}{I_2}\right), \qquad \mathbf{L} = \mathbf{R}_2\mathbf{L}, \qquad \mathbf{Q} = \mathbf{Q}\mathbf{R}_2^T; \tag{2.33}$$

$$\mathbf{R}_3 := \mathbf{R}_z\left(\frac{1}{2}\delta t\frac{L_3}{I_3}\right), \qquad \mathbf{L} = \mathbf{R}_3\mathbf{L}, \qquad \mathbf{Q} = \mathbf{Q}\mathbf{R}_3^T; \tag{2.34}$$

$$\mathbf{R}_4 := \mathbf{R}_y\left(\frac{1}{2}\delta t\frac{L_2}{I_2}\right), \qquad \mathbf{L} = \mathbf{R}_4\mathbf{L}, \qquad \mathbf{Q} = \mathbf{Q}\mathbf{R}_4^T; \tag{2.35}$$

$$\mathbf{R}_5 := \mathbf{R}_x\left(\frac{1}{2}\delta t\frac{L_1}{I_1}\right), \qquad \mathbf{L} = \mathbf{R}_5\mathbf{L}, \qquad \mathbf{Q} = \mathbf{Q}\mathbf{R}_5^T; \tag{2.36}$$

Once the orientation matrix ($\mathbf{Q}^{n+1}$) is obtained the calculation of torques $\mathbf{T}^{n+1}$ at time $t+\delta t$ is done and propagate the angular momenta another half step completing one integration step as given in Equation 2.37 [47]. For symmetric rigid bodies, the equations simplify to Equations 2.32-2.34 [47].

$$\mathbf{L}^{n+1} = \mathbf{L}^{n+\frac{1}{2}} + \frac{1}{2}\delta t\mathbf{T}^{n+1} \tag{2.37}$$

2.8 Handling Long Range Interactions

Long range interactions are interactions that decay no faster than r^{-n} where n is the dimensionality of the system. Monopole-monopole, dipole-monopole and dipole-dipole interactions are some of the interactions showing this behaviour. Several ways of handling this problem have been formulated, the most popular being the Ewald summation and its variants which look at the interaction of an ion with its all its periodic images [24][39]. This method is, however, computationally expensive and tends to overemphasise the periodic nature of the system [24][39]. Other alternatives include the reaction field method that assumes that the interaction from molecules beyond a cutoff distance can be handled in an average way using macroscopic electrostatics. However, these methods usually require an initial starting value for the permittivity beyond the cutoff (more on this method in section 2.8.1). Other alternatives include brute force cutoff where the interactions beyond the cutoff are simply ignored. However, the movement of ions/dipoles in and out of the cutoff sphere results in the pumping of energy and non Newtonian trajectories [52], which show up as artificial effects at the cutoff distance.

One of the methods to relieve the problems with the brute force cutoff is to use switching functions that taper off the interactions between a lower and upper cutoff by multiplying by a polynomial function. Usually the upper and lower cutoffs are taken as close as possible to each other so that the potential is unaffected up to the lower cutoff, and then tapers to zero at the upper cutoff. Another alternative is to use a shifting function whereby the entire potential is shifted from r_{ij}=0 until the cutoff. However the true potential is not computed. Depending on the system under study a switching function or a shifting function maybe utilised [5][24][39].

2.8.1 Reaction Field Method

The reaction field method is an alternative method to handle long range electrostatic interactions but relatively less costly than the Ewald summation methods [24][39][53] and gives similar results for dipolar systems. The reaction field method had a long history of being used in solvent systems using both MC and MD methods [54][55][56][57][58][59][60][61][62].

In the reaction field method the field on the dipole consists of two parts: the first is a short range contribution from the molecules situated within a cutoff sphere R and the second arises from molecules outside R which are considered to form a dielectric continuum ϵ_{rf} producing a reaction field within the cavity. The size of the reaction field acting on the molecule i is proportional to the moment of the cavity surrounding i and is given by Equation 2.38 where the summation extends over the molecules in the cavity including i and r_c is the radius of the cavity. The contribution of the reaction field to the energy is given by Equation 2.39 and the torque by Equation 2.40. To prevent discontinuous jumps in energy, forces and torques when the number of molecules inside the cavity changes, a cubic switching function is employed [24][53][54][55] (See Equations 2.47-2.49).

The reaction field method can be used to calculate the dielectric constant of a system. In calculating the dielectric constant of a system the dielectric constant of the medium outside the cutoff is usually assumed to be ∞ although faster convergence of the dielectric constant would be achieved if ϵ_{rf} was close to ϵ_0, the dielectric constant of the system under study. However previous simulation studies have indicated that the calculated value of ϵ_0 is insensitive to the value of ϵ_{rf} provided $\epsilon_0 \leqslant \epsilon_{rf} \leqslant \infty$ [53][59][63][64].

The dielectric constant of the system can be related to the Kirkwood g-factor given by Equation 2.41. Under conducting boundary conditions where ϵ_{rf} =∞ the equation reduces as given by Equation 2.42, where y is given by Equation 2.43 where ρ is the number density,

μ is the magnitude of the dipole moment, k_B is the Boltzmann constant and T the absolute temperature. The Kirkwood g factor is usually calculated in simulations by evaluating the average of the square of the total dipole moment of the system $< M^2 >$ and is given by Equation 2.44

$$\varepsilon_i = \frac{2(\epsilon_{rf} - 1)}{2\epsilon_{rf} + 1} \frac{1}{r_c^3} \sum_{j \epsilon R} \mu_j \tag{2.38}$$

$$\mathcal{U} = -\frac{1}{2} \mu \cdot \varepsilon_i \tag{2.39}$$

$$\mathbf{T}_i = \mu \times \varepsilon_i \tag{2.40}$$

$$\frac{(2\epsilon_{rf} + 1)(\epsilon_0 - 1)}{(2\epsilon_{rf} + \epsilon_0)} = 3yG_k(\epsilon_{rf}), \tag{2.41}$$

$$\epsilon_0 = 3yG_k(\infty) + 1, \tag{2.42}$$

$$y = \frac{4\pi\rho\mu^2}{9k_BT} \tag{2.43}$$

$$G_k = \frac{< M^2 >}{N\mu^2} \tag{2.44}$$

2.8.2 Shifting Functions

The use of a cutoff beyond which the interactions are considered to be absent will result in a discontinuity in the energy and force and lead to pumping effects and loss of energy conservation; this is especially acute for long range interactions. One of the methods to overcome this problem is by the use of a shifted potential where the interactions are modified by subtracting from it a Taylor series of the original energy expanded about $r = r_c$ [65]. Where r_c is the cutoff distance and r_{ij} is the inter site distance. The resulting modified potential is given in Equations 2.45 and 2.46, where $U^{'}(r_{ij})$ is the modified potential and $\mathcal{U}(r_{ij})$ the original. The number of terms to be utilised in the Taylor series expansion depends on the type of potential, time step as well as the level of energy conservation required. With increasing terms in the Taylor series the shifted function becomes smoother with increased energy conservation but at a higher computational cost. A shifted potential was used in monopole-monopole interactions and dipole-monopole interactions in $LANKA$.

$$\mathcal{U}^{'}(r_{ij}) = \mathcal{U}(r_{ij}) - \left[\mathcal{U}(r_c) + (r - r_c)\left(\frac{d\mathcal{U}(r_c)}{dr}\right) + \frac{1}{2}(r - r_c)^2\left(\frac{d^2\mathcal{U}(r_c)}{dr^2}\right)...\right] \qquad r_{ij} \leq r_c \tag{2.45}$$

$$\mathcal{U}^{'}(r_{ij}) = 0.0 \qquad r_{ij} > r_c \tag{2.46}$$

2.8.3 Switching Functions

An alternative to the shifting function is the switching function where the potential is multiplied by a polynomial function S that tapers the potential to zero at the cutoff. Though this polynomial function can be applied to the entire range up to the cutoff it is rarely done since it affects the potential being computed and the resulting potential deviates from the "true" potential as in the shifting function. In fact, the shifting function can be considered a form of switching function. Usually the switching is done close to the cutoff, tapering the values from a lower bound to an upper bound with both the lower and upper bounds as close to each other in distance as possible. Several forms of switching functions exist [24][39][55][66]. The switching function used in dipole-dipole interactions is given by Equations 2.47-2.49, where r_u and r_l are the upper and lower bounds of the switching function.

$$S(r_{ij}) = 1.0 \qquad r_{ij} < r_l \tag{2.47}$$

$$S(r_{ij}) = \frac{(r_u - r_{ij})^2(r_u + 2r_{ij} - 3r_l)}{(r_u - r_l)^3} \qquad r_l < r_{ij} < r_u \tag{2.48}$$

$$S(r_{ij}) = 0.0 \qquad r_{ij} > r_u \tag{2.49}$$

2.9 Controlling Temperature

In the isobaric isothermal and canonical ensembles the temperature of the system is constrained whereas in the microcanonical ensemble the temperature fluctuates as it is the total energy that is conserved. The temperature of the system is related to the kinetic energy, where m_i is the mass of the particle and the index α runs over the different (x, y, z) components of the momentum $p_{i\alpha}$ [24] and N_c is the number of constraints.

$$K = \sum_{i=1}^{N}\sum_{\alpha}\frac{p_{i\alpha}^2}{2m_i} = \frac{k_B T}{2}3N - N_c \tag{2.50}$$

One of the easiest methods to control the temperature is through the use of velocity scaling, where at periodic intervals the linear and angular velocities are multiplied by a factor as given by Equation 2.51, where T is the desired temperature, K the kinetic energy, k_B the Boltzmann constant and g the number of degrees of freedom. This method is not suitable to be used during the production runs of a simulation and is usually used during the equilibration.

$$\lambda = \sqrt{\frac{g k_B T}{2 < K >}} \tag{2.51}$$

Several alternative schemes exist that guarantee the system to be in the canonical ensemble such as the Anderson thermostat and the Nose -Hoover thermostat[5]. The method used in $LANKA$ is by Berendsen et al. [67] which is a refinement of the velocity scaling method where at each step the velocities are scaled by a factor χ as given by Equation 2.52, where τ is the current kinetic temperature and T the desired temperature. The method pushes the temperature towards the desired temperature at a rate determined by t_T. While the method does not rigorously sample in the canonical ensemble, the ease of use has made this method widely applied in simulations [51][68].

$$\chi = \left(1 + \frac{\delta t}{t_T}\left(\frac{T}{\tau} - 1\right)\right)^{\frac{1}{2}} \tag{2.52}$$

2.10 Controlling the Pressure

For realistic simulations that can be compared to laboratory conditions it may be necessary to control not only the temperature but also the pressure of the system. The pressure of the system is related to the virial. The virial is defined as the expectation value of the sum of products of the coordinates of the particles and the forces acting on them as given by Equation 2.53 [39], where x_i is a coordinate and F_{xi} is the force along that coordinate. The pressure is related to the virial through Equation 2.54, where r_{ij} is the inter particle distance between site i and j and $\mathbf{F}_{ij}$ the force between them with N the number of particles in the system and V the volume of the simulation box.

$$W = \sum_{i=1}^{N} x_i \mathbf{F}_{xi} \tag{2.53}$$

$$P = \frac{1}{V}\left[N k_B T - \frac{1}{3}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} r_{ij} \cdot \mathbf{F}_{ij}\right] \tag{2.54}$$

Several methods have been developed to control the pressure in simulation studies such as the Parrinello-Rahman barostat, Andersen barostat, and the Berendson barostat [24]. The method used in *LANKA* is by Berendson et al. [67][24], where the pressure is controlled by coupling the system to an external 'pressure bath' [67]. An extra term is added to the equations of motion to produce a pressure change. The system made to obey Equation 2.55 where P_{ext} is the desired pressure and t_P is a coupling constant which determines the rate at which the system is pushed towards the desired pressure akin to t_T in Equation 2.51 for the temperature [24].

$$\frac{dP}{\delta t} = \frac{(P_{ext} - P)}{t_P} \tag{2.55}$$

At each step, the volume of the box is scaled by a factor χ, and the molecular centre-of-mass coordinates $\mathbf{r}$ by a factor $\chi^{\frac{1}{3}}$ as give by Equations 2.56 -2.57 where β_T is the isothermal compressibility of the system under study and δt the simulation time step [24]. Similar to the Berendson thermostat the barostat does not guarantee the isobaric isothermal ensemble although its ease of implementation has made it widely used.

$$\mathbf{r}' = \chi^{\frac{1}{3}}\mathbf{r} \tag{2.56}$$

$$\chi = 1 - \beta_T \frac{\delta t}{t_T} \left(\mathbf{P}_{ext} - \mathbf{P}\right) \tag{2.57}$$

2.11 Speeding up of computations

Having a cutoff does not necessarily reduce the time taken for the computation significantly. It is still necessary check if two particles are within the cutoff at every step, that is N(N-1) computations every step. However, in a simulation the neighbours of a molecule/atom do not alter significantly for several steps except possibly in low density gaseous systems. So if a list is maintained of all the neighbours of a particular site and updated periodically then the time "wasted" determining if two particles are within the cutoff is saved. First proposed by Verlet, the method stores all the neighbours of a particular molecule/atom within a particular distance in a large array. The distance is usually (slightly)larger than the cutoff (cutoff +"skin") thereby precluding any possibility of a molecule/atom initially outside the cutoff radius coming within the cutoff distance but being excluded from the potential energy computation between updates to the neighbour lists. The updates may be done at fixed time intervals (usually every 10-20 steps)or automatically by updating whenever the total magnitude of the two largest displacements of sites exceeds the difference between the cutoff and neighbour list distance (”skin”). The size of the ”skin” is inversely related to the frequency of updates and the number of non-interacting sites included in the neighbour lists.

Another method often used in conjunction with the Verlet neighbour lists is the cell linked lists. In this method the cubic simulation box is divided into a regular lattice of $M \times M \times M$ cells and the particles are assigned to one of the cells; a list of particles in each cell is maintained by the use of linked lists. The length of each box is usually taken to be slightly larger than the cutoff radius. This ensures that the neighbours of a site in a cell must lie either in the current cell or one of the surrounding cells with symmetry consideration ensuring that only half the neighbouring cells need to be checked. This method requires that the simulation box is at least $4 \times r_{cutoff}$ for it to be used, else the same neighbour will be accessed more than once. For situations where this condition is not met such as when using a large cutoff then the more expensive all pairs method must be used to build the neighbour list. The all pairs method was used to build the neighbour lists when considering lipid-lipid interactions, DNA-DNA interactions and mixed lipid-DNA interactions due to the use of a large 40 Åcutoff, while for other interactions the cell method was utilised to build the neighbour list in $LANKA$.

2.12 Conclusions

Molecular dynamics provides an elegant means of observing the dynamics and thermodynamics of particles provided that the forces (and torques) acting on the particle at a given instance are known and if a suitable algorithm/mechanism is available to propagate the motion (translational and rotational) with the passing of time; it provides us with a motion picture of the behaviour of particles

Chapter 3

Lipids and DNA: Biological Background

3.1 Lipids: Introduction

Lipids are biological biomolecules of intermediate molecular weight (100-5000 amu) that contain a substantial portion of aliphatic or aromatic hydrocarbons. This broad definition of lipids encompasses everything ranging from hydrocarbons, steroids, soaps, detergents, phospholipids, gangliosides and lipopolysaccharides. This large range of chemically divergent group then results in a large range of physical behaviour. These range from their solubility in water and other solvents to their mechanical properties. It is this range of divergent properties that has fascinated man for the past hundred years. The main focus of this study will be on a particular class of lipids called the glycerophospholipids, the different head groups of glycerophospholipids are given in Figure 3.1.

Most lipids in the human body are obtained from the food we consume; there are no special genes coding for the lipids in the body. Lipids have the ability to form interesting structures with unique properties; these structures and their properties are determined by the biological solvent water whose peculiar properties drives their formation. The structures formed by lipids are held together by non mechanical forces, making lipids a type of soft matter. Since the interactions between lipids are weak, thermal energy has a great influence on the formation and breaking of lipid structures. For many decades lipids were known as playing the role of cell compartmentalisation; however that view is being challenged by new findings that point to lipids playing an important cellular role [36].

Lipids often are found as membranes in biological systems; these membranes also con-

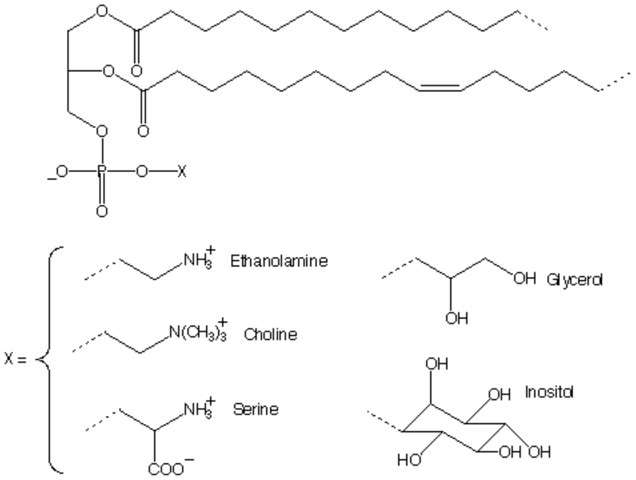

Figure 3.1: Glycerophospholipid Chemical Structure [69]

tain proteins and sugars whose weight ratio to lipids varies form membrane to membrane. Membranes are the most abundant cellular structure in all living matter [10][36].

Membranes lack a well defined structure; they are self assembled aggregates in which subtle elements of structure arise out a state of substantial disorder, so the role of entropy should not be overlooked in discussing the structure of lipids.

The lipids show a remarkable diversity in behaviour. This behaviour is brought about by the diversity seen in the head group that may be charged or zwitterionic at neutral pH to the saturation and length of the tails to the type of glycosylation. This leads to the lipids being the most chemically diverse group in the cell. For example, tri-acylglycerols are are storage and fuel lipids, while phospholipids and sphingolipids are structural and functional lipids. The large quantity and diversity of lipids within living organisms begs the question of why ?

3.1.1 Phospholipids

Phospholipids are complex molecules of biological origin containing a phosphorous in the head group and linked to the hydrocarbon tail through a a glycerol or spingosine backbone. These molecules are found in most (exception being Archaea microorganisms) living membranes and membrane organelles playing diverse roles. These range from maintaining the cell integrity by forming a barrier between the cell and the external environment and controlling the passage of substances to and from the cell, as well as regulating membrane proteins to still undefined roles within the nucleus.

3.1.2 Cholesterol

This is a type of lipid that is quite distinct from other types of lipids in the cell, based on sterol chemistry where the backbone is made of a steroid ring and a short hydrocarbon chain with a hydroxyl -OH group serving as the polar head group. The cholesterol molecule can be considered as a bulky and stiff tail lipid with a small head group. The different chemistry in cholesterol gives it properties different from lipids based on fatty acids and this is used by nature to control the properties of cell membranes by varying their cholesterol content, varying from 20-50 percent in animal cells [36].

3.1.3 Lipid composition in membranes

Depending on the organism and the role of lipids, the composition of lipids in membranes varies greatly. Cholesterol is found in plasma membranes of all animals. Prokaryotic plasma membranes do not contain any sterols.

The amount of charged lipids is about 10 percent of total lipids although the ratio between Phosphatidylinositol (PI) and phosphatidylserine (PS) varies greatly. The longer the lipids chain lengths, greater the number of double bonds. Lipid variety can not only be seen from cell to cell but also within the same bilayer. There is great lipid asymmetry in both the inner and outer bilayers, sphingomyelin (SM), phosphatidylcholine (PC), cholesterol and glycoproteins are enriched in the outer monolayer and PS, PI and Phosphatidylethanolamine (PE) are enriched in the inner monolayer [36].

3.2 Lipid Phases

Despite the bewildering range and complexity of biological lipids they all show similar lyotropic and thermotropic phase behaviour. The phases formed are determined to a large extent by the hydrophilic/hydrophobic balance between the head group and the tail group which determines the preferred interfacial mean curvature of the polar/non-polar interface and the packing constraints imposed by the head and tail components. The primary focus of this study will be the influence of the geometric constraints (also called packing constraints) on the phase behaviour of lipids.

If the lipid has a non-cylindrical effective shape then when lipid are arranged in a bilayer the two monolayers will develop an intrinsic curvature. This results in a development of a built in frustration called a curvature stress field [36]. If the spontaneous curvature of the two monolayers is different then the bilayer itself will become asymmetric and develop a spontaneous curvature. If the bilayer cohesive forces are unable to withstand these curvature stresses the lipid bilayer will assume a non-bilayer structure. This behaviour results in a rich structural polymorphism.

The tendency of lipids to aggregate is primarily driven by hydrophobic effect. Lipids show both lyotropic and thermotropic mesomorphism; the final resulting structure will depend primarily on the temperature and water concentration (though the phases can be altered by the presence of ions and pH) and shows two primary types of transitions. One type of transitions occurs when lipids change from one morphology to another and this is usually driven by the lipids tendency for spontaneous curvature. The transitions occurring within the lamellar phase are primarily the melting transition and are driven by entropy due to the increased disorder in the hydrocarbon tails due to chain rotation isomerism.

The various phases the lipids adopt may be classified according to long range order, the disorder in tails, and for curved surfaces whether these are the normal or the inverted types of curvature. This classification is based on the that proposed by Lazzati [70]. An upper case Latin letter characterises the long range order: L, one dimensional (lamellar); H, two dimensional hexagonal ; P, two dimensional oblique; Q, three dimensional cubic; T, three dimensional rectangular; R, three dimensional rhombohedral; C three dimensional crystalline. Short range conformations of the lipid tails are characterised by Greek letters as subscripts to the long range Latin letters: α, disordered(fluid); β, ordered untilted, β' tilted. Roman numerals used to characterise if the system is oil in water (normal-I) or water in oil (inverse-II).

3.3 Packing Parameter

The packing parameter theory is an elegant way of explaining in a qualitative manner the rich mesomorphism in lipids. First introduced by Isrealivichi and co workers [33][34][35] it is widely used as an alternative theory or in addition to the curvature theory at the interface. The packing parameter is defined by Equation 3.1 where a is the effective area of the head group, l the length of the hydrocarbon tail and V the volume of the lipid hydrocarbon region (see also Figure 3.2). The packing parameter is a quantitative measure of the compatibility between the sizes of the heads and tails of the lipids. If the lipid has an effective shape approximating a cylinder where the head groups and tails have approximately equal sizes then the volume of $a \times l$ will be approximately equal to the volume of the lipid tail group, resulting in a packing parameter value of ≈ 1. If the packing parameter value deviates from unity non lamellar phases can be expected. If the lipid has a relatively large head group

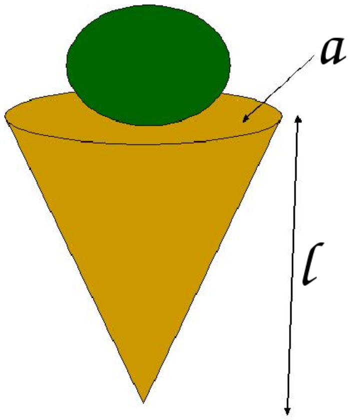

Figure 3.2: The definition for the packing parameter: a is the cross sectional area and l the length of the extended hydrocarbon tail

with respect to the tail (or if the lipid is a single tailed lipids such as dodecylphosphocholine (DPC)) resulting in a large effective cross sectional area a then the resulting volume of the enclosing cylinder will be greater than the actual volume V of the lipid tail giving a packing parameter value of less than unity. If the lipid has a bulky tail with a small head group region the volume of the enclosing cylinder will be less than the actual volume of the lipid hydrocarbon region resulting in a packing parameter value of greater and 1. If the packing parameter is less than unity, lipid phases such as micelles, hexagonal phases and cubic phases can be expected. Israelachvili and co workers [33][34][35] have predicted micelle like structures for packing parameter values $\leq 1/3$ and hexagonal structures for $> 1/3$ and $\leq 1/2$. While for packing parameter values greater than unity inverse phases such as inverse micelles, inverse hexagonal and inverse cubic phases can be expected. A lipid with a packing parameter value ≈ 1 it should give structures resembling bilayers. The structures shown with varying packing parameters are given in Figure 3.3.

$$P = V/a \times l \tag{3.1}$$

3.4 Lyotropic Phase Transitions

The lyoptopic phases found in lipid-water systems depends on the concentration of water as well as the type of lipids in the system. At low lipid concentrations the lipids exist as monomers, beyond the critical micelle concentration (CMC) they self aggregate to form micelles. With further increase of lipid concentration the system may undergo a transition to

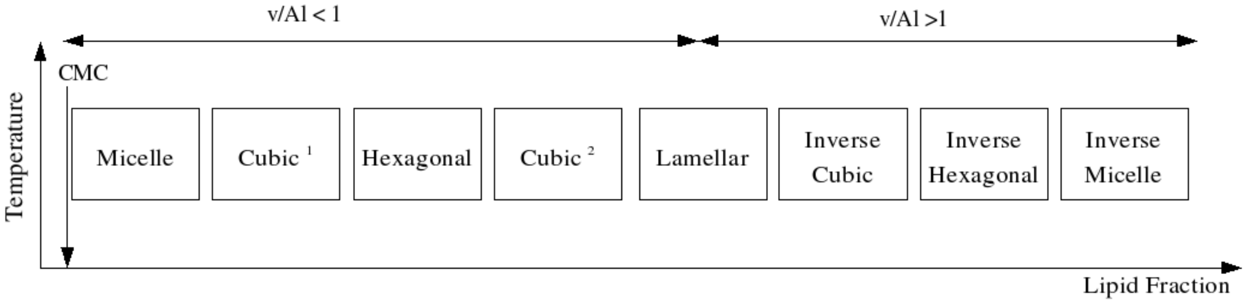

Figure 3.3: Variation of the phases with lipid concentration and the role of packing parameter on lipid phases

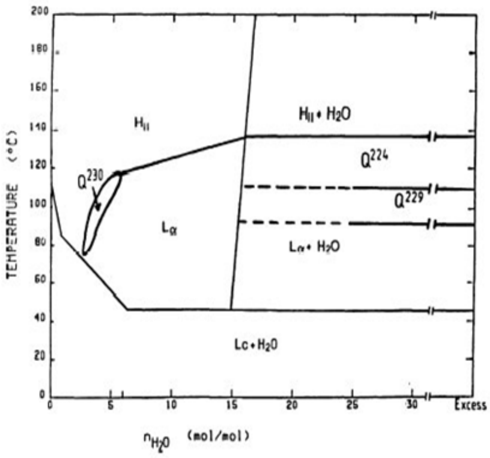

Figure 3.4: Phase diagram of dihexadecyl-phosphatidylethanolamine/water system. Figure from Templer et al. [71]

a cubic then an hexagonal phase followed by a lamellar phase. With further increase in the lipid concentration the lamellar phase then converts to an inverse cubic, inverse hexagonal, and finally an inverse micelle phase. Not all phases listed above are shown by a single type of lipids; phases starting from micelles ending at the lamellar phase are usually shown by single tail lipids while phases from the lamellar phase to the invere phases are shown by diacyl/dialkyl lipids. A schematic of the above listed lyotropic phase behaviour is shown in Figure 3.3 with a actual phase diagram of dihexadecyl-phosphatidylethanolamine/water system shown in Figure 3.4 clearly showing the different phases phospholipids can adopt in nature .

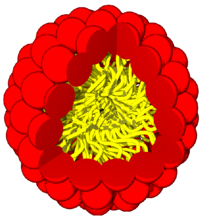

Figure 3.5: Micellar. Figure courtesy of Dr. S. Hant

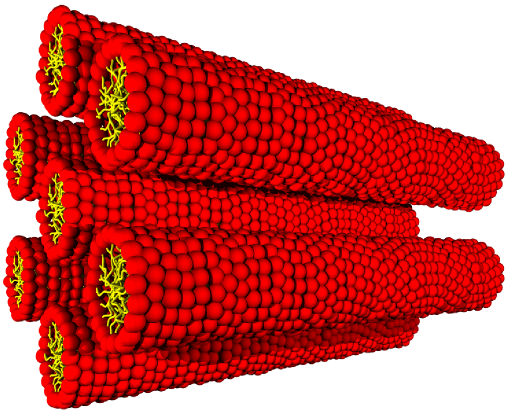

Figure 3.6: Hexagonal Phase. Figure courtesy of Dr. S. Hant

3.5 Micelles

Micelles are structures with a hydrophobic core and a miceller corona made of hydrophilic head groups that may be ionic or zwitterionic. Micelles may be in a range of structure from the ideal (theoretical) spherical to more commonly found ellipsoidal and spherocylindrical forms. The non ideal elliposoidal and spherocylindrical forms arise as a result of packing constraints, leading to in some instances to exposed hydrocarbon tails, a schematic of a idealistic (theoretical) spherical micelle is shown in Figure 3.5. Micelles are the simplest of the mesoscopic structures adopted by lipids/amphiphiles. At low concentrations of lipids the lipids exist as individual sites called unimers aggregating with increasing lipid concentration to form micelles; the concentration at which unimers aggregate to form micelles is called the critical micelle concentration(CMC).

3.6 Hexagonal Structures

Hexagonal structures are long cylindrical structures with the hydrophobic tail in the core and the hydrophilic head groups on the exterior of these long tubular structures. These cylindrical structures are then arranged in a hexagonal arrangement as shown in Figure 3.6.

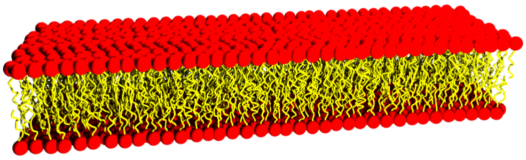

Figure 3.7: Lamellar Phase. Figure courtesy of Dr. S. Hant

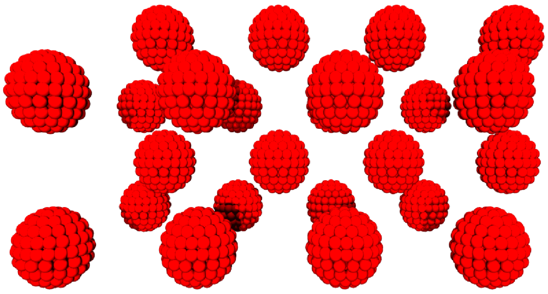

Figure 3.8: Cubic Micelle Phase. Figure courtesy of Dr. S. Hant

3.7 Lamellar Structures

This is the most common structural form found in living cells, flat 2-D structures with the hydrophilic head groups on either side with the hydrophobic tails back to back as shown in Figure 3.7.

3.8 Cubic Structures

Several cubic structures exist with different space groups. Cubic structures are bicontinous where the water is divided into disconnected regions, one on each side of the lipid bilayer which is curved everywhere (infinite periodic minimal surfaces). A schematic of a cubic micellar phase is shown in Figure 3.8.

3.9 Inverse Hexagonal Structures

These are similar in structure to hexagonal phases but with the hydrophilic head groups on the inside and the hydrophobic tails outside. The hydrophilic core usually contains a column of water as shown in Figure 3.9.

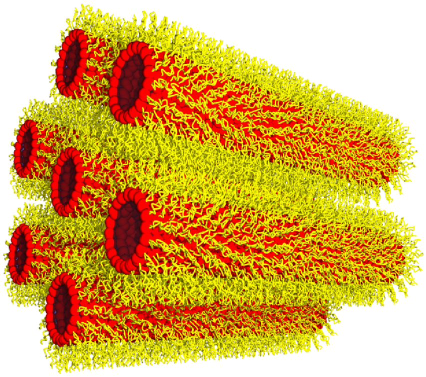

Figure 3.9: Inverse Hexagonal Phase. Figure courtesy of Dr. S. Hant

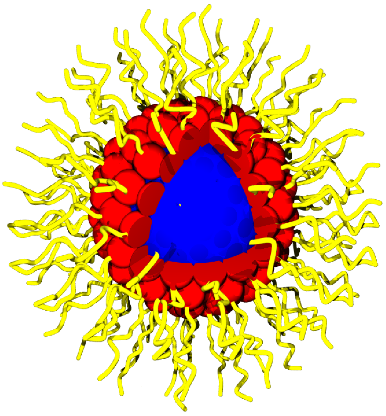

Figure 3.10: Inverse Micelle Phase. Figure courtesy of Dr. S. Hant

3.10 Inverse Micelles

These are spherical or ellipsoidal structures with the hydrocarbon tail on the outside and the hydrophilic head groups on the inside in contact with water as shown in Figure 3.10.

3.11 Thermotropic Transitions

While lipid in any morphology may show thermotropic behaviour, it is the thermal behaviour of lipid bilayers that has fascinated biophysicists for decades. In excess water a single phase can exist over a a large range of temperatures while two phases can exist only at a fixed temperature (Gibbs phase rule) and this ensures that thermotropic transitions are first order transitions showing sharp changes in several physical properties of the bilayers. Of the thermotropic transitions the most important is the the melting transition (T_m) where the lipid bilayer enters a fluid phase from a gel phase.

A generalised transition scheme starting from a low temperature system with increasing temperature may go from a crystalline (L_c) to a hydrated lamellar gel phase (L_β) phase. The gel phase may be of the tilted form ($L_\beta^{'}$) where the hydrocarbon tails are tilted respect to the bilayer normal or untilted (L_β). This is followed by a pre-transition to a rippled gel phase ($P_\beta^{'}$) followed by the main transition to the liquid phase L_α). Further increases in the temperatures may result in the bilayer structure converting itself into non lamellar structures, mostly inverted type structures. The formation of inverted structures may qualitatively be explained by the packing theory where the increase of temperatures leads to increased splaying of the tails resulting in non-cylindrical structure that favour the formation of non-lamellar structures.

3.12 Bilayer Thickness

The bilayer thickness or height depends on several factors, most importantly on the length of the hydrocarbon tails, saturation, the level of hydration, the presence or absence of cholesterol and temperature.

3.13 Lipid Fluidity

The lipids in the bilayers are in the L_α state, where they are not in a simple 2D fluid state but in a smectic A phase liquid crystal phase. Constantly moving around with long range lateral order. The types of motions that lipids undergo vary depending on the time scales observed and has been explained using a combination of 2D fluid like and and cage jumping mechanism between locally confined regions in the liquid crystal. Lipids may undergo rattling motion at picosecond timescales with rotation about its axis and molecular protrusions while at longer time scales may diffuse in the bilayer plane [72]. A rarer event is the flip flop process where the lipids switch from one monolayer to the other this process might take anything up to a few hour to a few days.

3.14 Lipid Rafts and Domains

Some parts of membrane samples treated with non-ionic detergents were found to be resistant to their action. These resistant fragments have given rise to the idea that biological membranes are not always in a liquid-crystalline phase. Instead, sphingolipid and cholesterol-rich membrane fragments appear to exist, at least partially, in the liquid-ordered phase or a phase with similar properties. The existence of rafts and domains have led to the modification of the widely accepted structure of the cell membrane; lipids and proteins diffusing randomly lacking any structure in the lateral direction. These raft like domains usually contain proteins, and high concentrations sphingolipid and cholesterol. Fatty acid chains of lipids in rafts tend to be more saturated than the lipids surrounds them and this allows for tighter packing, this in addition to the large quantities of cholesterol result in phase separation. The tight packing and phase separation are the factors giving rise to detergent resistant property of rafts [36][73][74].

These raft like domains floating in a sea of fluid bilayer have linked with several biological functions such as facilitating intra-membrane communication between the two leaflets of the membrane, cell surface adhesion and motility as well as cell surface signalling, intracellular trafficking and sorting of lipids and proteins [36]. The size estimates of rafts are heavily dependant on the experimental method; single particle tracking methods have given values around 26 nm while single dye methods measuring fluorescence of a fluorescent lipid probe have estimated rafts to be from 0.2-2.0 μm in diameter. The fraction the cell membrane covered by rafts also varies widely with estimates from 13% to 50% of the cell surface depending on the cell type and experimental method used. The types and concentrations of lipids and proteins is also found found to be heavily dependant on the cell and experimental method used to analyse them [36][73][74].

Another important type of domain is the *caveolae* domains rich in cholesterol and gycosphingolipids formed by the protein caveolin, appearing as invaginations of membranes and been found to be involved in cholesterol transport, cytosis and signal transduction. The presence of the protein caveolin distinguishes this from the other raft like domains [36][74].

3.15 Bicelles

Bicelles are a novel lipid structure consisiting of long and short chain phospholipids used for the biophysical and biochemical investigation of membrane associated biomolecules that have seen increasing use in place of unilamellar vesicles, multilamellar vesicles and micelles. Bicelles are used heavily in particular to better understand the role of membrane bound proteins. Membrane bound proteins are difficult to study because they exist in a complex matrix of different types of lipids in vivo. Nature uses different types of lipids to control the properties of membranes and hence the behaviour of membrane bound proteins [36] and so use of both experimental model membrane systems using unilamellar vesicles, multilamellar vesicles and micelles have become common as well as use of simulation studies. A good experimental/simulation mimic should allow for the incorporation of different lipids and other membrane components to accurately represent the natural cell membrane [75][76]. Experimental methods such as x-ray crystallography and NMR have been used to determine the structure if membrane proteins. The difficulty in purification and crystallisation of membrane proteins have limited the use of x-ray crystallography. The use of x-ray crystallisation is also questionable given that the membrane proteins are strongly dependent on their surroundings. NMR methods are better suited to study membrane-protein systems but there are several practical difficulties in using cell membranes, which has resulted in the use of micelles, unilamellar vesicles and multilamellar vesicles.

While micelles are simple structures and allow for study of membrane associated proteins using techniques such as NMR the packing in micelles is different to membrane lipid packing, increased entanglement in the center and high curvature of micelles will have an effect on capturing realistic protein-lipid interaction. Unilamellar vesicles are hard to prepare and are unstable, this has led to the use of multilamellar vesicles which are easier to make. However both unilamellar and multilamellar vesicles are still cumbersome to use in NMR studies given their low reorientation rates (results in line broadening). These shortcomings in micelles and vesicles have led to the use of bicelles in the recent past [75][76].

Bicelles can be prepared by mixing together long chain phospholipids and shorter length phospholipids or detergents. A schematic of the commonly accepted structure of bicelles is given in Figure 3.11, it consists of a planer central core usually made up of long chain phospholipids and surrounded by short chain lipids or detergents at the rims shielding the hydrophobic tails of the long chain lipids; a discoidal shape. An alternative structure where the longer tail lipids arrange edge-to-edge (lamellar structure) with pores stabilized by the shorter chain lipids/detergent; the so called Swiss cheese model as also been propsed [75]. The structure of the bicelle may be controlled by varying the ratio of the long and short chain lipids as well as the total concentration of lipids. Bicelles can either be neutral or may have a charged surface depending on the type of lipids used. Bicelles are better mimics of membranes than micelles as the have a planer core akin to cell membrane surfaces while the surface of the micelle is highly curved [76] and their ease of preparation, and the possibility of controlling hydration, pH and ionic strength allow to mimic a cellular environment better than micelles or vesicles. Bicelles have seen extensive use in NMR (soild and solution), electron paramagnetic resonance (EPR) and circular dichroism (CD) experiments to investigate lipid-protein systems [75].

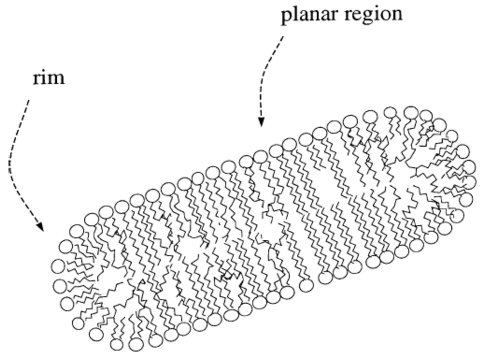

Figure 3.11: A schematic of a bicelle [76]

3.16 DNA: Introduction

DNA, a structure first described in the 1950s by Watson and Crick, has today established itself as the central molecule in the dogma of life [5]. DNA controls the types of proteins being synthesised in the cell through the transcription process and ensures that genetic material is passed on to the progeny through its ability to self replicate while maintaining the integrity of the genetic information encoded in the base sequence. There are several levels in the DNA structure starting from the base pair level up to the cellular-level arrangement in chromosomes within the nuclei [5][10].

The classical DNA structure described by Watson and Crick consists of a flexible ladder wrapped around an imaginary central axis. The rungs of the ladder are made from nitrogen base pairs bonded together by hydrogen bonds. The rails of the ladder made up of the DNA backbone consists of alternating deoxyribose sugar and phosphate groups. The DNA strand is a polymer made up of nucleotides. Each nucleotide is made up of a deoxy ribose sugar, nitrogen base and a phosphate group. Four types of nitrogen bases are commonly found in DNA; Cytosine (C) and Thymine (T) of the pyrimidine type and Adenine (A) and Guanine (G) of the purine type. According to the Watson-Crick pairing, Adenine pairs with Thymine and Cytosine with Guanine. A purine base pairing with a pyrimidine base ensures that the DNA structure is uniform. Once paired the nitrogen bases are held together by the formation of hydrogen bonds; two hydrogen bonds between Adenine and Thymine and three between Guanine and Cytosine. The nucleotides are connected together through the phosphodiester bonds that are formed between phosphate groups of one nucleotide with the deoxy ribose of the adjacent nucleotide. The resulting polynucleotide stand is then paired with a complementary polynucleotide strand to form DNA. The two strands forming the DNA have polarity; one end of the strand terminates with a $C5^{'} - OH$ ($5^{'}$ end) and the other end with a $C3^{'} - OH$ ($3^{'}$ end). The two strands then lie anti parallel to each other.

3.17 Canonical Forms of DNA

Slight variations in the helical geometry give rise to several canonical forms of DNA. The most prevalent types are A, B and Z forms of DNA of which the B form is the most commonly found and the one deduced by Watson and Crick. The B form of DNA occurs under high relative humidity (92% of water) while the A form occurs at round 75% relative humidity. The Z form of DNA whose biological role within the cell is still under investigation, occurs under high salt concentrations and when specific sequences are present in the DNA. The A and B forms are right handed while the Z from is left handed [5]. A listing of the characteristics of the three canonical forms is given in Table 3.1 [5].

Property	A Form	B Form	Z Form
Helix sense	Right	Right	Left
Mean base pair/turn	11	10.0	12 (6 dimers)
Rise-per-base pair/Å	2.6	3.4	3.8 (average)
Helix diameter/Å	≈ 26.0	≈ 20.0	≈ 18.0
Helix pitch/Å	28.0	≈ 34.0	≈ 45.0
Twist/residue	33.0°	36.0°	-60.0°/dinucleotide
Base pair inclination	20.0°	0.0°	-7.0°
Sugar pucker	$C3'$-endo	$C2'$-endo	$C2'$-endo(C)/$C3'$-endo(G)
Glycosyl rotation	anti	anti (higher)	anti (C)/syn (G)
Major groove	narrow and deep	wide and deep	convex
Minor groove	very wide and shallow	narrow and deep	very narrow and deep

Table 3.1: Comparison of Canonical DNA Characteristics [5].

3.18 Interactions of Ions and Water with DNA

DNA is built up of fragments that on first glance should not exist in a single structure; for example the DNA backbone is made up of negatively charge backbone particles spaced approximately 7 Å apart giving rise to a repulsive energy of around 50 kcal mol^{-1} in the gas phase, that would destroy the structure. However, this unfavourable interaction is counteracted by the presence of the solvent and the counter ions in the medium. The stabilising counter ion environment may be provided by histone proteins, bioinorganic molecules like spermine or by monovalant and divalent cations [77]. The high dielectric medium of water surrounding the DNA provides a screening medium for the ions and the backbone charges; experiments and simulations have indicated that the bases and phosphate groups have organised hydration shells surrounding them [5]. The presence of counterions and the high dielectric screening provided by water allows DNA supercoiling, allowing the phosphate groups to come together. In addition the ions have been found to be responsible for modulating the B to Z transition [77].

Several views of cation binding/interactions with DNA have been put forth since the discovery of the double helix structure of DNA in the 1950s. They can be broadly classed into two. The earlier understanding based on the theory by Manning (Manning condensation) assumed that the ions were a delocalised cloud around the DNA, not in direct contact with the DNA, with their hydration shells intact lying along the entire length of the polymer (uniform distribution) [78][79]. Subsequent theories, in part spurred on by the stability of cation-RNA interactions, indicated that the ions were tightly bound to the DNA and during the binding process lost part of their hydration shell. However the current understanding is that the real situation might be far more complicated than first anticipated. It is seen that ions, while interacting in a sequence specific manner, do not bind strongly to the backbone or the bases but have measurable residence times in the strand although the percentage of ions in the groove or its vicinity is quite low [77][80][81][82][83][84]. The ions are found to be in a highly diffusive state moving around the DNA strand but not localised in a particular site. The renewed interest in cation-DNA interaction during the past decade was in part the result of the work by Beveridge et al. [85] who suggested through their MD simulation that a single Na$^+$ ion resided deep within the minor groove with long residence times in fact penetrating the "spline of hydration".

Ions can bind with the negatively charged backbones or the electron donors in the bases such as N and O. Alkali metals, alkali earth metals and transition metals can interact with the DNA helping to stabilise the secondary and tertiary structures of DNA. Transition metal with their empty d-orbitals are known as free radicals and are known to bind directly to the

DNA while the alkali earth and alkali metals cations bind through their hydration shells or in a partially dehydrated state [77][81].

The minor groove has proved interesting to theoreticians/simulators, in part due to the well ordered "spline of hydration" seen in the Dickerson dodecamer [86] and the A tracts seen in the Dickerson dodecamer that have been identified with increased flexibility [5]. This is turn has resulted in a large body of work investigating the interactions of monovalant and divalent cations with the minor grove.

The most common techniques used for investigating DNA-cation interactions are solution NMR and X-ray crystallography, and capillary electrophoresis, and molecular dynamics simulations [77]. X-ray crystallography is by far the most common experimental method used. However if and when ions have been located in X-ray crystal structures of DNA it has been far less than what is required for charge neutrality. This has been explained in part by the low resolution of X-ray crystal structures, difficulty in detecting ions, particularly Na^+ cations, and the low occupancy of the ion binding sites [77].

Denisov and Halle [84] using magnetic relaxation dispersion (MRD), indicated that Na^+ preferentially binds to the A tracks in B-DNA with correlation times up to 50 ns with up to 5% of Na^+ ions in the minor groove in competition with the waters in the spline of hydration. Stellwagen et al. [87] using diffusion rates of duplexes in capillary electrophoresis, also indicated a preference for the A tracks in B-DNA though others have not seen such a preference [80]. There is still a lot of controversy with regard to detection of the monovalent cations in crystal structures, particularly Na^+ even though some have reported possible location of Na^+ in the minor groove of the B-DNA A tracks. However, the evidence of heavier alkali metals binding in the minor groove seems well established (Cs^+ and Rb^+) [77][81]. One source of continuing controversy is whether the locations and interactions of the heavier alkali metals and other substituents such as Tl^+ are representative of the lighter alkali metals such as Na^+ and K^+. Some authors have in fact argued against the binding of any cations in the minor groove [80].

Simulations have also been extensively used to resolve this issue although different workers have drawn different conclusions. Early simulation work was probably too short to provide an accurate measure of the DNA-cation interactions as well as having issues with the accuracy of the force fields used. However since the mid 90s a whole raft of reliable simulations have been reported, with specialised force fields that are able to handle the electrostatics in the system better than the first generation force fields, improving their reliability.

A 15 ns simulation by Wilson and co workers [88] as well as a 10 ns simulation by Feig and Pettit [89] indicated that there is some narrowing of the minor groove on entry of Na^+

although this has not been seen in the work by Beveridge et al. [90] using the same Dickerson sequence as the Wilson group [88], but the definitions used to determine if the ion is in the minor groove may be the cause of the disparity. The Beveridge group [90] also did not see the opening of the minor groove in the absence of water (structured water) and cations, as seen in the work by Wilson and co workers [88], and binding of ions in the minor groove is not considered sufficient to cause the narrowing of the groove. Feig and Pettit find that up to 20% Na^+ entry into the minor groove with residence times from 2-200ps [89] while Beveridge [90] finds 5-10% entry, and Wilson et al. [88] report residence times up to 6 ns in their 10 ns simulations. Again there is wide disparity among the simulations about the nature of the Na^+ interaction with the DNA although it seems likely that Na^+ ions can occasionally enter the minor groove but their effect on minor groove narrowing is still an open question. Laughton et al. finds that there is a possible a Na^+ binding site inside the minor groove with residence times of 10 ns. No exchange of Na^+ and water was observed in the minor groove although the simulation was for only 10 ns and possibly too short to observe the exchange [80]. The release of ions from the minor groove was difficult and found to be an energetic process. It is also found that waters are more stable within the minor groove than ions, with the ions trapped in the minor groove having an insignificant effect on DNA structure.

Egli and co workers have studied Dickerson dodecamer with Mg^{2+} and Ca^{2+} using X-ray crystallography and found that one Ca^{2+} ion binds deep in the minor groove widening it slightly. There is strong evidence that the interaction of divalent cations is strongly mediated by water and strong correlation exists between the position of the ion and the minor groove width with preferred binding to the GC tracts instead of the more favourable electron rich AT tracts; an issue that is still unresolved though possibly partly explained by the sizes of the minor groove containing AT tracts and GC tracts [91][81]. The ease of detection of Mg^{2+} and other alkali earth metals has meant the role in binding of divalent cations is more established than their monovalant counterparts and usually the divalent ions remain hydrated and more ordered in the minor groove.

In the major groove of the B DNA Ca^{2+} and Mg^{2+} were observed by Dickerson et al. [92]. Ca^{2+} was found to make direct contact while Mg^{2+} remains hydrated, explained by the ease with which Ca^{2+} could lose its hydration shell and has a less restricted coordination geometry. Binding of Ca^{2+} also compresses the major groove, in part due to the direct binding of Ca^{2+} without its hydration shell, although some others find fully hydrated Ca^{2+}. Ca^{2+} and Mg^{2+} are preferably found to bind at GpN and ApN steps in the major groove [81].

MD studies by Petit [81][89] indicate preferential binding of divalent ions to G tracts with

10-20 % occupancy and residence times of around 14-36 ps, while in the A from of the DNA longer residence times and greater ordering is observed, possibly due to the large opening to the major groove. Berveridge also report great affinity to the G tracts in his simulation, with 10 % occupancy [93]. The wider major groove results in increased exchange between sites of ions thus leading to lower residence times. Beveridge et al. [90] indicate that there is some occupancy of Na^+ in the major groove with slight narrowing of the groove for the Dickerson dodecamer with preferential binding to G-C sites. The preferential binding to G sites was also observed in the work by Feig and Petitt [89] who looked at both the A and B forms of DNA, with longer residence times for the A form.

Experiments and theory predict that the properties of water on the surface of the DNA helix is different from the bulk. It has been proposed that a layer of tightly bound water exists around the DNA with up to two layers of hydration containing anything from 5-12 waters per nucleotide. Water is more organised around the bases than the phosphates. Quasi electron scattering studies have indicated that it is easier to remove waters from the bases than those bound to the phosphates, indicating a stronger affinity for the phosphates. NMR studies have confirmed the "spline of hydration" in the minor groove first seen by Dickerson et al. [86] while few highly diffusive waters have been confirmed in the major groove. However, the hydration of the phosphate groups is probably too dynamic to be detected by NMR methods [94]. It has been shown through analysis of a multitude of crystal structures that the distribution of waters around the bases is more dependent on the chemistry of the bases and less on the conformation though the structural differences in the different canonical forms do play a role in the organisation of water around the bases. The more sterically restricted minor grooves of the B and Z forms have a higher density of water than the A form while the deeper major groove in the A form results in more ordered waters than in the B and Z forms [94].

Water has also been shown to act as a bridging agent between the phosphate sites helping to screen the charges between the interacting phosphates, especially in the Z and A forms of DNA, while in the B form, given the greater distance between phosphate sites, water is involved in hydrogen bonding with the phosphates. The increasing overlap between the phosphate hydration shells is thought to play an important role in the B to A transition at low hydration conditions. Water has the highest binding to the phosphate oxygens, followed by the oxo groups in the bases and then endo and exo cyclic nitrogens in the bases. The slowing of the diffusion of water around the charged backbone can be explained by the theoretical concept of elctrostriction. Early theoretical studies as well as analysis of crystal structures have revealed that up to three tetrahedrally arranged water molecules around each charged oxygen in the phosphate backbones exhibit characteristic hydration cones. Around 2-3 water molecules are found around the bases, and therefore around 8-9 tightly bound waters per nucleotide exist with low mobility and partial ordering. The second hydration shell is more weakly bound and more diffusive than the first. Depending on the distance from the helix, water molecules will have widely ranging lifetimes, from a few picoseconds in the rapidly exchanging outer layers to several nanoseconds in the more structured minor grooves (particularly A tracts in the minor groove) [5][94][95].

3.19 Persistence Length

The DNA can be viewed either as a rigid rod or a worm like chain depending on the length of the double strand under consideration. As a measure of the stiffness a persistence length may be defined. The persistence length is essentially the length scale over which the polymer directionality is maintained. Experimental measurements have indicated that DNA has a persistence length of 50 ± 10 nm if the solution contains at least 100μM Mg^{2+} or 10mM Na^{+} ions, showing little variation with ionic concentration at these "high" salt concentrations. Several experimental techniques have been used to determine the persistence length such as light scattering, cryoelectron microscopy, scanning force microscopy, force measuring force tweezers, and transient electric birefringence [5][96][97]. The persistence length usually decreases with increasing temperature due to increased thermal motion, making the double strand more flexible before the two strands separate(melt). However, there is still no agreement on the behaviour of persistence length at low temperatures, with some experiments indicating an increase (based on thermal motion) and others indicating a decrease with a shallow maximum appearing around 20 °C [97].

3.20 Conclusions

Lipids show a rich diversity in the phases they adopt. These phases can be varied by the changing of temperature and lipid/solvent concentration (lyotropic behaviour) while pH, ions and anaesthetic molecules may also have an effect. The phase behaviour of lipids can qualitatively be described using the packing theory which can be used to good effect in designing models to capture the rich phase behaviour of lipids. Lipids also show interesting behaviour in multicomponent mixtures, which nature has utilised to control the behaviour of the membranes.

DNA serves as the central molecule in the dogma of life; many of its complex structural and dynamic behaviour results from the highly divergent groups that are held together to form it. The complex interactions occurring within the molecule as well as with the surrounding environment are yet to be completely unravelled.

One conclusion that an be reached regarding the interactions of lipids and DNA is that these are highly complex systems requiring study using a multitude of experimental techniques ably supported by the use of appropriate theory and modelling.

If the lipid-DNA problem is to be tackled using simulations/modelling the simulations must be done in several levels of granularity. The models presented in this work are capable of tackling problems in microsecond to millisecond timescale where it would be expected to observe the self-assembly of such mixtures and the possible emergence of structures predicted by theoretical considerations such as hexagonal and bilayer intercalated DNA structures [8].

Lipid mixture simulations in the length scale of micrometers and time scales of microseconds and extending further if possible will allow for us to observe formation of domains and rafts and will make it possible to observe (possibly) processes such as flip-flops between the leaflets in a bilayer and permeation of water through the bilayer. Long simulations extending into the microsecond time regime become feasible with coarse grain lipid models as discussed in this work [36].

Chapter 4

Lipid and DNA Models: Literature Review

The existence of lipids and DNA in the nucleus is well established, and given their polar nature and their quantities within the nucleus they would be expected to interact with each other resulting in meso and macro structures being formed [1][2][3][4]. To capture these interactions a coarse graining approach was considered the most appropriate. Literature pertaining to coarse grain lipid and DNA models are summarised in this chapter.

4.1 Introduction

Notwithstanding the merits of experimental breakthroughs, true microscopic understanding can only arise in close conjunction with theories and simulations. Theories provide a basis for interpretation of experimental results, and simulations can elucidate atomic-scale processes that give rise to thermochemical properties. In addition, the simulations can also be used to test theories by directly simulating model systems in full detail without the simplifications or assumptions often needed to solve analytical equations[24][22].

Molecular mechanics simulations provide a means of studying lipids, protein and other biomolecules from a single molecule upwards. In fact, simulations of biomolecules, in particular proteins, lipids and DNA have been around for several years and are now routinely used for validation, understanding and extending experimental observations. There are several methods commonly employed to perform computer simulations of biomolecules. The two extremes in tackling this problem are the all-atom simulation in which the biomolecules and the embedding solvent is modelled explicitly, and the other is phenomenological models. However, atomistic simulations of molecules are computationally expensive, accessing time

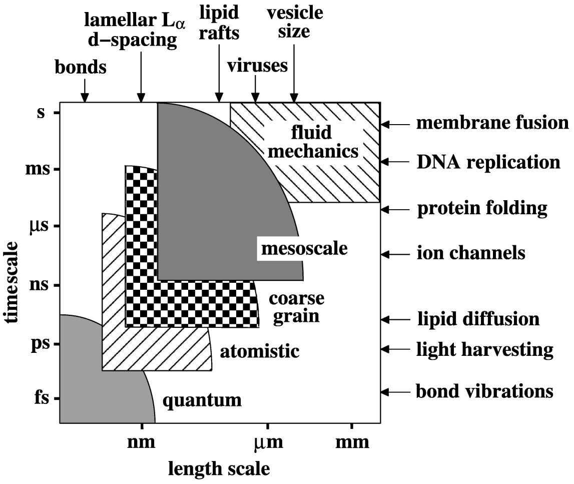

Figure 4.1: Schematic diagram of temporal and spatial scales accessible by simulation techniques. Also indicated are some characteristic biomolecular structures and events associated with these spatial and temporal scales. Figure from Nielsen et al. [99].

scales of nanoseconds and spatial scales of nanometres. This severely restricts the phenomena that can be studied using these simulation methods. Many events taking place within the living cell occur over longer spatial and temporal scales. Phenomenological models are unable to provide details of individual particles and are used to elucidate information on the macroscopic scales.

Several methods have been proposed to bridge the gap, one of which is the use of coarse grained models. Coarse graining is able to access larger spatial and time scales by subsuming some of the interaction sites, thereby reducing the calculation overhead drastically. Coarse graining results in the loss of a great deal of atomistic detail, retaining only the properties pertinent to bring about the results in the meso and macro scales. The most important question in the design and development of mesoscale models is determining which degrees of freedom are retained and which are lost due to the subsuming of sites. Often it is difficult to determine *a priori* the degrees of freedom required to capture a particular property using a coarse grain model. It is therefore often done in a repetitive manner using trial and error testing. Figure 4.1 summarises the relationship between simulation techniques and accessible time/space scales. Furthermore, in coarse-grained representations dihedral angle potentials are usually absent, and electrostatic interactions if present are highly simplified. Most phenomena of lipid bilayers such as lateral diffusion of lipids, undulations, phase separation and rafts (fundamental to account for the biological role of bilayers), DNA folding and supercoiling as well as the calculation of DNA persistence length are beyond the time and length scales accessible to atomistic MD simulations[98], and therefore require coarse grain models.

Coarse grain models of any biological system can be developed in two ways; top down or bottom up. The bottom up approach starts from an all atom description and works its way up through simpler and simpler model while attempting to maintain certain macroscopic and mesoscopic properties that are deemed important. In the top-down approach the models attempt to reproduce mesoscopic and macroscopic properties without considering the microscopic nature of the molecule being modelled. While the bottom up approach uses repetitive methods such as force matching [100] to match RDFs [101][102], the top-down approach is done in a more haphazard fashion, simply trying to capture some properties by fine tuning the potentials. Most implicit solvent models and idealistic explicit solvent models fall into this category. It is the models developed by the top down approach that this review will focus on where lipids and DNA are represented by a few sites (beads), and where the models represent a particular class of lipids or a generic DNA structure, rather than a particular lipid or canonical form of DNA.[14][16]

4.2 Coarse-Graining of Lipids

Lipids are amphiphilic and have a polar hydrophilic head group and a hydrophobic tail group. The hydrophilic head group is usually a zitterionic or charged species. The hydrophilic end interacts favourably with water forming hydrogen bonds while the hydrophobic end, interacting unfavourably with water, shields itself from contact with water. The unfavourable interaction with the hydrocarbon tail is due to the inability of the tails to form hydrogen bonds with water. This is called the hydrophobic effect, and is one of the primary factors driving the formation of different phases.

Lipids have been modelled for many decades using all-atom models based on different force fields of varying complexity, primarily focussing on the lamellar phase with some work on other lipid phases such as micelles [68][103][104][105][106][107][108][109]. These have often been explicit solvent model systems consisting of a small number of lipids usually with preassembled bilayers, with the exception of the all-atom simulation of Marrink et al. [17] which was a self-assembled system. These all-atom simulations have been used to elucidate vital information about the dynamics and structural properties of lipids. At the other end of the length scale, lipid membranes have also been represented as continuum sheets where individual lipid molecules are not explicitly represented but are shown as a single thin continuous sheet with curvature elastic energy (deformation energy). These models have no explicit representation of solvent, but solvent effects are present in determining the elastic properties that are used to define the model. The most common of these continuum models is the Helfrich model and its variants[110]. The Helfrich model has only two experimentally measured parameters, the bending rigidity and the surface tension, and parameters related to these serve as inputs to models. These types of ideal bilayer sheets have provided vital information about the mechanical properties of the bilayer at macroscopic length and time scales, such as thermal and shape fluctuations of red blood cells and vesicles [111] [14].

Lipids form many types of phases, the most important being the lamellar phase. In the lamellar phase the lipid is a 2D flat structure. The bilayer thickness is only 2-4 nm while the length of it can be up to several micrometers. Lipid membranes are multicomponent systems with varying concentrations of different lipids in addition to the proteins and sugars that are a part of the cell membrane. The various lipids found within a cell membrane are used by nature as a means of modifying the function and structure of membranes. Lipid membranes, long considered as a random aggregation of lipids with no distinct features in their lateral arrangement, have now been shown to have both structural and functional heterogeneity in the lateral direction [36]. This heterogeneity is shown via the formation of domains and rafts, which have been shown to play a crucial role in facilitating communication between the monolayers, cell surface signalling, cell surface adhesion and intercellular trafficking [36]. This structural heterogeneity in the lateral direction and the multicomponent nature of lipid bilayers, as well as the need to understand the behaviour of lipids over longer temporal scales, has resulted in increasing use of alternate modelling techniques such as coarse graining to capture the complex behaviour of lipids. Coarse graining of lipids in the mesoscopic length and time scales has different levels of granularity. Depending on the granularity, the information that can be elucidated form these models will vary, though at the cost of increased computational effort. Models containing greater levels of detail (lower levels of coarse graining) will result in greater agreement with the underlying molecules atomistic simulations and experimental observations [20][51][102][112]; these models also usually have explicit consideration of charges. With increasing level of coarse graining the model will lose its ability to represent a specific lipid but will generally represent a specific class of lipid species and with further coarsening it will simply represent an amphiphile. It is the highly coarse grain models that have been published on which this review will focus. Several excellent reviews on both implicit and explicit solvent coarse grain models of lipids have appeared in the literature[14][16][15][113].

Explicit Solvent Lipid Models

Explicit solvent lipid models are models in which the surrounding solvent is explicitly represented. In coarse grain models the solvent is often a coarse grain representation of real water representing a collection of more than one water molecule by a single site. The grouping of water is often arbitrary, or is made identical to the lipid head group site, interacting favourably with the head groups while repelling the tail site. The grouping of water is done to reduce the number of interaction sites, thereby increasing the time and length scales accessible for the simulation. Often, owing to the complexity of handling electrostatics, the solvent is simply represented as a sphere, completely ignoring the dipolar nature of water. The electrostatic interactions in the head group, if present, are scaled by the dielectric constant of water to artificially impose a screening effect in the system.

Implicit Solvent Models

The major restriction on the size of systems in simulations stems from the large number of atoms included in the simulation cell. The low ratio between the number of lipids and total number of atoms is due to the number of water molecules. Thus a great fraction of time is *wasted* on simulating water. The major difficulty in establishing $waterfree$ simulations is the need to mimic the hydrophobic effects that prevent the amphiphilic molecules from leaving the aggregate [114].

Models that include solvent particles are able to enforce the formation of bilayer and other lipid phases in the same way nature creates these structures; through hydrophobic effects. In the absence of the solvent, the hydrophobic effect has to be included through the use of effective potentials acting between the lipids.

A review of very simple explicit and implicit lipid models is presented below.

4.2.1 Drouffe Model

Introduction

This is an early implicit solvent model that shows stable fluid bilayers and vesicles [115].

Model

This is a single hard-sphere model with three parts: a hydrophobic layer surrounded by two hydrophilic parts. The model interacts through a distance and orientation dependent anisotropic potential: hard sphere potential, and to maintain a fluid bilayer structure an isotropic density dependent multi-body potential. The multi-body potential mimics the hydrophobic effect and favours a close-packed structure formation.

Simulation and Results

Molecular dynamics simulations show self assembly of both bilayers and vesicles; however, without the multi-body potential the self assembled bilayers are unstable at higher temperatures. The model shows transition from a solid close packed structure to fluid phase with concomitant increase in lipid diffusion. By analysing the fluctuation spectrum of the vesicle the authors were able to extract a value for vesicle rigidity.

4.2.2 Brannigan and Brown Model

Introduction

This is a single site implicit solvent model of lipids. The lipids are represented by rigid sites as shown in Figure 4.2.

Model

Each lipid is represented by rigid asymmetric soft spherocylinder [116][117][118]. The interactions were developed considering that the bilayer molecules are both amphiphilic and anisotropic. Bilayer structures were maintained by pair potentials consisting of three terms: an excluded volume repulsion, an orientation dependent attraction, and attraction between hydrophobic ends.

Simulation and Results

Monte Carlo simulations in the canonical and constant tension ensembles were performed. By variation of the energetic parameters the model was able to capture various lipid phases such as micelles, monolayers and bilayers. The bilayer in a fluid state was quantified by bond order, diffusion constant and correlation functions. The rigidity of the model results in bending moduli being higher than single component biological systems and is comparable to membranes containing cholesterol. A stress profile similar to those observed in solvated simulations and predicted by theories of elasticity was observed. The model under content tension conditions was used to determine the transition temperature for a gel to liquid crystal transition, and to determine the role of molecular steric interactions in the model phase behaviour [116]. Mixed lipid systems were simulated by varying the length of the lipid species, [117] and their thermal and elastic properties studied over the full composition range.

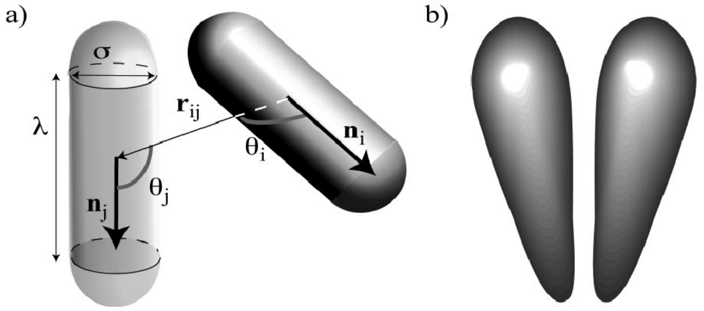

Figure 4.2: A schematic diagram of Brannigan and Brown spherocylinders (a) shows the parameters used in the model (b) shows the "effective shape". Figure from Brannigan and Brown [116].

4.2.3 Noguchi and Takasu Model

Introduction

This is a solvent and electrostatic-free rigid amphiphilic model using multitude potential to investigate a series of vesicle processes such as vesicle formation, fusion and fission [119][120][121][122].

Model

The model is made of up three spheres, one representing the hydrophilic head and the other two the hydrophobic tail. The model interacts through a soft core repulsive and a hydrophobic attractive interaction modelled using a multi-body potential. The multi-body potential helps maintain the bilayer structure which in the absence of the solvent would be unstable[119].

Simulation and Results

Brownian dynamics simulations were performed looking at the formation of vesicles starting from a random isotropic mixture passing through various lipid/amphiphilic morphologies dependent on temperature. The formed vesicles exist in a fluid state showing both transverse diffusion (flip-flops) and lateral diffusion [119]. The model identifies two possible pathways for vesicle fusion dependent on the temperature of the system. At low temperatures the model supports the pore opening pathway as suggested by the modified stalk model while at higher temperatures the pore opening proceeds through stalk bending [120]. The model was also used to investigate the effect of mechanical forces on vesicles, mimicking optical trapping experiments. Two nano particles were placed inside a vesicle, one was fixed while the other was moved with a constant force until fission occurs; this enabled the study of the intermediates formed during the fission process [121]. Nanoparticles were also used to study the adhesion process to vesicles' hydrophilic surfaces; this serves as a simple model of phagocytosis. This study also shows how the adhesion process may promote vesicle fusion and may serve as a crude model for proteins promoting fusion processes in the cell [122].

Extending the work of Nouchi and Takasu, Wang and Frenkel [123] made the lipid model flexible. The model is able to show the transition from solid to liquid phase and exists as a two dimensional liquid at room temperature. Using Monte Carlo simulations on a pre-assembled bilayer the model was used to calculate area compressibility, bending modulus and elastic ratio in a range of biological membranes. This model was also used to study the formation of pores in membranes. The model identifies both hydrophobic and hydrophilic pores, hydrophobic pores starting off the process growing into hydrophilic pores when their size grows larger than the molecular dimensions; this is in line with experimental observations. The free energy of the hydrophilic pore is qualitatively described by classical nucleation theory and the computed line tensions in hydrophilic pores are in the range of experimental measurements [124].

4.2.4 Cooke and Deserno Model

Introduction

This is a three site solvent and electrostatic free model used to capture bilayer self assembly using simple pair-wise interactions.

Model

The model is made up of three sites, one head particle and two tail particles connected by finite non extensible linear elastic bonds. The size of the beads is fixed using a repulsive Weeks-Chandler-Andersen potential and the model flexibility controlled using an angle potential [125]. The authors present two forms of attractive interaction between the tails leading to unassisted self assembly, one a cosine potential while the other is a modified Lennard Jones potential [126], both potential forms having broad attractive wells.

Simulation and Results

Molecular Dynamics in the canonical ensemble were performed using the Langevin thermostat. Constant surface tension simulations were also carried out using the Andersen barostat. The model was validated by initially capturing the spontaneous self assembly of bilayers which agree with pre-assembled bilayers. The authors conclude that the range of attraction and not the exact functional form is essential for self assembly. The bilayer was quantified using order parameters, cross bilayer density profiles, overlap parameters, bending modulus, diffusion coefficient and flip-flop rates [125][126].

The model was also used to capture the transition from gel to liquid phase, and the lateral thermal area expansivity is in agreement with lecithin bilayers. By tuning the interaction between the tail particles the authors simulated a mixed system and were able to show the formation of domains and induce a budding process [125].

The model also produced spherical micelles, branched and unbranched worm like micelles and bilayers in qualitative agreement with the correct packing parameter order, by variation of the size of the head particle [127]. The model was also extended to investigate the coupling between local membrane curvature and the distribution of lipids with non cylindrical shapes; on a simulated vesicles, lipids with larger head groups were on the outside while the lipids with smaller head groups were on the inside monolayer. Though the authors find that the coupling between the shape and curvature is rather weak [127]. Using similar particles as the lipids, the authors constructed simple curvature inducing protein like structures and showed that these proteins can experience attractive interactions that arise purely as a result of membrane curvature [128].

4.2.5 Farago Model

Introduction

This is a rigid implicit solvent model consisting of three sites [114].

Model

The model consists of a rigid trimer of equal sized spheres glued together, one hydrophilic and two hydrophobic sites (Figure 4.3) interacting through the Lennard Jones potential. The interaction between hydrophobic tail and hydrophilic tails is purely repulsive while like atoms interact via an attractive LJ potential. The increased hardcore repulsion between hydrophilic particle 1 and hydrophobic particles 2 and 3 compensates for the absence of water and prevents the lipids escaping from the membrane plane.

Simulation and Results

Monte Carlo simulations at constant area and temperature were performed. Pre-assembled bilayers were simulated with the model validated by looking at the phase behaviour of the model with projected area. The model is able to show a solid to liquid transition when moving from a low projected area to a high projected area, the solid-phase is characterised

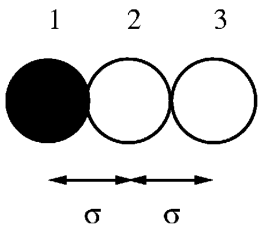

Figure 4.3: A schematic diagram of the lipid with the coloured circle representing the hydrophilic head and the open circles representing the hydrophobic tails. Figure from Farago [114].

by low mobility and hexagonal order and the fluid phase by increased mobility and the loss of hexagonal order. At larger projected areas the model develops a positive tension and shows the formation of pores and flip flop processes, and supports a possible conjecture between the two that has been alluded to in experiments. The spectrum of thermal undulations was in agreement with the Helfrich Hamiltonian [129], the bending modulus is in the range of biological membranes. The developed model was further utilised in the investigation of lipid-DNA mixtures as described in the next section [130][131].

4.2.6 Brannigan, Phillips and Brown Model

Introduction

This is a flexible implicit solvent model similar in spirit to the explicit solvent model by Goetz et al. [132].

Model

Individual lipids are represented as semi-flexible chains of five beads in this model [133]. Bead one is hydrophilic and the second bead acts as the interface between hydrophilic and hydrophobic components; the remaining beads constitute the hydrophobic tail (Figure 4.4). Three different pair potentials were applied: the repulsive core, an attractive tail-tail, and a longer range attractive interaction between tails and the interface that was crucial for the formation and stability of the bilayer. The angle potential used was similar to the Goetz et al. model [132] with no torsion potentials were applied.

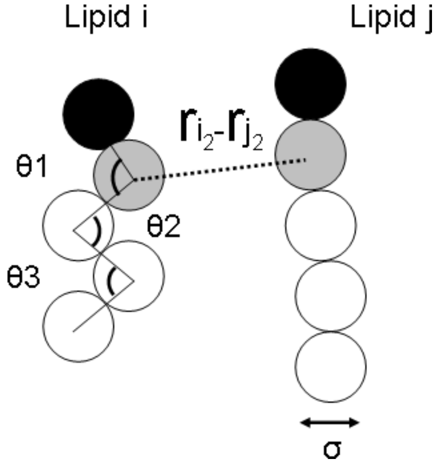

Figure 4.4: A schematic diagram showing the parameters used: polar heads are black, interface beads in gray and the hydrophobic tails white. Figure from Brannigan, Phillips and Brown [133].

Simulation and Results

Monte Carlo simulations under the constant tension constant volume conditions were performed. Self assembled bilayers in the fluid state were formed, validated by the diffusion constant, with area per lipid and compressibility in the biological range. The stress profile is in qualitative agreement with atomistic models and quantitative agreement with the explicit-solvent coarse grain model of Geotz et al. [132] and indicates that the interfacial beads are essential for bilayer stability. Bending rigidity is comparable to various biological systems depending on the flexibility of the model.

4.2.7 Sintes and Baumgartner Model

Introduction

This is an implicit solvent tethered model with no electrostatic interactions; it was developed to capture lipid mediated interactions between different shaped inclusions [134][135][136] that serve as crude models of proteins.

Model

The lipids are modelled by five spheres each representing a loose grouping of 3-4 heavy atoms; the head group is represented by a single sphere. The model interacts via the hard core potential; the proteins in the model are represented by hard cylinders [134][135] or truncated cone shapes [136]. Bilayer structure is maintained by tethering the head groups to artificial interfaces but allowing for diffusion along the x-y plane.

Simulation and Results

Monte Carlo simulations of lipid-protein systems were validated by calculating the mean area per lipid and average tilt angle obtained experimentally. The model was used to study the range of attraction and the influence of the shape of the protein on the attractive forces, identifying two types of lipid mediated attraction; a depletion induced attraction in the short range and a fluctuation induced attraction the medium to long range.

4.2.8 Ayton's Model

Introduction

The authors Ayton et al. [137] have attempted to interface microscopic models with macroscopic continuum models using a feedback loop. The idea is that properties such as diffusion which are macroscopic in nature are calculated using macroscopic models and they are input to the microscopic models. The process is reversed for microscopic models with them being used to calculate properties such as diffusion coefficients, and they being used as input to the continuum models. Below the details associated with a simple lipid model are given; the equilibrium model simulates a bilayer membrane whose stability is thermodynamic and is best described as a lyotropic concentration dependent liquid crystal.

Model

The model uses the Gay-Berne potential [25][42] consisting of ellipsoids for the lipids and spheres for the solvent. The cross interaction between ellipsoid particles and spherical particles was modified to promote de-mixing by omitting the attractive term in the potential. The membrane particles are ellipsoids of length 20 Å and diameter 5 Å. A dipolar charge distribution embedded within a sphere of diameter 5 Å was placed at one end of the Gay Berne particle making that end polar and the other end non-polar.

The charge distribution was constructed by placing a positive charge q+ = 0.705 and a corresponding negative charge q- =-0.705 at a distance of 0.5 Å from the centre of the terminal sphere such that the resulting dipole moment μ= ql was perpendicular to the long axis of the ellipsoid. The solvent particles were modelled with a charge distribution identical to the membrane particles. (Figure 4.5.)

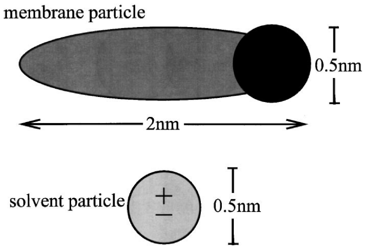

Figure 4.5: A lipid represented by an ellipsoid and solvent represented by a sphere used in Ayton's simulations. Figure from Ayton et al. [137]

Simulation and Results

A series of molecular dynamics equilibrium runs under NVT conditions were performed on a preassembled bilayer to find the upper density bound at which the membrane was in a gel state or solid phase. Below the low density bound the membrane decomposed to an isotropic mixture of solvent and membrane particles.

At higher densities the polar tips of the membrane particles are in direct contact with the polar solvent, likewise the opposing non polar end of the membrane particles tend to point away from the solvent. Simulation runs of 100 ps were performed to ensure that the membrane remained thermodynamically stable. At the highest density simulated most of the lipid molecules lie within the plane although some were found in the solvent layer as well. The stability of the lipid bilayer phase with respect to an isotropic mixture was determined using an order parameter, and fluidity confirmed using mean square displacement perpendicular to the director. The order parameter decreased with decreasing density indicating that the bilayer is less ordered at lower densities enabling the calculation of the density range in which the bilayer is stable.

The results from the equilibrium MD simulations above was then subject to Non Equilibrium Molecular Dynamics (NEMD) simulation using the the density range calculated above. The NEMD was used to calculate the density dependent modulus E.

The density dependent modulus E calculated using NEMD is used as the input for the continuum level simulation Material Point Method (MPM). This allows for the effective jump from microscopic time and spatial scales to macroscopic time and spatial scales.

4.2.9 Sun and Gezelter Model

Introduction

This is a simple molecular level model of rigid lipids, the model is based on the earlier model by Ayton et al. [137] (See Section 4.2.8) but with extensions to capture the role of the ordering of head groups in the formation of ripple phases [50].

Model

This is a two site lipid model, the hydrophobic tail represented by an ellipsoid and the polar head by a sphere at one end of the ellipsoid. The head is made polar by embedding a dipole perpendicular to the main axis with a fixed orientation. The solvent is the model proposed by Marrink et al. and is a non polar sphere representing four water molecules (in contrast to the Ayton model in Section 4.2.8 which uses a dipolar sphere for the solent) The model is simulated by the Gay Berne potential [25][42][43], Lennard Jones potential [24] and the dipolar potential [24].

Simulation and Results

Molecular dynamics simulations in the NPT ensemble were performed. Preassembled bilayers were constructed by allowing two separated monolayers to come together in a vacuum, then solvating the system. Variation of the head diameter, dipole strength and temperature and their effect on the bilayer structure and order were investigated. Increasing the head diameter increases the curvature and leads to the formation of symmetric and asymmetric ripple phases from flat bilayers. Larger head groups give area-per-lipids in the experimental range [138] with amplitude and wavelength of ripples in the experimental range [139]. Increasing the dipole strength increases the curvature of the monolayer and order and with increasing strength forms inverse micelles. The effect of temperature on order and bilayer structure was investigated; the effect is weak though increasing temperature destabilises the bilayer.

4.2.10 Loison and Schmid Model

Introduction

In this model a simple amphiphile to simulate a large number of multiple bilayers and their thermal fluctuations in a smectic A phase [140].

Model

The amphiphile is represented by a tetremer with two hydrophilic head beads and two hydrophobic tail beads covalently linked. The solvent bead is similar to the head bead and represents a loose collection of three water molecules but does not capture the electrostatic properties of water. All beads are identical and are differentiated by the interaction well depth. All particles interact through a soft potential with a LJ type repulsive part and a short ranged attractive interaction; similar particles have an attractive potential while dissimilar particles repel. The bonding is captured using the LJ and FENE potentials but no angle potentials exist in this model.

Simulation and Results

Molecular dynamics simulation in the NPT ensemble with the pressure maintained independently in the normal and tangential directions to the bilayer were carried out. Simulations containing a large amphiphile ratio self assemble into bilayers. Multiple bilayers were simulated, one containing up to fifteen bilayers and the other up to five, and order parameters, total density profiles and partial volume fractions analysed. The total density profile shows that the amphiphile and solvent layers are well separated. The distances between the membranes fluctuate and the in-plane correlation length extracted is in agreement with the discreet harmonic model and from that the compressibility modulus was calculated. The model was used to investigate the energetics of pore formation [141] and the molecular structure of the pores show that the amphiphiles situated in the rim of the pore reorient to shield the hydrophobic tails. The mean effective free energy of the pore and the line tension were calculated and found to be in agreement with mesoscopic line tension theory. Extending this work further, the authors looked at the effect of inserting a hydrophilic polymer into the system and identified formation of defects in the membrane system in relation to the length of the polymer strand [142].

4.2.11 Lenz Model

Introduction

The model [143] attempts to capture the liquid to gel transition; two versions of the model are proposed, one consisting of a solvent free model where the lipids are tethered to a surface to maintain bilayer structure as in other models [134][135][136] and the other consisting of "phantom" solvent beads. The tethered model prevents bilayer deformations and undulations

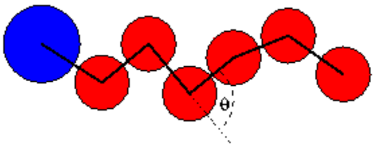

Figure 4.6: Schematic diagram of a Lenz lipid molecule . Figure from Lenz [143].

while the "phantom" bead model overcomes this limitation with minimum computational cost. Both models do not consider electrostatics in the solvent or lipid head groups.

Model

The model [143] consists of two subparts, the lipid and the solvent. The lipid consists of six tail beads and one slightly larger head bead bonded together by a finite extensible nonlinear elastic type spring potential (Figure 4.6). All beads interact via the truncated 12-6 LJ potential. The interaction parameters were chosen to crudely represent a fully saturated lipids of 16 to 18 carbon length acyl chains. The solvent model is represented by beads identical to the head particle and they do not interact with each other.

Simulation and Results

Constant pressure Monte Carlo simulations were carried out. The model reproduces qualitatively the gel and liquid crystal bilayer phases, correctly describing the strong decrease of the lipid tail ordering and bilayer thickness, as well as the increase of the area per lipid. The translational and conformational degrees of freedom need to be modelled correctly for qualitative reproduction of the generic phase behaviour. The model shows that the solvent does not influence the existence and location of the transition but is required for the stability of the liquid crystal phase.

4.2.12 Smit Model

Introduction

The Smit model is an early off-lattice simulation of surfactant systems that shows the spontaneous formation of monolayers and micelles based on the Lennard Jones potential [144] [145].

Model

The model is made of two types of particles represented by spheres, oil-like and water-like particles. Oil-like and water-like particles are joined together by harmonic springs to form surfactant molecules. Particles interact via the Lennard Jones potential only with like particles attracting and while the unlike ones repel: the model does not consider angle bending or any electrostatic potentials.

Simulation and Results

The system was simulated under the canonical ensemble using molecular dynamics. The simulated system consisted of oil, water and surfactants with the concentration of the surfactants varied. Depending on the concentration of the surfactants the system is able to simulate the formation of micelles and monolayers and shows a depletion layer at the water interface. With increasing surfactant concentration the interfacial tension between the oil and water layer decreases in line with experimental observations. The system was able to demonstrate that self assembly may occur in the absence of hydrogen bonding.

Using a similar system of surfactants in water the authors demonstrate the spontaneous formation of several micelles and follow the time evolution of the micelles. The model shows that short range repulsion between the head groups is sufficient for the formation of micelles [146].

4.2.13 Goetz and Lipowsky Model

Introduction

The model [132] is essentially a Lennard Jones (LJ) fluids consisting of 'solvent' and surfactant molecules; the models simulate both the self assembly process and bilayer properties. The surfactant molecule consists of a head group and one or more tails which are modelled by chains of particles interconnected by harmonic bond potentials. In addition it incorporates the bending stiffness of these chains.

Model

The model is built up from three types of *particle* which can be hydrophilic or hydrophobic. One particle in the tail corresponds to 3-4 CH_2 groups, similar in style to the later developed Marrink model [20] with each solvent sphere representing two water molecules and identical to the head particle. All particles have the same mass and are taken to be spherical

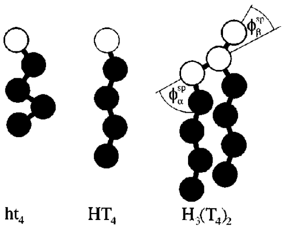

Figure 4.7: Types of surfactants proposed by Lipowsky: ht_4, HT_4 and $H_3(T_4)_2$. The white particles [(h) and (H)] represent the head group, the black particles [(t) or (T)] the tail. In particular, the $H_3(T_4)_2$ molecules can be taken as crude lipid models. Capital letters indicate that the chains are flexible due to the presence of a bending potential. Figure from [132].

(Figure 4.7). The interaction between one hydrophobic and one hydrophilic particle is modelled by a repulsive softcore potential. All other interactions are modelled using the Lennard Jones (LJ) potential. In the model the head groups and tail particles are connected by a harmonic bond potential to form surfactant molecules. The bending stiffness between each three consecutively bonded particles is modelled using an angle potential making the surfactants semi-flexible. No dihedral potentials or electrostatics are considered in this model. Different surfactant molecules constructed by combining different numbers of particles including $H_3(T_4)_2$ surfactants as crude models of a lipid molecule.

Simulation and Results

Monte Carlo (MC) simulation was done to study the self assembly behaviour, by performing several simulations at different surfactant concentrations for all three types of surfactant molecules. The type of aggregates formed depend on the concentration. The model shows that the binary LJ fluids studied here leads to the formation of micelles and bilayers. Fluidity in the membrane system is confirmed by both lateral and trans membrane (flip flop) diffusion. It also identifies tensionless bilayer states as appropriate for real bilayers. For such tensionless states, the stress profile was analysed and found to be reasonable apart from the unphysical negative peak at the bilayer midplane. The area-per-lipid is smaller than biological systems but the area compressibility is of the same order magnitude. The effect of chain length and stiffness on the lipid structure was also investigated.

Extensions to the Goetz and Lipowsky model

Goetz, [147] extending the previous work on lipids, [132] set about extracting a value for the bending rigidity κ from shape fluctuations of a bilayer. The authors derive a simple relation to calculate the bending rigidity from the height of the bilayer and compressibility, without analysis of shape fluctuations. The model provides for the first time an explicit connection between movement of individual molecules and elastic surface behaviour.

Otter, [148] using the rigid molecule $ht4$ proposed by Goetz and Lipowsky [132], calculated the bending rigidity κ using equilibrium and non equilibrium methods. The authors find that the Helfrich expression for the bending free energy is valid only for small amplitude fluctuations. Non-equilibrium methods were found to be appropriate for the calculation of the bending rigidity for bilayers with large amplitudes, equilibrium methods were found to be appropriate for small fluctuations. The calculated area-per-lipid was considerably larger than those reported by Goetz and Lipowsky [132]; differences were also observed in the compressibility modulus from the Goetz model. However, the bending rigidity compared favourably with the Goetz model. Otter, [149] also finds that the calculated area compressibility in bilayers containing undulations is smaller than the expected value, due to the neglect of the undulations in calculating the compressibility from area fluctuations.

Imparato et al. [150][151] studied bilayers with two components to determine the dependence of their elastic properties on membrane composition. The two amphiphiles considered (A and B) differ in their tail lengths, allowing for the focus on the contribution from the mismatch of amphiphile tails (Figure 4.8). All sites have the same value for σ and ϵ except when the interacting particle is H_B, where the value of ϵ was different. This was done in order to study the behaviour of the smaller molecule B, which could be considered as cholesterol or fatty acids. The work finds that there is an increase in the lateral diffusion on addition of the

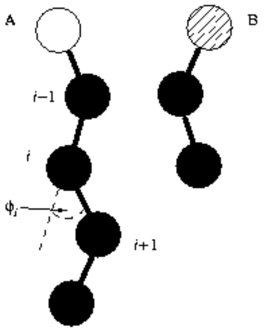

Figure 4.8: The two components A and B that make up the bilayer in the Imparato model. Figure from Imparato et al. [151]

shorter amphiphile due to the increase in mean free volume, supporting the assumption that diffusion is driven by free volume fluctuations. The model was able to simulate the flip-flop process concluding that the process is an activated one, and calculated the activation energy for such an event. The rate of flip-flop decreases with increasing chain stiffness, though the characteristic time for the process is only weakly dependent on it. The model was able to predict the bending rigidity taking into account all types of fluctuations, including different tail conformations, curvature fluctuations, molecular protrusions and composition fluctuations arising from lateral diffusion. The study also investigated the relationship between bending rigidity and lateral diffusion in relation to the mole fraction of amphiphile B. The authors find that the membrane composition and inverse bending rigidity do not follow a simple lever rule and observe a decrease in bending rigidity on addition of the second amphiphile due to the increase in molecular protrusions. This work also highlights that there is a non-monotonic functional dependence of the intrinsic area on membrane composition.

4.2.14 Stevens Model

Introduction

This is a bead and spring double-tailed lipid model in the spirit of the Goetz et al.[132] models. It is used to study the mechanism of membrane fusion by a biased method of forcing two vesicles to merge,[152] as well as structural and dynamic behaviour of phosphatidylcholine type neutral lipids [98].

Model

Three types of beads are in the system. The tail is composed of hydrophobic beads of type T and the head is composed of hydrophilic beads of type H. The solvent is represented by a single bead S; the bead S and H types are equivalent in this work. The beads interact via the LJ potential with the same interaction energy with the cutoff determining if the interactions are attractive or purely repulsive. The bond potential is the sum of the attractive finitely extensible nonlinear elastic (FENE) potential and the purely repulsive LJ potential. Angle terms are included via a harmonic potential. A double bond is modelled using a kink in the tail chain. The model does not consider explicitly the effect of electrostatics.

Simulation and Results

Molecular dynamics simulations were performed in the NVT and NPT ensembles. The system self-assembled into a bilayer starting from a random configuration. The system was shown to simulate not only the liquid crystal phases but also the gel phase. The area per lipid variation with temperature and tail length correspond favourably with experimental trends. The melting temperature increases with increasing number of carbon atoms in the lipid tail. The bending modulus lies in the experimental range of phosphatidylcholine lipids and also shows the expected trends on varying the tail length. Double bonds in the hydrocarbon tail were modelled by kinks, increasing the disorder and area per lipid in the bilayer, again in line with experimental observations. The gel phase shows lower diffusion rates than the liquid crystal phase, increased tilt relative to the liquid crystal phase and increased bilayer thickness, decreased area per lipid, and local triangular lattice ordering.

Using the same model, the author simulated a binary mixture of lipids by varying their tail lengths. The author was able to capture the formation of domains using such a simple binary mixture. The longer tail lipids form the gel phase and the shorter tailed lipids the liquid crystal phase and their complementary coexistence is in line with NMR experiments and confirmed by two dimensional RDFs of the bilayer. The work also sheds light on how two lipids may adopt their structures at boundaries to reduce the cost of boundary mismatches [153].

4.2.15 Izvekov and Voth Model

Introduction

The model [112] uses an approach called ”multiscale coarse graining” (MS-CG) because the potential parameters are systematically derived from atomistic level interactions. The basis of MS-CG is a new method for force matching, which can determine a pairwise effective force field from a given trajectory and force data regardless of their origin. The MS-CG method was used to obtain a CG model of the DMPC lipid bilayer. The force fields describing the interactions between all CG sites in the system (water-water, water-DMPC and DMPC-DMPC) were obtained from atomistic trajectory and force field data from a single MD simulation of the bilayer. This approach provides a systematic way to coarse-grain underlying atomistic force fields and it utilises only those atomistic interactions as input. The atomistic MD simulation was carried out with 64 DMPC and 1312 water molecules under NPT conditions.

Model

The coarse grain lipid molecule contains 13 interaction sites. The water molecules were mapped into a single CG interaction site associated with their geometrical centre rather than their centre of mass. For simplicity the intra lipid CG sites were linked by harmonic bonds. The non-bounded sites interact via non-bonded potentials, helping the lipid maintain its molecular geometry. (Figure 4.9.)

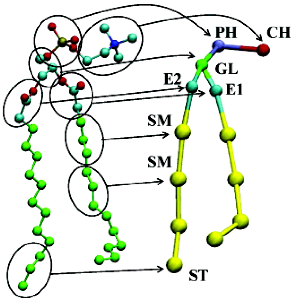

Figure 4.9: Atomistic (left) and coarse grain (right) model of DMPC molecule proposed by Izvekov and Voth. Figure from Izvekov and Voth [112].

Simulation and Results

For the purpose of force matching, an atomistic MD simulation was carried out. The coarse-grain MS-CG system was the same size as the atomistic one. The MS-CG system was simulated under NVT conditions with 64 DMPC molecules and 1312 water sites. The coulombic interactions are subsumed into effective non-bonded terms, because of screening effects from the environment which cause the force matching algorithm to yield effective charges on the charged groups in the lipids which are smaller than those inferred from atomistic charges. The MS-CG method reproduces correctly the structural properties of the lipid bilayer. The solvation numbers of both the atomistic and MS-CG simulations agree as do the water density profiles. The gain in speed is about fifty times compared to the underlying atomistic simulation of the same size. The gains will be larger for larger systems as MS-CG only contains short range interactions making it highly efficient.

4.2.16 Shelley and Klein Model

Introduction

The model proposed by Shelley et al. [102] was subsequently further extended to study several other biological problems [154][155][156]. The method involves selecting simplified representation of water, alkanes and dimyristoylposphatidylcholine (DMPC) molecules and developing for each type of molecule a model that mimics key physical and structural properties from experiments.

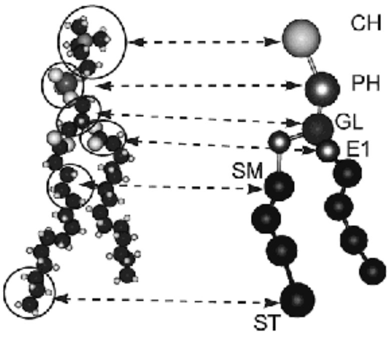

Figure 4.10: Atomistic (left) and the corresponding coarse grain (right) model of DMPC molecule from Shelley et al. Figure from Shelley et al. [102].

Model

Most of the detailed nature of water was eliminated leaving a representation that provides an efficient medium for other molecules to exist and interact with. Specifically, it was required that the model for water carry momentum so its behaviour can be consistent with hydrodynamics, have the correct density at the desired temperature, and maintain a liquid/vapour interface over the desired temperature. A symmetric site referred to as W represents a loose grouping of three water molecules. Interactions between W sites are described by a LJ 6-4 potential. Alkane portions of the molecule are represented by sites representing triples of carbon atoms and their accompanying hydrogen atoms. $(CH_2)_3$ is referred to as SM sites, while $(CH_2)_2 - CH_3$ is referred to as ST sites. Harmonic bond length potentials were used to link the SM and ST sites to form the hydrocarbon chains. The force constants and the distance of minimum potential was adjusted to reproduce the average and standard deviation of the corresponding distance obtained from atomistic simulations. Bond angle potentials are used to maintain the proper chain stiffness and overall length. For hydrocarbons a cosine angle potential was used. Cosine angle potential parameters were adjusted to reproduce the average and standard deviation of the corresponding angle obtained by atomistic simulation. Non bonded interactions were modelled using LJ 9-6 potential with no non-bonded interactions for directly bonded groups. The interaction of W with SM and ST was also modelled using the LJ 9-6 potential.

Single spherically symmetric sites were used to represent each of the choline (CH) phosphate (PH), glycerol backbone (GL namely CH2-CH-CH2) and ester groups (E1 and E2, O2CCH2). The hydrocarbon chain was modelled using the alkane molecule above. The sites in the hydrocarbon chain were linked together with harmonic bond length potentials. The choline and phosphate groups were assigned charges of +e and -e and a dielectric constant of 78 for all electrostatic interactions. Interactions among CH, PH, GL and E groups as well as between these groups and the W, SM and ST groups were parameterised on the basis of radial distribution functions obtained from atomistic simulations (Figure 4.10).

Simulation and Results

The models were tested at 303.15 K in the NVT ensemble using MC simulation. Neighbour lists were used to speed up the calculation and "tinfoil" Ewald periodic boundary conditions were used to treat electrostatic interactions. 50 DMPC and 428 W sites were simulated using the MC method.

The CG simulation was able to reproduce the overall structure of the bilayer produced by the atomistic simulation in spite of leaving out considerable amounts of atomic level information. A possible defect is the high pressures leading to higher water densities. The model semi quantitatively reproduces the density profile of an aqueous DMPC bilayer. A simulation starting from random configuration leads to a bilayer with a few defects demonstrating the efficiency of the CG model.

Extensions to Shelley's model

Shelley [102] extended the model to simulate DHPC by including an additional SM site (representing $(CH_2)_3$) in the hydrocarbon chain. Three studies were performed, one of the anaesthetic halothane in a DMPC bilayer and two studies of self assembly of phospholipid liquid-crystalline phases, a lamellar phase and reverse hexagonal phase. Halothane was represented by H and was introduced into the preassembled bilayer from [102] maintaining a 2:1 ratio of phospholipid to halothane. MC simulations were done, though several adjustments were required to bring the model into qualitative agreement with the atomistic study. This indicates that the current methodology is not sufficient to provide predictive models in the absence of careful parameterisation using extensive data.

When self assembly was carried out using MC simulation there were defects in the bilayer formed; when the simulation was switched to MD the defects healed and a defect free membrane resulted. DHPC and nonane were used to simulate formation of a reverse hexagonal phase. MC failed to assemble the system into a reasonable configuration while switching to MD produced some collective organisation. Simulation was continued by increasing the size of the simulation box. Once again, MD showed order while MC showed only some local aggregation. MD shows that some columns that form in this phase appeared to be polygonal while they adopt a bilayer like structure locally. The self assembly process for this phase initially proceeds through local aggregation, then bilayers, and finally to form a reverse hexagonal structure. This study suggests that MD is more efficient for modelling collective motion within dense molecular structures.

Lopez [155], using the model by Shelley [102], made a comparison of the electron density profile of the CG model with that of an all-atom model. It shows that a CG description of the lipid and reduced number of sites can provide quantitatively similar results to those of atomistic simulations. This indicates the applicability of CG methods to large systems. In this simulation the electrostatic interactions were not cut off since they were found to contribute significantly to system properties. Comparison of the diffusion constant with that of atomistic simulation shows that it is two orders of magnitude faster. Reduction of the interaction sites from 118 in the atomistic simulation to 13 in the CG model reduces the calculation overhead by another two orders of magnitude.

4.2.17 Marrink Model

Introduction

A model that possibly the most widely used coarse grain model though it is considered to be accurate only to a semi quantitative level [20].

Model

Four-to-one mapping of heavy atoms on average represented by a single interaction site. To better capture the different types of atoms found in lipids four sub types were defined; polar, non-polar, apolar and charged types. Solvent represented by a LJ sphere and represents four water molecules, lacks solvent electrostatics and dielectric screening artificially imposed using a dielectric constant of ϵ_r=20. The four-to-one mapping results in a 12 site DPPC lipid model with the polar head group charges represented by charged LJ spheres and interact via the shifted Coulombic potential while all other particles interact via the LJ potential.

Simulation and Results

The model is able to show spontaneous self assembly and the final results are consistent with all-atom simulations and experimental measurements of DPPC. Mechanical properties such as area compressibly, bending modulus and line tension were computed and were within an order of magnitude to experimental work. Dynamical properties such as the permeation of water across the bilayer was also simulated and were of the same magnitude as experimentally values for DPPC vesicles.

The model was also used to study the formation of a gel phase by lowering of the temperature as well as the phase transition from a multi lamellar to inverse hexagonal phase with increasing temperature or lowering of solvent concentration. The model particles were used to build a coarse grain model of DPC lipids and was able to simulate the formation of micelles. The Marrink model has in recent years been extended to model coarse grain DNA, proteins and dendrimers [9].

4.2.18 Orsi Model

Introduction

A ten site coarse grain model using the Gay Berne potential with explicit charges for the head group immersed in an explicit solvent model [51][157][158].

Model

The model consists of ten sites representing a DMPC lipid. The polar head choline and phosphate are represented by LJ spheres with embedded charges. The glycerol and hydrocarbon regions are represented as soft ellipsoids using the Gay Berne potential [25]. The glycerol region ellipsoids have embedded point dipoles to capture the charge distribution in this region while the two hydrocarbon tails have three neutral ellipsoids each; with each ellipsoid representing a grouping of four consecutive methyl groups. The solvent is represented by single interaction sites with a one-to-one mapping to water; the soft sticky dipole model [157]. The sites interact via the Lennard Jones and Gay Berne potentials while electrostatic interactions are handled using charge-charge, charge-dipole and dipole-dipole potentials with switching and shifting functions to avoid any artifacts at the cut-off.

Simulation and Results

The model spontaneously self assembles to form defect free bilayers under constant NPT conditions. The reduced interaction sites enable the model to be approximately two orders of magnitude faster than all-atom models of the same size. The model was parameterised to reproduce the experimental area-per-lipid, volume-per-lipid and order parameters. The resulting bilayer is then used for predicting many other lipid structural and dynamic properties. The electron density profile is in good agreement with experimental measurements and the model predicts a head dipole of 16.1 D in close agreement with experimental values for DPPC. Area compressibility, volume compressibility, bending modulus, in agreement with experimental measurements. Lateral pressure profile in agreement with all-atom models and membrane curvature elasticity in agreement with theoretical predictions were also predicted by this model. Both microscopic and macroscopic diffusion in agreement with experiment and the diffusive behaviour is in agreement with the free volume theory. Electrical properties such as the membrane dipole potential and associated transbilayer electric field were investigated using this model

The model has now been extended to model DOPC lipids as well as to study the permeation of β blockers and steroid hormones across the bilayer using multi resolution models; where the bilayer is represented using the coarse grain model described above while the permeants are represented using all-atom models [159][160][161].

4.2.19 Summary of Lipid Models

The above review has highlighted some of the many coarse grain models that have been published in the literature. The review has focussed on highly simple lipid models represented by single sites or by a small number of sites representing amphiphilic molecules rather than a particular class of lipids. These models have used a variety of potentials and techniques using both implicit and explicit solvent in an attempt to capture the phase behaviour of lipids. The model developed in this work also lies at this level of granularity and is similar to the models developed by Ayton et al. [137] and Gezelter et al. [50]. A comparison of the model developed in this work and the models by Ayton et al. [137] and Gezelter et al. [50] is given in Chapter 5 to put the developed model in perspective.

4.3 Coarse Graining of DNA

In the classical DNA double helix described Watson and Crick, a flexible, ladder like structure is formed with the polymer wrapped around an imaginary central axis (See Figure 4.11). The two rails of the ladder consist of alternating sugar and phosphate units; the rungs of the ladder consist of nitrogenous base pairs held together by hydrogen bonds [5]. The two strands run anti-parallel to each other. Adenine (A) pairs with thymine (T) and guanine (G) pairs with cytosine (C) (See Figure 4.11). The spaces formed between the helical backbone and imaginary cylinder that encloses the DNA are termed major and minor groves; they have different dimensions because of the sugar-based linkages' asymmetry with respect to the base-pair plane [10]. There are three main canonical forms of DNA, though numerous others have been identified. The B-DNA is considered the dominant form under physiological conditions [12].

DNA is highly charged and interacts strongly with its solvent environment and other

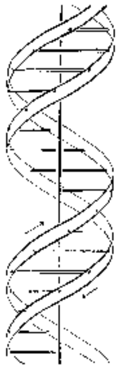

Figure 4.11: Schematic diagram of DNA double helix, the two ribbons represent the sugar phosphate backbone and the horizontal lines the base pairs holding the two strands together. Figure from Watson and Crick [11].

solutes over long distances [13]. The large surface to volume ratio means that solvation effects also play a important role in the DNA structure. Solvent may also pay a role in stabilising the helix; this has been confirmed by simulations of DNA helix *in vacuo* [162], so both solvation effects and electrostatics need to be treated properly by simulation methods [163]. More and more evidence is being presented today that the counterions in solution interact in a complex sequence and structure specific manner [18]. However the long life-spans of bound counterions, the conformational sampling issues, and force field balance issues make it a difficult issue to handle [164]. Some DNA simulations neglect solvent effects and use implicit solvent methods to increase the speed of computation; one such example is the use of generalised Born implicit solvent model. [18]. However given the spline of hydration seen in the minor groove and the significant interaction of DNA with water, it is likely that some explicit solvent may be required. One method of overcoming the *waste* in computation time simulating water is the use of a minimal explicit solvent model, where a narrow layer of explicit water is added around the DNA molecule giving a hydration of $\sim$ 15-20 waters per base pair with the bulk solvent being treated implicitly. This level of hydration is lower than normal hydration conditions for DNA and would be expected to represent the DNA under dehydrating conditions [18][165][166].

DNA has many levels of structural hierarchy, from length scales of nanometres for several base pairs to micrometres for several thousand base pairs.

DNA can be regarded as a one dimensional polymeric chain and thus theoretical treatment of its thermodynamic properties naturally originate from polymer physics. DNA, however, has a non negligible cross section which leads to high flexural rigidity (and concomitantly a large persistence length); it is therefore often referred to as a "stiff" polymer. The worm like chain (WLC) model includes flexural rigidity via a quadratic bending potential. Many of the present day models are variations of this. When it was realised that the intrinsic twists of DNA could not be neglected, several models were introduced that include both bend and twist rigidity. None of these models thus far include the chirality of the molecule, that would allow an asymmetry in the tension upon over-twist or under-twist, although some attempts in that direction have been made with models of fluctuating ribbons. The low level of parameters in these continuum level descriptions and the concomitant low level of detail make them generally applicable to length scales on or far beyond the average pitch length. Solvent effects are treated in only an average fashion and excluded volume effects are generally neglected.

For atomic-level detail atomistic force fields are used. The all-atom MD simulations can study 10-20 base pair sequences and are limited to timescales of the order of nanoseconds and length scales that are well below the persistence length. This means atomistic simulations yield local mechanical properties.

One way to bridge the gap between continuum level and atomistic simulations is particle based models. Each "particle" (bead) represents a group of atoms; these models are roughly an order of magnitude faster than atomistic models but still contain the essential physics of the system under study. By representing a group of atoms by beads and connecting them by springs, the degrees of freedom and hence the computational cost may be lowered tremendously.

Solvent molecules can be included in molecular dynamics simulations but at high computational cost. For simulation of multiple or longer DNA duplexes, or for longer timescales, it is computationally essential to decrease the number of degrees of freedom drastically [23]. This maybe done by simplifying the solvent as well as the model itself.

Many simulations of simple DNA models have been already attempted, with the level of simplification determined by the properties to be studied using the model; several such simplified models are given below.

4.3.1 Huertas Model

Introduction

This is a partially flexible, double-helical model [167] describing the conformational and dynamic properties of DNA. The model is partially flexible, enabling it to experience deformations and internal motions such as bending and twisting. The model is computationally attractive as it only requires a few parameters to be specified, but the helical symmetry is pre-imposed and thus the model can at best describe physics of length scale larger than the helical pitch.

Model

In the model each nucleotide is represented by one element (bead) and the known geometrical features of the double helix are incorporated in its equilibrium conformation. Each bead is bonded to seven other springs that connect to four beads in the same strand and three beads in the opposite strand, thus stabilising the helix. These interactions keep the two strands intertwined and may be assigned to base-pair stacking. The model assumes that all four types of nucleotides have the same hydrodynamic radius, σ. The springs are stiff but allow for some flexibility and internal motion. The molecule used contains up-to twenty base pairs. The backbone is described by twelve beads per helical turn, interconnected by harmonic springs. Solvent particles were not explicitly considered in this model.

Simulation and Results

Brownian dynamics simulations were used to sample the conformational space and monitor the overall and internal dynamics of short DNA pieces. From Brownian trajectories, calculation of the dimensions of the helix and estimation of its persistence length was made. The translational diffusion coefficient and various rotational relaxation times including both overall rotation and internal motion were calculated. Even with limited parameterisation the calculated properties agree well with experimental results. The model is able to access time scales in the range of milliseconds.

Other Single Site Models

Fang Hai-Ping et al. [168] use a single-site per nucleotide model to investigate the role of hydrogen bonding and base stacking interactions in the stretching to B-DNA like structure from S-DNA, and, unzipping of DNA to form single stranded DNA. Savelyev and Papoian [169] use atomistic force fields to derive an effective hamiltonian for a CG DNA model using renormalisation group ideas. Both Fang [168] and Papoian [169] models lack electrostatics and solvent models.

Trovato and Tozzini have developed a single bead model [170] in an implicit solvent for simulating nano circles and plasmids with no explicit charges. As in the above models each bead represents a nucleotide placed at the location of the phosphate group. Each bead is bonded to the the adjacent beads in the same strand to form a poly nucleotide chain and then bonded to the opposite poly nucleotide chain to form a double helix. Microsecond timescale

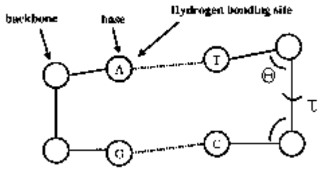

Figure 4.12: Schematic diagram of DNA dimer, each backbone and base is represented as a sphere. Figure from Drukker et al. [23]

simulations as well as simulation of the small complete plasmid have been performed. The model allows for the investigation of the transition from a double to a single strand, as well as a denaturation bubble with results in agreement with experimental measurements. The model reproduces correctly the structural parameters such as rise and twist angles.

4.3.2 Drukker Model

Introduction

This is a two-site implicit solvent model with no treatment of electrostatics that aims to capture the dynamic behaviour of DNA including: hydrogen bonding, opening of the duplex, and the fraying of the ends of the duplex.

Model

The model building blocks are single DNA strands, each strand consisting of nucleotides each of which is represented by two sites. The first site is the backbone site representing the sugar and phosphate groups while the second site represents one of the four bases A, T, C or G, see Figure 4.12. The backbone sites are connected by bonds representing the backbone structure, and each base is bonded to its corresponding backbone site; to maintain the helix structure additional bonds were added along the next nearest neighbours mimicking the long range repulsion along the backbone.

Hydrogen bonding is the most important interaction and is dependent on both the distance and angle of the two nitrogen bases.

Results and Simulation

Molecular dynamics simulations in the microsecond time range become accessible for small DNA decamer duplexes. Simulations look at the melting of helices made up of either C-G or A-T. The study also looked at the effect of mismatches, length and the effect of helical structure on the melting behaviour of duplexes.

The melting temperature and its variation with base pair composition, length of DNA and the presence of mismatches, is in general accord with experimental or thermodynamic models.

Other Two Site Models

Dorfman and Kenward [171] simulated a solvent and electrostatics free two site model using Brownian dynamics. The backbone is represented by a sphere and the base by the other sphere. No differentiation of base sites are done, treating all base sites as identical. The model was used for the study of: hairpin formation, effect of mismatches on the binding section of the hairpin, effect of temperature on melting of hairpins, and kinetics of hairpin formation.

4.3.3 Mergell's Model

Introduction

This model [21] proposes ”DNA like” molecules with simplified interactions resolved at the base or base pair level. Interactions between neighbouring bases (base pairs) were studied using the Gay Berne (GB) potential. The sugar phosphate backbones are reduced to semi-rigid springs connecting the edges of the disks/ellipsoids, see Figure 4.13.

Using MC simulations local stacking and global helical properties are explored as functions of model parameters. The model allows for the representation of equilibrium structure and linear response, while preserving the possibility of local transitions due to external forces.

Qualitatively the general considerations suggest a B-DNA like ground state and transition to a skewed ladder under stretching force. The observed conformation is a balance between base stacking interactions and backbone conformation. Roll, tilt and rise depend on base stacking while twist depends on the constraints imposed by a rigid backbone. Slide and shift movements are sequence dependent and are not considered in the model.

Model

The DNA is viewed as a twisted ladder in which all bars lie in one plane; this is used to illustrate the relation between local parameters such as slide, roll, etc [172][5] and overall shape of the resulting helix.

Figure 4.13: Illustrative diagram of the underlying idea; the base pairs are shown as ellipsoids and the backbone is treated as semi-rigid springs connecting the edges of the ellipsoid. Figure from Mergell et al. [21].

The DNA molecule is described as a stack of thin, rigid ellipsoids. The size of the ellipsoids was chosen to reproduce experimental results as well as for perfect stacking, see Figure 4.13. The attraction and excluded volume between base pairs is modelled using the GB potential. The electrostatic interactions between neighbouring base pairs are neglected, since stacking interactions dominate under physiological conditions. The sugar-phosphate backbone is nearly inextensible. The backbone is represented by elastic springs with non zero spring lengths.

No minor or major grooves exist in the model, and due to the ellipsoidal shape of the base pair and the fact that the internal base-pair degrees of freedom cannot relax, a non B-DNA like ground state is obtained where roll and slide motions are involved. The right handedness of the model is due to excluded volume interactions between bases and backbone.

There are only three free parameters: the GB energy depth ϵ, which controls the stacking interaction, the spring constant κ which controls torsional rigidity, and the depth of the ellipsoid which influences the fluctuations of the bending angles. The other parameters are fixed to reproduce experimental values for B-DNA.

Simulation and results

All interactions were local and studied using MC simulation. Trial moves consisted of local displacements and rotations of one ellipsoid by small amplitudes as well as global moves to change positions and orientation of large parts of the chain. Two randomly chosen points in the molecule define the axis of rotation. It was found that global moves significantly improve the simulation efficiency.

The equilibrium structure of the chain was studied as a function of model parameters. Various conformations were minimised to verify that the proposed Hamiltonian prefers the B-DNA form. Only local interactions were considered, with calculation restricted to two base pairs. Three local minima were identified: stacked-twisted, a skewed ladder, and unwound

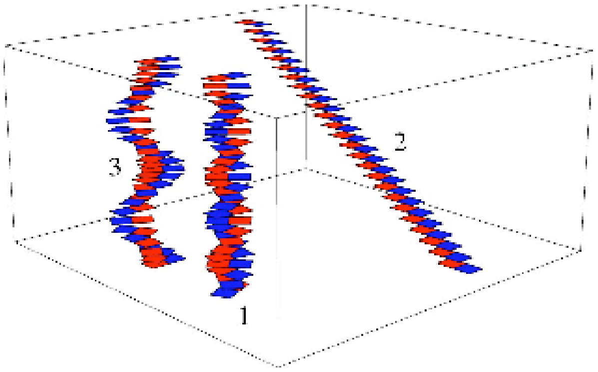

Figure 4.14: DNA geometry (1) Twisted ladder (2) Skewed ladder (3) Helix. Figure from Mergell et al. [21].

.

helix (See Figure 4.14).

Without an external pulling force the global minimum is found to be the stacked-twisted conformation. Rise movements were found to be independent of all three variables used. MC was also used to study the effects of thermal fluctuations. With increasing temperature, it was seen that twist angles decrease while the mean value of rise increases. Probability distribution functions show rise and twist motions that do not follow a Gaussian behaviour, with deviation mainly determined by κ. The GB energy depth ϵ was also found to play a major role in the bending persistence length while κ does not. It was also found that transverse slide and shift fluctuations contribute to the longitudinal stretching modulus of the chain in this model.

The DNA when stretched increases its contour length by a factor of 1.8 to the normal length to form a conformation known as S-DNA. The exact structure of S-DNA is still under discussion. The model identifies two possible transition paths to S-DNA. Either the chain untwists and unstacks resulting in an untwisted ladder 1.8 times the equilibrium length, or the chain untwists and the base pairs slide against each other resulting in a skewed ladder with the same length. The latter could be energetically more favoured since it retains the possibility of partially conserving base-pair stacking. However, in simulations of stretching both scenario have been observed depending on which strand was pulled.

The transition from B-DNA to S-DNA mainly depends on the GB energy depth ϵ by controlling the stacking energy. An optimum value for all three variables in the B to S transition as well as the effect on rise, slide, and shift etc. as a function of applied forces was also determined using this model. The model produces the correct persistence lengths for

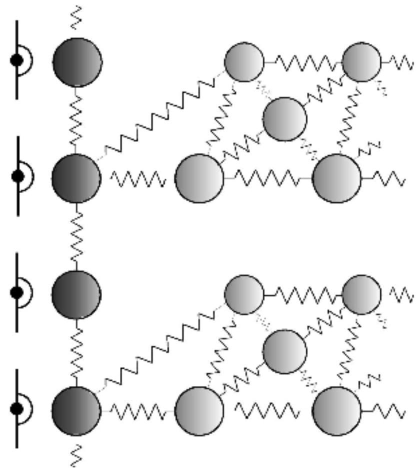

Figure 4.15: Schematic diagram of the backbone connectivity with base-pairs (light beads) and backbone (dark beads). Figure from [22].

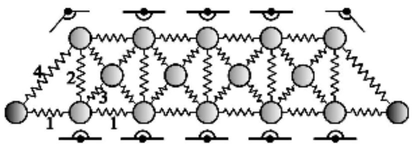

Figure 4.16: Schematic picture of the base pair (light beads) with the connection to the backbone(dark beads). The symbols above and below the base pair denote the angular potential. Figure from Tepper and Voth [22].

the B-DNA and gives the correct mean values for the base pair step parameters known by X-ray diffraction studies. The model predicts a S-DNA structure with partial base stacking; this is in good agreement with atomistic simulations.

4.3.4 Tepper and Voth Model

Introduction

This is a simple bead and spring model with explicit solvent but no electrostatics where hydrophilic/hydrophobic interactions drive helix formation.

Model

The complete model consists of three types of beads, A (backbone), B (base pair) and C (solvent), see Figure 4.15. The backbone of the DNA is represented by two bead while the nitrogen base pairs are represented by fourteen beads with the base pairs being considered coplanar see Figure 4.16. The solvent used in the model is the "coarse grained water" proposed by Goetz and Lipowsky et al. [132] The beads are held together using harmonic potentials while harmonic angle potentials along the backbone and in the base pairs represent the long range phosphate-phosphate repulsion and maintain planarity in the base pairs.

The model uses the shifted form of the Lennard Jones potential for attractive interactions and the r^{-12} term of the Lennard Jones potential for repulsive interactions between the backbone and the basepairs, and the solvent and the base pairs. The repulsive interactions drive the helix formation in this model. To study the effect of particle size, the r^{-12} potential was replaced by the repulsive Weeks-Chandler-Andersen (WCA) repulsive potential.

Simulation and Results

Canonical (NVT) simulations show a straight ladder consisting of twelve base pairs forming a helix in around 100 ps with one complete helical turn. The helix forms both right handed and left handed helices due to the absence of chirlaity in the sugar group.

Further simulations were done to test the effect the mass distribution over the solute and the effect of interactions with the solvent, as well as the effect of base pair interactions on helix stability.

The model is able to simulate many dynamic properties, correctly reproducing experimental results that show folding is always initiated at the ends. The model also shows helical properties like twist, bending and stretching rigidity as well as structural properties like helical pitch. The off-centre placement of the base pair results in the formation of a major and minor groove, see Figure 4.16.

The model finds that the spontaneous helix formation is caused by the interplay of three contributions, hydrophobic/hydrophilic interactions between base pairs, backbone and solvent; phosphate-phosphate repulsion along the backbone; and base pair stacking energy. The final structure is mainly dictated by the geometrical dimensions and connection topology of the base pairs and backbone see Figure 4.17.

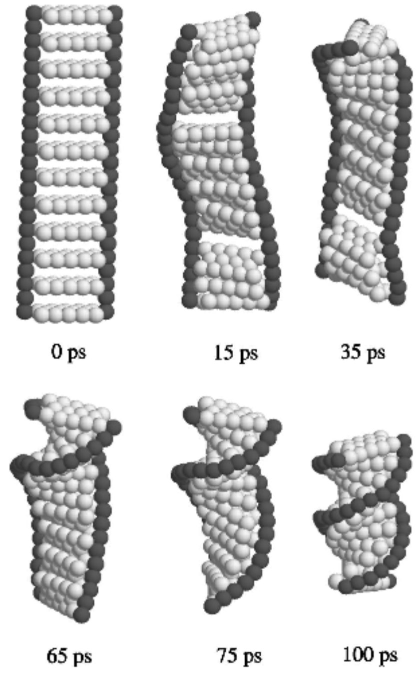

Figure 4.17: Representative snapshots of helix formation, starting from a straight ladder conformation. Note the appearance of the major(bottom half) and minor (top half) grooves in the final state. Figure from Tepper and Voth [22].

.

4.3.5 Knotts and Pablo Model

Introduction

This is a solvent-free three-site model with electrostatics parameterised to predict mechanical and structural properties of B-DNA in the mesoscopic region, in the spirit of G$\bar{o}$ type Models [19].

Model

This is a three-site model with four different base sites for the four types of nitrogen bases. The model is mapped from B-DNA atomistic coordinates, capturing correctly the major and minor grooves. Two strands make up the double helix with hydrogen bonds between them. Phosphate groups contain a negative charge while ionic effects are captured using the Debye-Huckel approximation. The derivation of parameters from the B-DNA precludes transitions to other canonical forms of DNA.

Simulation and Results

Traditional canonical molecular dynamics and replica exchange molecular dynamics simulations were performed. The model correctly captures the melting behaviour with varying salt concentration. The calculated persistence length is off by a factor of two, although the model captures qualitatively the persistence length behaviour with varying salt concentration; the persistence length of single strands is in agreement with experimental measurements. Bubble formation and rehybridization dynamics were studied using this model. Sambriski et al. [173] extended the Knots force field to look at renaturation of DNA specifically by including a weak intermediate range potential acting between inter-sugar pairs to aid the renaturation. Looking at short strands of DNA with varying content they were able to identify that the renaturation occurs as a second order reaction with the rate determining step being the nucleation followed by a rapid zipping action. Further work by Sambriski [174] using expanded ensembles shows that free energy profiles of varying content show similar behaviour, while the reduction of ionic concentration reduces the free energy barrier to renaturation as a result of rigidization of the single strands. The extended ensemble suggests a critical nucleus of two base pairs for renaturation. Transition path sampling helps identify mechanisms of renaturation in both random and repetitive sequences and identifies a critical nucleus of four base pairs.

4.3.6 Tan and Harvey Model

Introduction

This is a simple three-site solvent and electrostatic-free model developed to capture the dynamic and structural behaviour of relatively large DNA strands, and to study how distant sites in a long DNA strand are brought close together.

Model

The model consists of three sites representing the planer nucleotide pair, the three sites representing the central helix, a marker in the major groove, and a marker indicating one of the backbone sites [175]. The model was built to match the superhelical density of B-DNA in nature. Bond and angle stretching force constants mimic elastic moduli of the DNA double helix. The model was able to simulate four helical parameters, roll, tilt, twist and rise. A 50 Å bonded exclusion radius prevents strand cross-overs, corresponding to an ionic strength of 0.115 M.

Simulation and Results

DNA strands of several hundred base pairs were simulated using molecular dynamics, looking at both structural and dynamic effects and how these are affected by sequence changes. The model was used to simulate both homogeneous and heterogeneous sequences, to identify motions such as slither and branching, and their sequence effects. The model reproduces the correct persistence length of B-DNA [176][177][178].

4.3.7 Lattice DNA Models

Amaral et al. [179] look at single stranded DNA using a two site per nucleotide electrostatics-free lattice model. The sugar is represented by a sphere and the bases by pins attached to it; the phosphodiester bonds are represented using rigid rods connecting the sugar groups, the base may represent any of the four nitrogen bases. The model was used to look at how base-base heterogeneity affects mechanical and kinetic properties at mesoscales.

Zhang and Collins [180] propose a two-dimensional lattice model where the nitrogen bases and the deoxyribose sugars are treated as rigid units connected to a flexible rod representing the phosphate backbone linkage; two such strands form the double stranded DNA model. The nitrogen bases are represented by two sites for A and T, and three sites for C and G. Canonical molecular dynamics simulations were used to study the dynamics of the double strand, in particular the melting behaviour. The melting behaviour was in agreement with experimental trends with A-T strands showing greater tendency to melt than C-G or mixed C-G and A-T strands. The model shows sharp melting curves with melting temperatures varying with composition, though the melting temperatures were around 100 K higher than experimental values, with melting taking place over long time scales.

4.4 Comparison of Lipid and DNA Model Granularity

A tabular view of the different models discussed above is given in Tables 4.1-4.3 indicating the different levels of granularity of published models. This allows for the comparison of lipid and DNA models in order to select the most appropriate models to mix and investigate lipid-DNA mixtures.

	Drouffe [115]	Brannigan [116]-[118]	Takasu [119]-[122]	Deserno [125][126]	Farago [114]	Philip [133]	Sintes [134]-[136]
No. of sites/lipid	1(for 2 lipids)	1	3	3	3	5	5
Electrostatics	×	×	×	×	×	×	×
Self-assembly	√	√	√	√	×	√	×

Table 4.1: Comparison of Implicit Lipid Models

	Ayton [137]	Sun [50]	Schmid [140]	Lenz [143]	Smit [144] [145]	Goetz [132]	Stevens [98]	Voth [112]	Shelley [102]	Marrink [20]	Orsi [51]
No. of sites/lipid	2	2	4	7	7	11	11	13	13	12	10
Electrostatics	√	√	×	×	×	×	×	×	√	√	√
Self-assembly	×	×	×	×	√ (mono)	√	√	×	√	√	√

Table 4.2: Comparison of Explicit Lipid Models

	Huertas [167]	Drukker [23]	Mergell [21]	Voth [22]	Knotts [114]	Harvey [176]-[178]	Sansom [9]
No. of sites/nucleotide	1	2	1 (nucleotide pair)	18 (nucleotide pair)	3	3 (nucleotide pair)	5/6
Explicit solvent	×	×	×	√	×	×	√
Electrostatics	×	×	×	×	√	×	√

Table 4.3: Comparison of DNA Models

4.5 Simulations of lipid-DNA complexes

The formation of lipid-DNA complex is driven by the electrostatic attraction between negatively charged DNA and cationic lipid head groups, and through the entropic gain associated with the concurrent release of the tightly bound conterions from cationic lipids (CL) and DNA [181]. The structures formed by the lipid-DNA mixtures maybe also determined by the neutral lipids involved in the complex. Most neutral lipids when complexed with CLs lead to a L_{α}^{C} phase. Addition of the neutral lipid DOPE (or other PE based lipids) drives the spontaneous curvature negatively, thereby inducing a transition to the H_{II}^{C} phase [131].

Viral vectors continue to be used for transfection as they poses a larger transfection efficiency (TE) than the cationic lipid-DNA complexes. The relationship between the physical state and the TE of the complexes have been studied through a combination of experimental and theoritical methods for L_{α}^{C} systems, that show a lower TE compared to H_{II}^{C} . Improving the TE of L_{α}^{C} complexes is important as most neutral lipids tend to form L_{α}^{C} phases [131]

Very few simulations of lipid-DNA complexes have been attempted, these include a single all-atom simulation and two coarse grain simulations; these are briefly discussed below.

4.5.1 Khalid and Sansom Model

This is the most recent and by far the most successful coarse grain lipid-DNA model [9], based on the same principles of the popular Marrink et al. [20] lipid model. As in the Marrink lipid model, four heavy atoms were represented by a single site. The DNA molecule was represented by a single particle for the phosphate, two for the deoxy-ribose sugar, and two and three for pyrimidine and purine bases, respectively. The particle types were chosen based on their chemical nature as in the Marrink parametrisation [20]. Non bonded interactions were modelled through the LJ potential while charged groups in addition to the LJ potential also include a shifted Coulombic potential. All particles within 0.7 nm of each other are held together by harmonic restraints based on the elastic network model; this helps maintain the correct geometry in the model. The parameters for lipid molecules used in this work are also based on the Marrink model. The developed DNA model was validated by computing the persistence length of a 300 base pair duplex; the calculated persistence length was in excellent agreement with the experimentally measured values for B-DNA.

The MD simulation consisted of 116 DPPC lipids, 116 DMTAP (dimyristoyltrimethylammonium propane), a DNA dodecamer, 3073 coarse grain solvent molecules and 136 charge neutralising Cl^{-1} ions. Self assembly runs show that the DNA preferentially locates it self at the lipid bilayer interface in agreement with the work by Klein et al. [182]. The formed lipid-DNA complexes are in agreement with the all-atom simulation by Klein et al. [182]. The simulation is also able to show the reorientation of the DPPC head groups to interact with the DNA backbone. Potential of mean force (PMF) calculations indicate the high energy barriers that the DNA experiences as it crosses the hydrophobic core of the bilayer. The high energy barriers to the crossing of the DNA through the bilayer indicates that the possible mechanism of DNA release may involve an alternate mechanism; one that does not involve transmembrane crossing.

Poulain et al. [183] have used the DNA model developed by Khalid et al. to capture protein-DNA docking interactions.

4.5.2 Farago Model

The coarse grain model by Farago et al. [130] is one of the more recent attempts to capture the interactions between the DNA and lipids. The model is based on "water free" lipid models proposed by the same author [114][129] that showed spontaneous bilayer formation. The lipid/DNA model proposed studies isoelectric complexes where the total charges of the DNA and the cationic lipids neutralise each other with no counterions being added. A fraction of the lipid head groups carry a charge of $+e$. The DNA is modelled as a rigid rod with uniform axial charge density and radius. Simulations were performed at room temperature and bulk water uniform dielectric constant of 78. The model examined the ability of the system to self assemble starting from a random configuration consisting of 240 lipids of which 90 were cationic and 3 DNA strands each carrying a total charge of $-30e$, exactly balancing the charge of the cationic lipids. MC simulations showed the lamellar complex formation is rapid with most cationic lipids lying facing the DNA as expected, thought the run was too short for full association. On the other hand, preassembled complexes seem to be stable with respect to dissociation and lipid loss even in long MC runs. The model was used to calculate interaxial spacing at different densities and was consistent with experimental work. By increasing the fraction of cationic lipids the model was able to demonstrate the formation of pores which may explain the experimentally observed higher percentage of complex disintegration at higher cationic complex ratios.

4.5.3 Klein Model

A MD atomistic study of lipid-DNA complexes consisting of DMPC, DMTAB (dimyristoyltrimethylamonium propane) and a 10 base pair DNA duplex was done by Klein et al. [182] with the solvent being the TIP3P water model [184]; the experimental work related to such a system was done by Rädler and co-workers [185][186]. The system consists of 24 DMPC and 20 DMTAP molecules with 1003 water molecules together with the DNA duplex. The DNA duplex was inserted between the preassembled lipid layers in the L_{α}^{C} phase. The system was simulated at isoelectric point with no counterions being present under NVT conditions with the electrostatics handled using the particle mesh Ewald method. A 5.5 ns simulation was carried out using the CHARMM (all hydrogen) force field and the complex remained stable throughout the simulation. The simulation was also able to show that there are strong interactions between the DMMC and DMTAP lipids allowing the DMPC lipids to come close to the DNA backbone and compete with the DMTAP molecules in neutralising the negative charge of the backbone. The results indicate the existence of DMTAP with equal probability around the DNA phosphate group. Water was found to be playing a bridging role between the lipid head groups and DNA phosphate group. The distribution of PC and DMTAP is determined by the interaction between the cationic DMTAP head group and anionic phosphate of zwitterionic DMPC head. The interaction induces a change in orientation of the P^{-}-N^{+} headgroup dipoles of DMPC with the N^{+} part pointing away from the bilayer plane and found to be $\sim 50\,^{\circ}(\pm 8\,^{\circ})$ allowing for favourable interaction with the anionic phosphate of the DNA.

4.6 Conclusions

The majority of the lipid models detailed above are bead and spring models. While they have been successfully utilized to study various properties of the lipids, they contain far too much detail to be used in the simulation of lipid-DNA-protein mixtures with the exception of the model by Khalid et al. [9]. The model most appropriate will be a simplified representations of lipids such as that reported by Ayton et al. [137]. This model is simple yet captures the essential details of the lipids to be used in the simulation of mixtures.

The DNA models discussed above are coarse grain models that have been developed in the recent past; they are specifically developed to simulate certain properties of DNA. The biggest shortcoming of a majority of the developed models is that they lack details such as explicit solvent and electrostatics; they have been shown to be important in determining the behaviour of DNA in the cell [13][162][163][18]. The model required to simulate lipid-DNA-protein mixtures would ideally require both the explicit solvent and electrostatics to be considered, making the available models unusable in their present form. The model to be developed could be based on the model by Mergell et al. [21] since it is also based on the Gay Berne potential [25], like the lipid model being developed, with suitable modifications to consider electrostatics and explicit solvent, both features which the Mergell [21] model lacks. The model developed in this work hopes to overcome the issues raised such as the lack of electrostatic considerations by including explicit charges in the backbone sites as well as charge neutralising counter ions, plus having an explicit solvent with the correct electrostatic, structural and dynamic behaviour. Another drawback in the models discussed above is the length of the strands considered; they are typically too short, and the model to be developed would need to consider longer strands consisting of hundreds of base pairs.

Chapter 5

Construction of the Lipid , DNA and Solvent Models

5.1 Introduction

The models that are used in simulations of bio-molecular systems are highly simplified versions of the real environment, but systematic study based on tractable quantitative tools can help discern patterns and add insights that are difficult to observe experimentally. The model developed needs to be appropriate to the problem being studied [5]. The problem of lipid-DNA mixtures requires that the models developed are simple enough to allow for large-scale molecular assemblies to be modelled, but complex enough to capture the complex interactions taking place within cell nuclei.

With the intention of simulating such interations a bespoke simulation package was developed called *LANKA*, the details of which is given below.

5.2 *LANKA*: Lipid And Nucleic Acid Komputer Algorithm

LANKA is a molecular dynamics package capable of simulating rigid body dynamics. The system is able to handle the simulation of both lipids and DNA together with associated solvent and ions. The models in *LANKA* were developed and parameterised from scratch with the intention of studying lipid phase behaviour, DNA dynamics, lipid mixtures and lipid-DNA mixtures.

LANKA is able to capture interactions between spherical particles using the Lennard

Jone potential and interactions between prolate and oblate particles using the Gay Berne potential and its extensions. The model is also able to handle long range interactions such as dipole and charge interactions. In addition to the above, where necessary the models also include bonded, angular and torsion potentials. Like most available MD packages it is able to simulate a system in the microcanonical, canonical and isobaric-isothermal ensembles starting from an initial set up. The temperature and pressure in the simulation box is maintained using the Berendsen thermostat and barostat. A MD simulation in $LANKA$ is started from an initial set of coordinates for the lipids, DNA, solvent and ions (and their velocities) and is propagated by calculating the forces and torques acting on them due to pair wise interactions between them. The details of the interaction potentials in $LANKA$ are given in Chapter 2.

The implementation of the DLM integrator [47] to handle rotational motion and the simplicity of the model has allowed the use of large time step of 40 fs, while still being able to conserve energy in the microcanonical ensemble. The simplicity of the model as well as the stability of the integrator coupled with the large time step has allowed the study of large systems over long time length scales using $LANKA$.

While systematic benchmarking of the model LANKA has not been undertaken and is a requirement for the future development usage of the model a crude estimate of the speed of the system can be obtained by measuring the average length of simulation undertaken on standard CPU machines. On a standard quad-core machine using a single processor a lipid-solvent system consisting of 450 lipids 27000 waters (4500 sites) will complete approximately 40 ns in a single day. This speed is with no code optimization or parallelization, both of which are essential developments for the future.

5.3 Solvent Model

An explicit solvent model was developed with an embedded dipole to capture the dipolar nature of water, since electrostatic forces dominate in solvation interactions of biological systems such as lipids and DNA [5]. The initial water model developed by the authors represented a one-to-one mapping where each water molecule was modelled by a single site with an embedded dipole to capture the polar nature of water . However given the highly coarse grain nature of the lipids and DNA and the time ”wasted” in computing solvent-solvent interactions, the solvent model was modified to represent six water molecules. The interaction between solvents was captured by using the Lennard Jones potential [39] and dipolar potential [24][187]. The explicit formulae of the interaction potentials are given in Chapter 2 Section 2.3.4. The parameters for the solvent model are given in 5.1 with the

$Parameter$	$Value$
σ/Å	5.4
ϵ/kcal mol $^{-1}$	0.75
dipole/D	3.8
mass/amu	108.0
I_x/amu Å^2	314.93
I_y/amu Å^2	314.93
I_z/amu Å^2	314.93

Table 5.1: Solvent Parameters

parameterisation process discussed below. The solvent-solvent interactions were truncated at 13.0 Å with a switching function for the dipolar interactions starting at 11.0 Å. Initial testing and parameterisation was done at 30 °C but the solvent model was also tested at 50 °C because the lipid model developed was a highly coarse grain representation of a DPPC lipid which behaves as a liquid crystal at 50 °C, and most experimental and simulation measurements pertaining to DPPC are carried out at this temperature (melting temperature of DPPC 41.3 ± 1.8 °C [188])

5.3.1 Density

The system of 5415 solvent sites was simulated under full anisotropic pressure coupling in the constant NPT ensemble for 200 ns using the Berendsen barostat and thermostat [67] with a coupling constant of 0.1 ps in both the thermostat and barostat, the isothermal compressibility β was set at 4.6×10^{-5} atm, the isothermal compressibility of liquid water at 25 °C [189] although an exact knowledge of this factor is not necessary, it can be included in the pressure coupling constant [24]. The calculated density of the water model at 30 °C was 0.95 ± 0.0 g cm^{-3}and the calculated number density was 0.03175 molecules Å^{-3} both in excellent agreement with experimentally measured values for real water (experimental density of water at 30 °C is 0.99 g cm^{-3} and 0.03346 molecules Å^{-3}) [190] . The variation of the model density with temperature was also investigated to ensure that the model was appropriate for use at various temperatures; the density of the model at 50 °C was 0.92 g cm^{-3} and number density of 0.0308 molecules/Å^3 (Experimental density ≈ 0.98 g cm^{-3} at 50 °C [191]). The solvent model was tested in the temperature range 10-90 °C to ensure that the model behaviour was acceptable as the lipid model mechanical properties were tested in

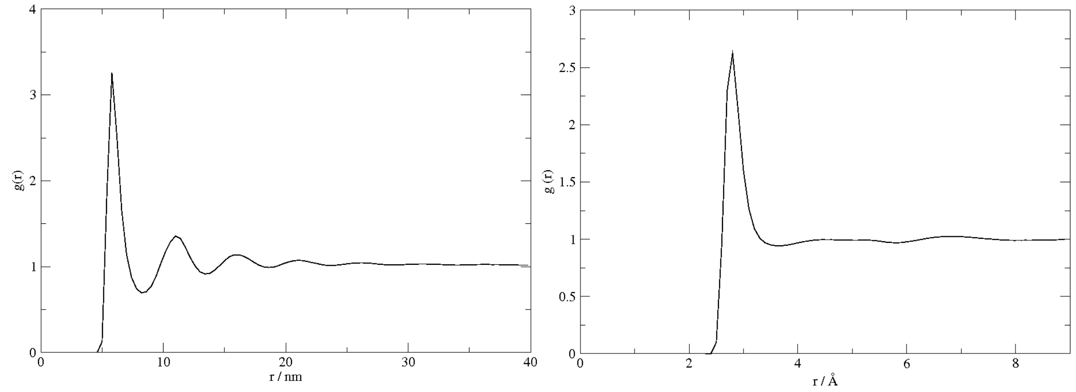

Figure 5.1: Radial Distribution Function of the solvent model with 5415 solvent molecules (left) and the TIP3P water model [184] (right)

this temperature range.

5.3.2 Radial Distribution Function

The positions of the solvent spheres may be characterised by a set of distribution functions, the simplest of which is the pair distribution function $g(r)$ [24]. This function gives the probability of finding a pair of solvent molecules a distance r apart, relative to the probability expected for a completely random distribution at the same density; it is useful for describing the structure of the system especially in liquids [39]. The radial distribution function (RDF) was calculated for the solvent model and is in qualitative agreement with other computational models [184][189][157] and experimental water [192][193][194][195] in spite of its simplicity. The plot of the RDF for the model is given in Figure 5.1. It is able to show the primary peak as well as the secondary peak which is similar to experimental data but given the larger size of the solvent sphere (representing six waters) the positions of the peaks are different [192][193] [194][196]. What is remarkable in the plot is the appearance of the second minimum, which other solvent models such as the often used TIP3P water model fail to show [184]. The RDF plot clearly shows that the developed solvent model behaves as a liquid at room temperature (the RDF plot at 50 °C is similar) unlike the widely used non-dipolar solvent models [20] [197].

5.3.3 Diffusion Coefficient

The diffusion coefficient was computed using the Einstein equation [39][40][46] as given by Equation 5.1 where D is the diffusion coefficient, N_m the number of particles in the simulation system, $\boldsymbol{r}_j(t)$ the coordinates of particle j at time t and $\boldsymbol{r}_j(0)$ the coordinates of particle

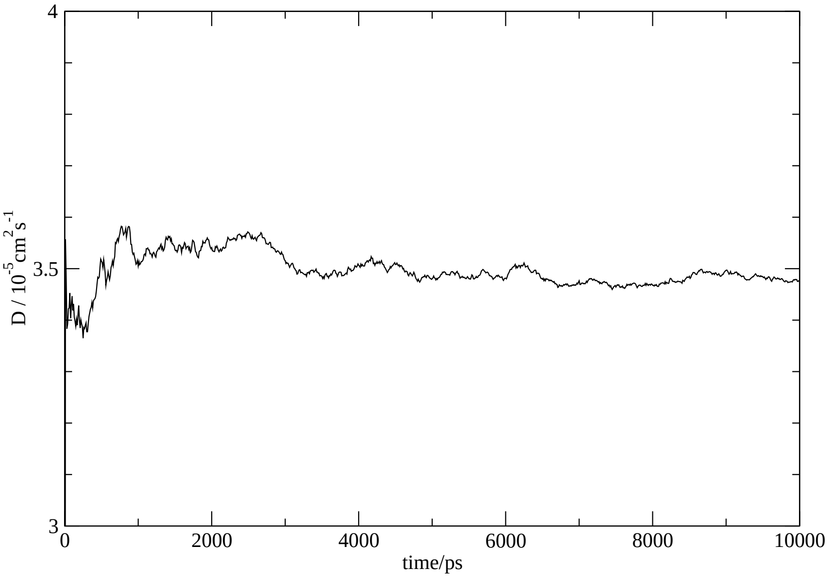

Figure 5.2: Diffusion Coefficient of the solvent model with 5415 solvent molecules

j at time beginning of the observation. The diffusion coefficient is in good agreement with experimental data although this was not part of the parameterisation process. The diffusion coefficient obtained is D = 3.4×10^{-5} cm^2 s^{-1} which is in good agreement with the experimental value of D = 2.99×10^{-5} cm^2 s^{-1}. A plot of the diffusion coefficient is given in Figure 5.2 where the diffusion coefficient is plotted as instantaneous values and is well converged. With the increase of temperature to 50 °C the diffusion coefficient increased to a value of 4.2×10^{-5} cm^2 s^{-1}.

$$D = \lim_{t \to \infty} \frac{1}{6N_m t} \sum_{j=1}^{N_m} [\boldsymbol{r}_j(t) - \boldsymbol{r}_j(0)]^2 \tag{5.1}$$

5.3.4 Dipole-Dipole Correlation

The dipole-dipole correlation functions were computed for the solvent system using Equation 5.2 for the solvent system to determine the correlation between the solvent dipoles, and is given in Figure 5.3 where $\boldsymbol{\mu}_i$ and $\boldsymbol{\mu}_j$ are the dipole moments of molecules i and j respectively separated by a distance r and $< \cos\theta(r) >$ is the average cosine angle between the dipole vectors of i and j. The distance r is calculated using the all pairs method and the data accumulated in 0.45 Å bins in a similar manner to the RDF calculation above. The figure shows dipole correlation with and without a reaction field. Both plots show a sharp first peak indicating strong polarisation. The outer anti polarisation will cancel the inner polarisation if there was no reaction field; the reaction field helps maintain a large dipole moment and

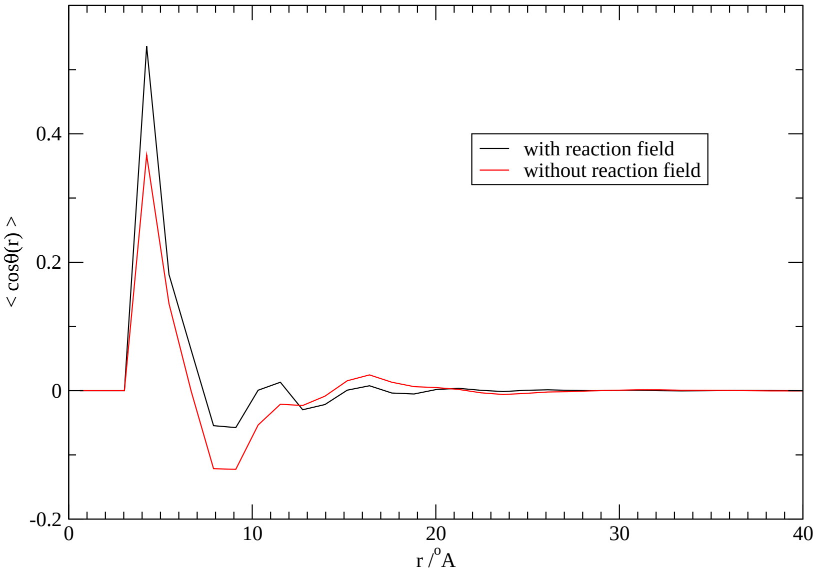

Figure 5.3: Dipole-Dipole Correlation of the solvent model with 5415 solvent molecules with and without the reaction field

hence a large dielectric constant [54].

$$< \cos\theta(r) > = \frac{\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j}{| \boldsymbol{\mu}_i || \boldsymbol{\mu}_j |} \tag{5.2}$$

5.4 Lipid Model

The lipid model is a highly coarse grain model of a saturated phosphtidylcholine lipid with around 14-18 carbons in the fatty acid chains in the tail. At this level of coarse graining it is not possible to assign the model to an exact type of lipid but the parameterisation and validation were based on DPPC lipids which contain 16 carbons in the fatty acid chains. DPPC was chosen for parameterisation as there exists a large body of both experimental and simulation work on this lipid as well as the nuclear lipids are usually fully saturated as in the DPPC lipid. The lipid is a two site model; this model replaces over a hundred atoms in a PC lipid by two sites. The two sites represent the very basic features expected of a lipid model, the hydrophilic head and a hydrophobic tail. The hydrophobic tail is represented by a rigid and long ellipsoid and at one end of the ellipsoid is a sphere with an embedded dipole representing the hydrophilic head. The dipole orientation is fixed perpendicular to the long axis of the ellipsoid, to ensure that the head dipole lies along the bilayer plane, experimentally the dipole vector on average lies 18 ° upwards from the bilayer plane [198]. The dipole magnitude was fixed at 18 D with experimentally measured values for DPPC

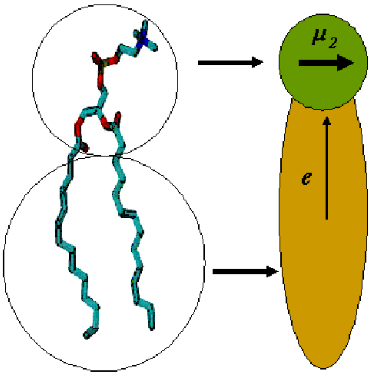

Figure 5.4: Schematic of the lipid model showing the head dipole vector and tail orientation vectors

being 18.7 D [50][70][199]. While allowing the dipole vector to be free rotating would be more realistic, it would mean incorporating an additional degree of freedom to the model and the fast motion of the dipole vector would require a smaller time step to be used. Not only was this simple model of lipid used to capture the bilayer phase but also some of the other phases such as micelles, inverse micelles and hexagonal phases. This was done primarily by changing head and tail diameters of the model. However, to better represent the increased head solvent interaction seen in lipids forming micelles and hexagonal phases the head solvent interaction was also increased in some instances. Changing the relative sizes of the heads and tails changes the packing parameters defined by Equation 5.3 [33][36][127][200] where v is the volume of the lipid, l_o the length and a_o the preferred head area per lipid and thereby changes the intrinsic curvature (see Chapter 3 Section 3.2). This feature can be used to capture many of the phases of the lipids as demonstrated in this work. The interaction between the lipids was captured by using the Gay Berne (GB) potential [25], Lennard Jones potential [39] and the dipolar potential [24][187]. The explicit formulae of the interaction potentials are given in Chapter 2 Section 2.3. The parameters used for the bilayer forming lipid model are listed in Table 5.2. The list refers to Parameter Set I; Parameter Set II differs only in the value of κ'_{Tail} (See Parameterisation Section below). A schematic of the lipid model used in *LANKA* is given in Figure 5.4 showing the coarse graining of over hundred atoms into a two site model with a fixed dipole vector.

$$P = v/a_o l_o \tag{5.3}$$

$Parameter$	$Value$
σ_{Tail}/Å	7.0
ϵ_{Tail}/kcal mol^{-1}	0.5
κ_{Tail}	2.74
$\kappa^{'}_{Tail}$	0.3
μ_{Tail}	2.0
ν_{Tail}	1.0
$\kappa^{'}_{GBLJ}$	0.2
σ_{Head}/Å	7.4
ϵ_{Head}/kcal mol^{-1}	0.9
dipole$_{Head}$/D	18.0
Head Distance/Å	9.6
mass/amu	800.0
I$_x$/amu Å^2	24884.12
I$_y$/amu Å^2	24884.12
I$_z$/amu Å^2	16589.42

Table 5.2: Bilayer Lipid Head and Tail Parameters

5.5 DNA Model

The DNA model represents the DNA at the base pair level. Each base pair is represented by a three site model. The DNA base pair is represented by oblate disks and on opposite edges of the disk by two spheres with embedded negative charges representing the phosphate groups; the disks primarily represent the planer base pairs. A schematic of the DNA model proposed is given in Figure 5.5. The representation of the two base pairs in the strand by a single disk precludes the study of hybridisation dynamics in this model. These base pair disks are then connected through the harmonic bonds to form DNA strands. To capture the correct angle geometry of the bonded group, an angle potential, and to prevent the formation of kinks due to the achiral nature of the model, torsion potentials, were introduced (See Chapter 2 Equations 2.20-2.22). The disks interact via the Gay Berne potential while the charged backbone sites through the shifted Coulombic and Lennard Jones potentials in addition to the above mentioned bond, angle and torsion potentials. The explicit formulae of the interaction potentials are given in Chapter 2. The parameters used for the DNA model are listed in Table 5.3.

The Gay Berne potentials representing the disks were chosen to model the DNA for several reasons, namely the ease of mixing with the lipid model based on the GB model and, at the level of granularity required, base stacking interactions dominate in DNA [21][37]. The GB potential has been used previously in modelling DNA strands at the base pair level but the model lacked an explicit solvent and electrostatics [21], and the model developed here improves on this model in this regard (See Chapter 4 Section 4.3.3). The propensity of the GB disks to favourably stack on top of each other could also be exploited by using a GB disks [28][30].

5.6 Ion Model

Ions are represented by spheres with embedded point charges. A system was developed to handle both negative and positive ions in order to simulate a ”realistic” ionic environment though most simulations were done with positive ions introduced as counter ions for the negatively charged DNA. The ions interact via the Lennard Jones and Coulombic potentials. The Coulombic and Lennard Jones potentials were modified using a shifting function to avoid artefacts and energy pumping due to abrupt truncation of interactions at the cut-off. The explicit formulae of the interaction potentials are given in Chapter 2 Section 2.8.2. The parameters used for the ion model are listed in Table 5.4.

$Parameter$	$Value$
σ_{Disk}/Å	18.0
ϵ_{Disk}/kcal mol $^{-1}$	1.0
κ_{Disk}	0.1833
$\kappa^{'}_{Disk}$	0.1
μ_{Disk}	2.0
ν_{Disk}	1.0
$\kappa^{'}_{GBLJ}$	0.2
$\sigma_{BackBone}$/Å	3.0
$\epsilon_{BackBone}$/kcal mol^{-1}	1.0
$charge_{BackBone}$/e	-1.0
$BackBoneDistance$/Å	9.0
$mass$/amu	800.0
I$_x$/amu Å^2	16589.42
I$_y$/amu Å^2	16589.42
I$_z$/amu Å^2	24884.12

Table 5.3: DNA Base Pair and Back Bone Parameters

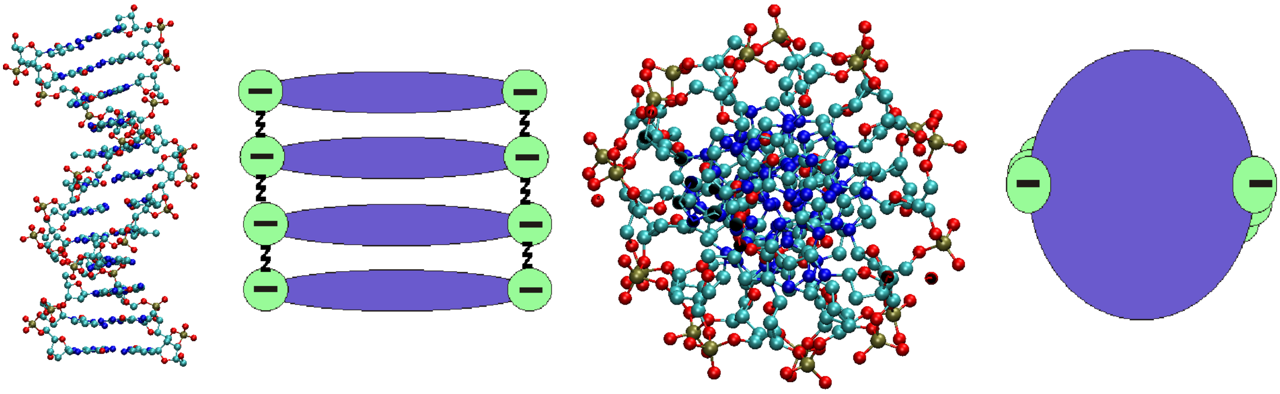

Figure 5.5: Schematic of the DNA model-side view and top view along with the CPK representation of the side and top views of the Dickerson B-DNA (PDB-Code:1BNA) [86].

$Parameter$	$Value$
σ/Å	5.4
ϵ/kcal mol $^{-1}$	0.27
$charge$/e	0.7
$mass$/amu	800.0
I_x/amu Å^2	2332.80
I_y/amu Å^2	2332.80
I_z/amu Å^2	2332.80

Table 5.4: Ion Parameters

5.7 Parametrisation

5.7.1 Truncation of Interactions

All interactions were truncated at different lengths to speed up simulations. Dipolar and Coulombic interactions are long ranged and abrupt truncation of these interactions would lead to non-Newtonian trajectories. Dipolar interactions were scaled using a switching function while Coulombic interactions were scaled by a shifting function leading to smooth trajectories with no pumping of energy due at the cut-off. All switching interactions were implemented 2.0 Å before the cutoff. A list of different cut-offs used in the system is given in Table 5.5. Further details of the switching and shifting functions used can be found in Chapter 2 Sections 2.8.2 and 2.8.3. The cutoff distances were determined based on the dimensions of the lipids and DNA models as well as electrostatic considerations. For instance the cutoff distance for a pair of lipid head groups was set at 40.0 Å to allow two lipids arranged end-to-end in a bilayer to be able to *see* each other as well as considering the large magnitude of the head dipole. All other electrostatic interactions were truncated at the same cutoff as the solvent-solvent interactions; to avoid possible artifacts due to differing cutoffs between the solvent and electrostatic interactions [201].

5.7.2 Parameterising the Solvent Model

The solvent model was parameterised based on capturing the correct dielectric behaviour and density of water. An initial FCC lattice was melted and the system run under full anisotropic pressure coupling to obtain the density. Since the mass of each solvent site was set to 108 amu the density was calculated in g cm^{-3} as well as the number density. The

$Parameter$	$Value$
$Lipid_{Tail} - Lipid_{Tail}/\text{Å}$	40.0
$Lipid_{Head} - Lipid_{Head}/\text{Å}$	40.0
$DNA_{Disk} - DNA_{Disk}/\text{Å}$	40.0
$Lipid_{Head} - Lipid_{Tail}/\text{Å}$	30.0
$DNA_{Disk} - DNA_{BackBone}/\text{Å}$	30.0
$Lipid_{Tail} - Solvent/\text{Å}$	30.0
$Lipid_{Head} - Solvent/\text{Å}$	30.0
$DNA_{Disk} - Solvent/\text{Å}$	30.0
$DNA_{Disk} - Ion/\text{Å}$	30.0
$DNA_{BackBone} - Solvent/\text{Å}$	13.0
$DNA_{BackBone} - Ion/\text{Å}$	13.0
$Solvent - Solvent/\text{Å}$	13.0

Table 5.5: Truncation Lengths of Interactions

parameters were optimized in an iterative fashion until the systems had a density close to experimental water at 30 °C. The initial value for the dipole moment was obtained by looking at the net dipole magnitude of the closest six neighbours in a well equilibrated TIP3P [184] water box. Simulation using a reaction field method [202][203] in the canonical ensemble used to fine tune the dipole moment and obtain the correct dielectric behaviour. Different starting configurations and different system sizes in the canonical ensemble were simulated, although previous studies using the reaction field method have shown that the calculated dielectric constant is largely insensitive to the system size [55][204] and it was also found to be the case in this study as well. The system sizes simulated consisted of 3000 sites (18000 waters), 4500 (27000 waters) and 7000 sites (42000 waters). The results reported are for a system consisting of 4500 sites each representing six water molecules and hence the simulation box consists of 27000 water molecules in a 100 Å cubic box with a cutoff of 13 Å and switching from 11 Å. The system sizes used are large in comparison to other similar solvent systems; this was done to reduce the effect of using a switching in simulations [55]. The box was initially run NPT for 50 ns then switched to NVT (to enable the calculation of the dielectric constant using a reaction field). The results were collected over 100 ns simulations; long simulations were done to ensure that the values of the calculated G_k were converged (Refer Section 2.8.1 for further details on the reaction field method). The calculated dielectric constant of the system

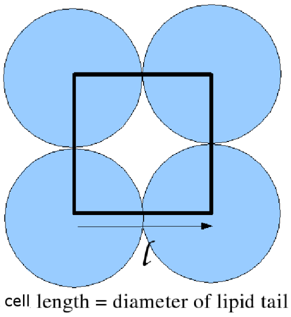

Figure 5.6: Simple cubic lattice used for parameterising bilayer lipid dimensions

at 30 °C is 76 ± 6 with a dipole moment of 3.8 D which is good agreement with experimental value of 78.0 for water [190]. The value for the dielectric constant of the medium outside the cutoff is assumed to be ∞ in this study since the value of the dielectric constant was not known.

Once suitable parameters for the Lennard Jones and dipole moment were determined the system was simulated in the microcanonical ensemble for over 50 nanoseconds to ensure energy conservation at both 30 °C and 50 °C. The dielectric constant at 50 °C was 72 ± 8.

5.7.3 Parameterising the Lipid Model

The initial goal of parameterising the lipid model was to obtain a stable and fluid bilayer structure. To this end the energy parameters for the lipid model were obtained from the original work by Ayton et al. [137]. However, the geometric parameters were based on the packing of a PC lipid in a simple cubic lattice as shown in Figure 5.6. Since PC lipids in the L_α phase have areas-per-lipid (A_L) of around 60 Å^2 the diameter of the lipid was taken to 8 Å while the length of the lipid molecules was based on bilayer thickness [138][205][206]. The dipole moment of the lipids in initial work was set to 1.0 D to make one end of the lipid model polar but was later increased to 18 D better reflect the highly polar nature of PC lipids [50][70][199]. The area-per-lipid, volume-per-lipid and bilayer thickness were used to parameterise the lipid model.

The system for self assembly was generated by placing lipids and solvents randomly on a face centred cubic lattice (FCC) with random orientation vectors for the tail ellipsoids and solvent dipoles [24][40]. However, the use of a lattice system meant that the lipid and solvent sites were required to ”grow”; where the diameter and the length of the lipids as well as the diameter of the solvent were steadily increased. During the ”growth” process the dipole moments for the solvent and lipids were removed. Once the lipids and solvent were at the correct sizes, the system was run in the canonical ensemble at elevated temperatures ranging from 500 °C to 1000 °C ensure that the system started from a completely random starting configuration. The randomness of the system was ensured using the P_2 order parameter [207][208][209][210] and visual inspection of the trajectory files; the order parameters were close to zero in all systems indicating a random mix of lipids and solvent, although not exactly zero due to the finite size effects [209]. This was followed by the switching on of the dipoles and production runs initially in the canonical and then in the isobaric-isothermal ensemble. However, the systems did not form bilayers. Instead the lipids showed some local clustering at best even with long simulation runs exceeding 500 ns.

As an alternate strategy, a preassembled bilayer was set up as indicated in Figure 5.7. The bilayer was set up in the centre of the box with solvent sites arranged on either side. The sites were placed on each layer in several passes as indicated in Figure 5.7. The system initially ran in the canonical ensemble followed by isotropic NPT ensemble and fully anisotropic NPT ensemble. The systems simulated all had excess solvent particles with a varying lipid to solvent ratio (ratio between lipids to solvent ranging from 1:27, 1:60, 1:100) to ensure a fully hydrated simulation system. However, maintaining the correct density starting from a lattice meant that the starting sizes of the solvent and lipids were far smaller than they were required to be as in the random self assembly attempts above. This meant that the particles; both lipids and solvents needed to be "grown" before production runs. This was done in small stages using extremely low time steps of around 1fs. With either solvent or lipids' translational movement "frozen" The freezing of the translational motion was also alternated with unconstrained MD, where the particle masses were artificially increased to reduce the translational motion of the particles. Usually the lipids were allowed to "grow" and relax to form a stable bilayer before the solvent system was "grown" and equilibrated. This method of building a bilayer while rather ad-hoc serves its purpose.

The initial setup of the bilayer was unstable, either breaking up during the "growth" stage or during the NPT runs while on some occasions the solvent particles permeate the bilayer leading to its break up. The individual potential energy contributions were then analysed and the interactions driving the breakup of the bilayer determined by looking at which energies became more negative during the simulation, instead of fluctuating about a mean value. These interactions often were the mixed interactions taking place between unlike Lennard Jones sites such as between head sites and solvent or between Gay Berne (tail) sites and Lennard Jones sites. Accordingly the system was scaled from standard Lorentz-Berthelot mixing rule values to the following; by 0.2 for the head-tail and by 0.1 for solvent-tail interactions while the head-solvent interaction strength was left as an adjustable parameter as given in Table 5.6. These values were arrived at by exhaustive repetitive testing where the scaling values were

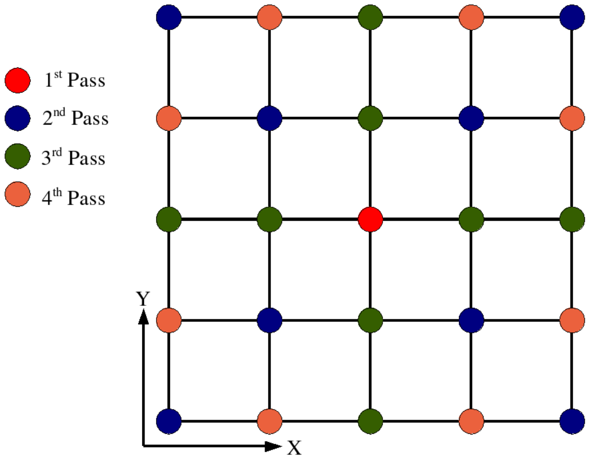

Figure 5.7: Setting up of the grid-scheme showing one plane in the lattice

$Parameter$	$Value$
$\sigma_{Tail-Head}$	$\frac{1}{2} \times (\sigma_{Tail} + \sigma_{Head})$
$\epsilon_{Tail-Head}$	$0.2 \times \sqrt{\epsilon_{Tail} \times \epsilon_{Head}}$
$\sigma_{Tail-Solvent}$	$\frac{1}{2} \times (\sigma_{Tail} + \sigma_{Solvent})$
$\epsilon_{Tail-Solvent}$	$0.1 \times \sqrt{\epsilon_{Tail} \times \epsilon_{Solvent}}$
$\sigma_{Head-Solvent}$	$\frac{1}{2} \times (\sigma_{Head} + \sigma_{Solvent})$
$\epsilon_{Head-Solvent}$	$\sqrt{\epsilon_{Head} \times \epsilon_{Solvent}}$

Table 5.6: Scaling for Mixed Interactions for Bilayer Lipids

reduced gradually, each time testing to ensure that the preassembled bilayers remained stable during long anisotropic NPT runs extending several hundred of nanoseconds.

Since the initial energy parameters obtained from the Ayton model [137] proved be unsuitable; leading to unstable bilayers; an alternative strategy was developed to obtain the Gay Berne potential. This involved looking at a coarse grain model of DMPC lipid tails where the hydrocarbon tail was represented by six Gay Berne sites, three ellipsoids per chain [51][158] and calculating the interactions for these systems as given in Figure 5.8, where a grouping of six ellipsoids arranged in the manner shown interact with another six ellipsoids in different arrangements; where the displacement between the molecules were varied and the resulting potentials fit to obtain starting values for the Gay Berne energy parameters. The energy minimum of 2.0 kcal mol^{-1} for the X configuration was taken as the starting value for the ϵ_0 for the ellipsoid tail in this two site model. The values for the head sites were obtained by looking at a head sites in a united atom DPPC model [211] obtaining a starting value of

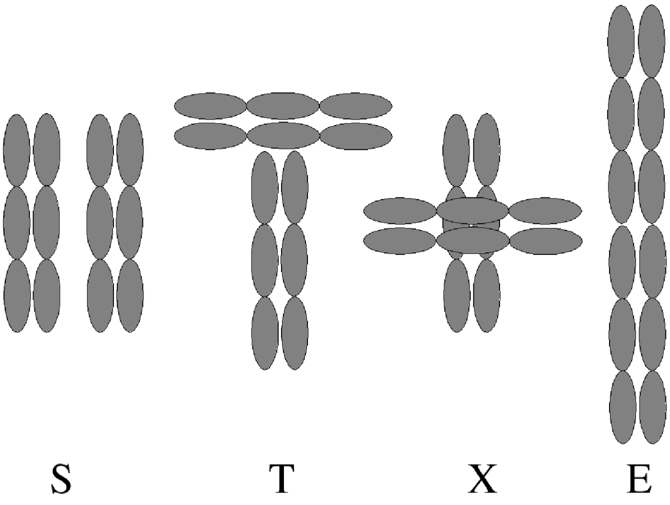

Figure 5.8: Different configurations used in parameterising the system

3.0 kcal mol^{-1}.

The exhaustive parameterisation resulted in two sets of parameters that were deemed suitable for self assembly runs culminating in bilayers. The two sets of parameters were different in only one parameter, the side-by-side to end to end interaction strength ratio in the Gay Berne potential that determines the strength of the side by side to end to end interactions. The two sets of parameters will henceforth be known as Set I and Set II. The parameters for the lipid model Set I are given in Table 5.2, Set II has a value of 0.9 for κ'_{Tail} with all other parameters identical to Set I as given in Table 5.2, all simulations were run at 50 °C with a 40 fs time-step.

5.7.4 Parameterising the DNA model

The DNA model was set up in a similar manner to the lipids. The initial values for the diameter of the disks were obtained by extracting the diameter of a B-DNA form x-ray crystal structures [10][37]. The Lennard Jones parameters for the backbone sites were chosen considering the packing constraints such as the spacing between adjacent base pairs and the length of the phosphodiester bonds from of x-ray crystal structures [37]. The DNA model backbone site represents primarily the phosphate backbone and also part of the deoxy ribose sugar, making it difficult to map the parameters to any of the existing coarse grain models [19][21][22][23]. The force constant for the phosphate-phosphate backbone bond was taken from other coarse grain DNA models [23] while the backbone-backbone bond distance was set initially at 6 Å based on experimental measurements [37] and other coarse grain DNA models [21][22] with length modifications introduced to capture the correct helix repeat. The backbone-backbone-backbone angle was determined by analysing the PDB structure of a B-DNA dodecamer (PDBCode:1BNA) [86]. and the angle was $\approxeq 154\,^{\circ}$. The angle potential was implemented as a cosine angle potential (See Equation 2.21 in Chapter 2) in order to maintain the correct helix geometry. The model lacks the chirality present in the sugar groups in real DNA and hence when the model coiled up the coiling may be initiated in opposite directions on either ends of the strand and may result in kinked structures as seen in Figure 5.9, an effect seen in other achiral coarse grain models of DNA [22]. To avoid the formation of kinks a torsion potential based on Ryckaert and Bellmans [44] was implemented. The Ryckaert and Bellmans potential coefficients favour the formation of trans alkane structures, the same coefficients were used but with their signs switched to favour cis like structures and avoid the formation of kinks [40][44] (See Equation 2.22 in Chapter 2), the coefficients used are as follows -1.000, -1.310, 1.414, 0.330, -2.828, 3.394 [24][40].

Initial attempts at parameterisation were aimed at obtaining a stable helix structure. Once the helix structure was obtained the resulting short DNA strand consisting of 22 disks representing 44 nucleotides were run under canonical as well as isobaric-isothermal ensembles. The energy parameters of the GB ellipsoid and backbone spheres in the initial helix generation were unrealistic (relatively large compared to the solvent and lipid models) and were carefully modified to better reflect the other parameters (bringing them closer to the energy parameters of the lipid model) being used in the lipid model by a trial and error method while still capturing the correct helix structure such as the correct helix repeat and base pair rise. The scaling used for mixed interactions between DNA, solvent and ions is given in Table 5.7. The mixed interaction parameters were either based on standard Lorentz-Berthelot mixing rules or were scaled by the same value as lipid-solvent interactions (See above Table 5.6). The only exception was the value for $\sigma_{Disk-Ion}$; which was scaled by $\frac{1}{1.2}$ instead of $\frac{1}{2}$, to prevent the penetration of the ions into the grooves of the DNA. These simulations were run using 40 fs time-step at 30 °C, the parameters used for this model are listed in Table 5.3.

5.7.5 Parameterising the Ion Model

Initial parameters for the ions were obtained from work by Dang and Smith [212]; the parameters from Åqvist [201] were not used as they have been shown to form anomalous ionic clusters [213]. However, given the coarse grain nature of the solvent and other solutes the size of the positive ion was increased to the size of the water particle to take into account the hydration shell around the ion, and the charge reduced to 0.7 e per Na^{+}; this is similar to the strategy adopted by Marrink et al. [20] and the use of the same diameter for the ion as the solvent is justifiable given that Na^{+} and Cl^{-} ions have a coordination number of

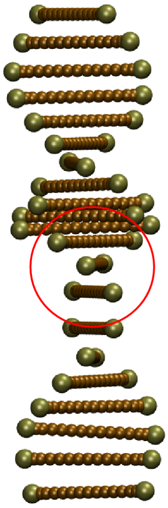

Figure 5.9: Kinked DNA Structure

$Parameter$	$Value$
$\sigma_{Disk-BackBone}$	$\frac{1}{2} \times (\sigma_{Disk} + \sigma_{BackBone})$
$\epsilon_{Disk-BackBone}$	$0.2 \times \sqrt{\epsilon_{Disk} \times \epsilon_{BackBone}}$
$\sigma_{Disk-Solvent}$	$\frac{1}{2} \times (\sigma_{Disk} + \sigma_{solvent})$
$\epsilon_{Disk-Solvent}$	$0.1 \times \sqrt{\epsilon_{Disk} \times \epsilon_{Solvent}}$
$\sigma_{BackBone-Solvent}$	$\frac{1}{2} \times (\sigma_{BackBone} + \sigma_{Solvent})$
$\epsilon_{BackBone-Solvent}$	$\sqrt{\epsilon_{BackBone} \times \epsilon_{Solvent}}$
$\sigma_{Disk-Ion}$	$\frac{1}{2} \times (\sigma_{Disk} + \sigma_{Ion})$
$\epsilon_{Disk-Ion}$	$\sqrt{\epsilon_{Disk} \times \epsilon_{Ion}}$
$\sigma_{BackBone-Ion}$	$\frac{1}{1.2} \times (\sigma_{BackBone} + \sigma_{Ion})$
$\epsilon_{BackBone-Ion}$	$\sqrt{\epsilon_{BackBone} \times \epsilon_{Ion}}$
$\sigma_{Solvent-Ion}$	$\frac{1}{2} \times (\sigma_{Solvent} + \sigma_{Ion})$
$\epsilon_{Solvent-Ion}$	$\sqrt{\epsilon_{Solvent} \times \epsilon_{Ion}}$

Table 5.7: Scaling for Mixed Interactions for DNA, Solvent and Ions

six [20][212][214]. The reduction of the charge is justified based on the screening provided by the coordinated water molecules around the ion. The Lennard Jones energy parameter for the new ion model was determined by applying standard Lorentz-Berthelot mixing rules [24] to the solvent model and Dang and Smith ion model [212].

5.8 Lipid Model Results

5.8.1 Lipid Bilayers

Lipid Bilayer Self Assembly

Snap shots of self assembly of a lipid bilayer is shown in Figure 5.10, these are for parameter Set I as given in Table 5.2. A similar observation was seen with parameter Set II but the bilayer segments appeared more fragmented compared to Set I. The system was started from a completely random mixture confirmed by calculating the order parameter ($P_2 \approx 0.1$) for the lipid and by visual inspection. The images clearly show the quick phase separation followed by slow self assembly of several bilayer fragments. It was observed that while the formation of bilayer fragments was fairly fast, they remain fragments with only very slow merging of these fragments. It is not apparent why this occurs though it could be due to the long lengths of the ellipsoid tails and their rigid nature. The measurements of lipid properties were done in preassembled bilayers containing 450 lipids and 4500 solvent (27000 solvent) particles after a long equilibration ($>$ 100 ns) under anisotropic NPT ensembles. Preassembled bilayers were used due to the computational difficulty in using the self assembled bilayer fragments though it would be expected to give similar results as the preassembled bilayers. This method of using preassembled bilayers has seen extensive use in all-atom simulations as well as several coarse grain lipid models. To ensure that the preassembled bilayers have the same properties as the self assembled bilayer fragments, the volume-per-lipid (V_L) and order parameters(P_2) were compared and found to be consistent.

An interesting observation was the effect of the removal of the net translation during the simulation that is required to prevent the so called ”flying ice cube” effect [215]. Initially the net translation was removed every single step, resulting in poor aggregation dynamics. However, decreasing the frequency of the net translational removal to every 2 ns improves the rate of formation of bilayers (and other aggregates). This is an interesting phenomenon that may affect other simulations involving rigid bodies of relatively large dimensions. Further simulations are required to validate this observation.

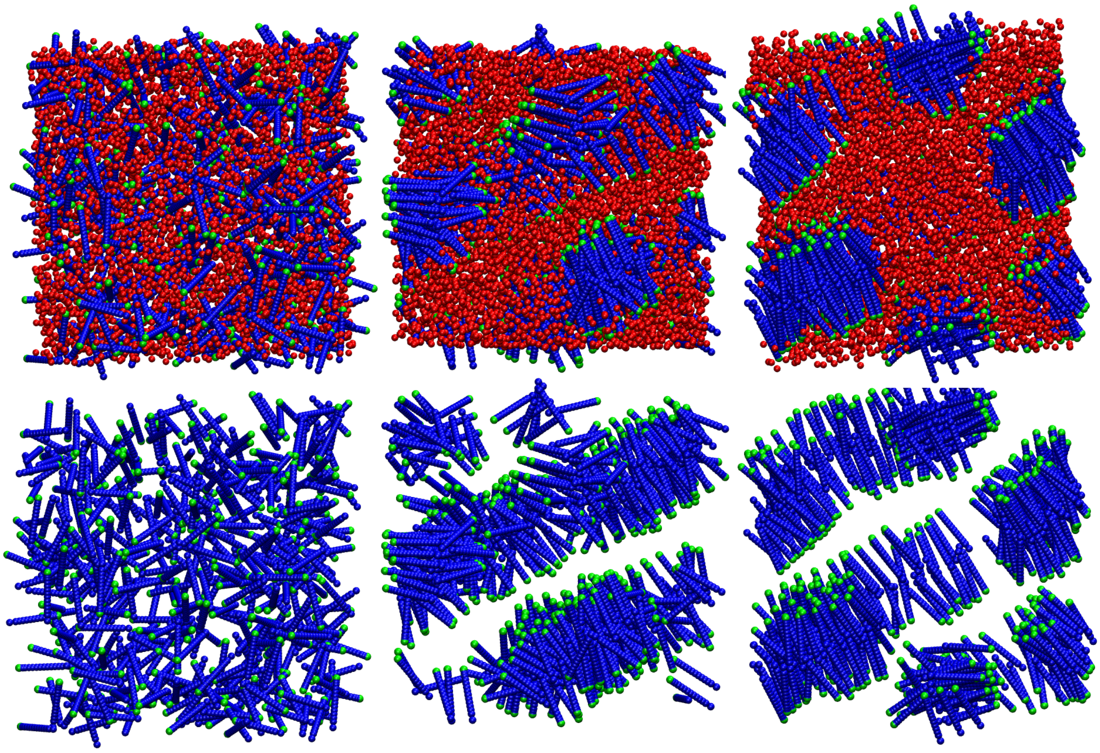

Figure 5.10: A snapshot of a 450 lipid with 27000 (4500 sites) solvent system with and without the solvent after 0 ns , 60 ns and 200 ns

Elastic Properties

The area and volume compressibility can be calculated for the bilayer patch containing 450 lipids and 4500 solvent sites (27000 water molecules) at 50 °C giving an indication of the ability of the membrane to deform due to external forces. The area compressibility can be calculated from the fluctuations in the membrane area per lipid and is given by $K_A = k_B T \langle A \rangle (N \langle (A - A_0)^2 \rangle)^{-1}$ where N denotes the number of lipids per monolayer and A_0 the equilibrium area. The experimental value for DPPC is 231 ± 20 mN m^{-1} while Marrink et al. [20] report a value of 260 ± 40 mN m^{-1} for their coarse grain model of DPPC. Orsi et al.[51] calculating K_A for a coarse grain DMPC lipids report a value of 297 ± 22 dyn cm^{-1} (297 ± 22 mN m $^{-1}$) while experimental values for a DMPC lipid are reported as 234 dyn cm^{-1}(234 mN m^{-1}). Experimentally it has been shown that K_A does not depend strongly on the length of the hydrocarbon tail as can be seen from the K_A values for DMPC and DPPC[216].

The values calculated from our model using parameter set I is 406 ± 68 mN m^{-1} and for parameter set II 1372 ± 46 mN m^{-1}. The parameter set II values are significantly higher than experimental values while parameter set I values are fairly good agreement with experimental values. The higher compressibility values may arise as a result of the rigid tail failing to capture the splaying that is inherent in real lipids and other coarse grain models, and the rigidity of the tails. However the parameter set II model is more closely related to bilayers containing cholesterol. It is well known that cholesterol increases the rigidity of lipid bilayers [217][218]. The preassembled bilayers used in calculating the area compressibility as well as lipid areas (see below) had naturally occurring undulations. This is likely to lower the area compressibility of the system (and area per lipid). However, experimental measurements are usually made in multi lamellar vesicles having naturally occurring thermal undulations, and, hence the values calculated in this system are closer to ”true” values while simulation measurements using smaller patches of bilayers claiming to reproduce experimental values for area-per-lipid and area compressibility need to be re-examined with larger systems [20][51]. The errors bars reported are standard errors using four simulations each 60 ns in length for the above calculation as well as all simulations henceforth.

The volume compressibility of the model was computed from volume fluctuations using $K_V = k_B T \langle V \rangle (N \langle (V - V_0)^2 \rangle)^{-1}$ where N denotes the number of lipids per monolayer and V_0 the equilibrium volume. The volume compressibility is several orders of magnitude smaller than the area compressibility of lipids and indicates that the lipids behave as dense incompressible liquids. The experimental value for K_V is in the range $10 - 30$ kbar while the developed model has a value of 3.1 ± 0.2 kbar for parameter set I and 6.4 ± 0.1 kbar for parameter set II at 50 °C indicating the bilayers are softer than experimental PC lipids. The variation of K_A and K_V in the temperature range 10 °C to 90 °C at 10 °C intervals is given in Figures 5.11 - 5.12 respectively and is in line with experimental observations of decreases with increasing temperature [98].

Bending Modulus

The bending modulus is a macroscopic property that determines the shape, strength and structural phases of lipids [219]. The bending modulus can be calculated using the equation $\kappa^b = K_A h^2 / b$ where K_A is the area compressibility, h the height /thickness of the bilayer and b the geometrical factor/elastic ratio. The value used for b is 48 as used by several other simulation studies [98][116][147][220]. A value of 24 is used in some experimental [221] and simulation work [51] for the calculation of the bending modulus where the height of the bilayer used is $h - h_0$ where $h_0 = 1$ nm is the nondeformable length determined from plots of $(\kappa^b / K_A)^{\frac{1}{2}}$ versus h. The use of 48 or 24 does not alter the final value of the bending modulus as the use of $h - h_0$ scales the value of κ^b by approximately 1/2.

An atomistic study of DPPC gave a value of 5×10^{-20} J for the bending modulus while a

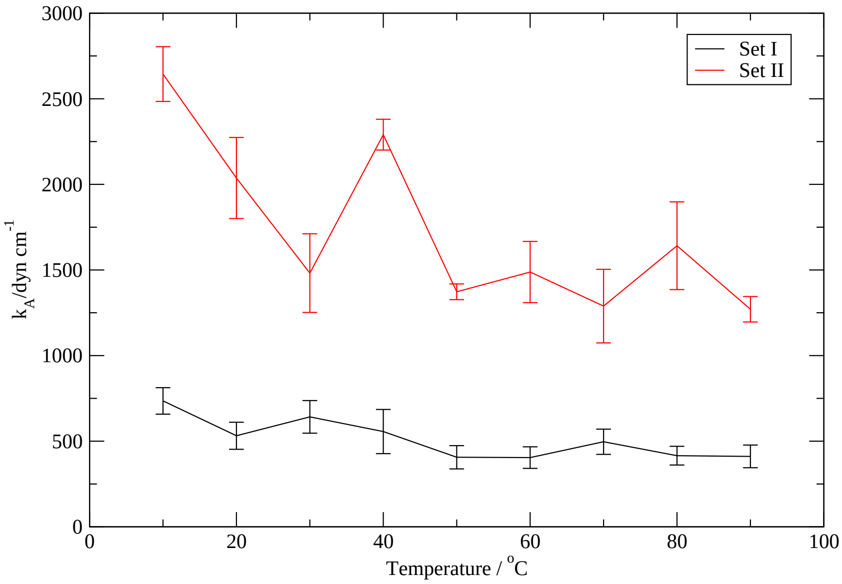

Figure 5.11: Area Compressibility variation with Temperature-standard errors are indicated in the plot

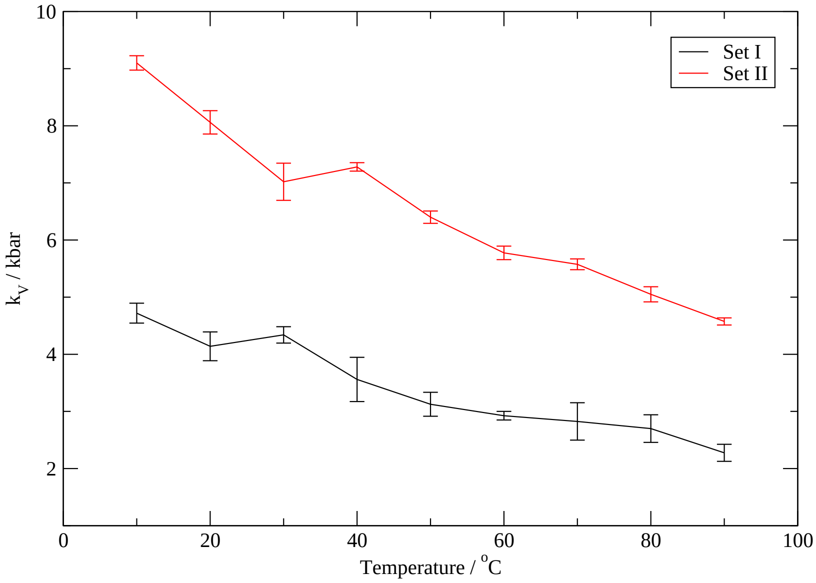

Figure 5.12: Volume Compressibility variation with Temperature-standard errors are indicated in the plot

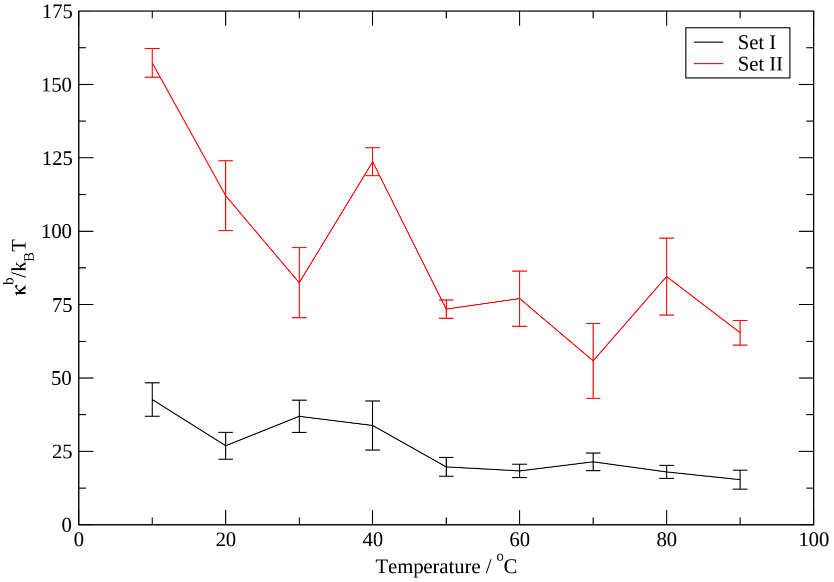

Figure 5.13: Bending Modulus variation with Temperature-standard errors are indicated in the plot

coarse grain model by Marrink et al. gave a value of $4 \pm 2 \times 10^{-20}$ J [20] and experimentally Bossev et al. report value of $4.24 \pm 0.1 \times 10^{-20}$ J (9.5 ± 0.18 k_BT) at 60 °C using neutron spin echo techniques [216]. The value calculated from this work using parameter set I is $8.79 \pm 1.4 \times 10^{-20}$ J (19.7 ± 3.2 k_BT) while parameter set II gives a value of $32.76 \pm 1.3 \times 10^{-20}$ J (73.5 ± 3.1 k_BT). Parameter set I values are in good agreement with other coarse grain [20][51]and atomistic [220] simulations and experimental [221] work for PC lipids. Like the compressibility values above this was not part of the parameterisation process. Parameter set II gives the bending modulus closer to bilayers containing cholesterol where the bending modulus is reported to be between 70-100 k_BT [116][217] .

The variation of k^b in the temperature range 10 °C to 90 °C at 10 °C intervals is given in Figure 5.13; k^b decreases with increasing temperature. The bending modulus shows a sharp decrease between 40 °C - 50 °C indicating a possible phase transition [98]. Experimentally it is known that the bending modulus increases as much 6-10 times below the melting temperature while it remains fairly constant above it, [219][222][223]; the sharp increase (the bending modulus doubles) could be indicative of a possible phase transition with the bending modulus fairly constant above 50 °C especially in parameter set I .

Area-per-Lipid and Volume-per-Lipid

The area-per-lipid and volume-per-lipid are fundamental measurable used to evaluate a lipid model. These parameters have been calculated using a whole raft of experimental techniques

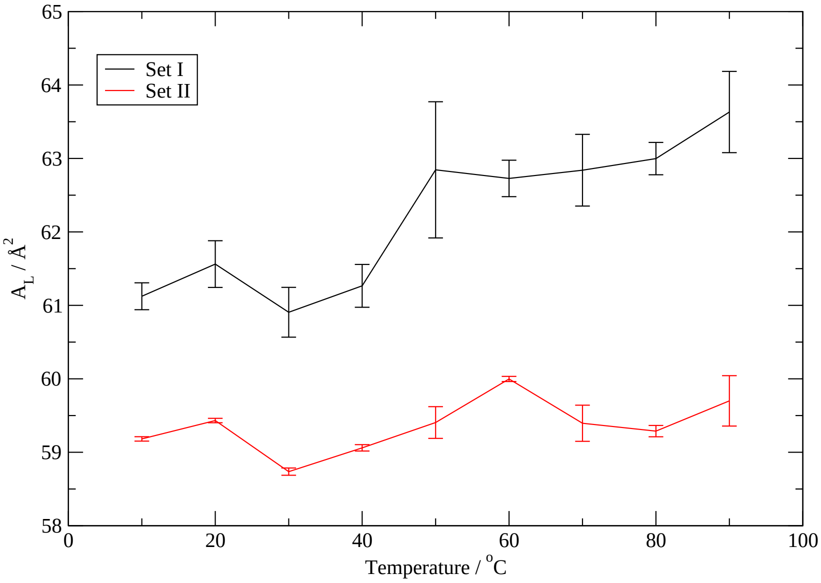

Figure 5.14: Area-per-Lipid variation with Temperature-standard errors are indicated in the plot

and have been extensively used for the parameterisation and validation of lipid models. The area and volume per lipids were computed by running a system of 450 lipids with 4500 solvent sites (27000 water molecules) under the isobaric-isothermal ensemble with anisotropic pressure coupling for over 100 ns. Four such simulations with different starting conditions were carried out and the averages computed from this. Different starting conditions were generated by running the initial random system generating simulation with different random seeds, for different time lengths and with varying temperatures and GB and LJ parameters. The computed value for the area-per-lipid for parameter set I is 62.8 ± 0.3 Å^2 and parameter set II gives a value of 59.4 ± 0.2 Å^2. The volume per lipid for parameter sets I and II respectively are 1226.75 ± 2.25 Å^3 and 1132.75 ± 2.17 Å^3 both values in good agreement with experimental and simulation values for DPPC lipids [138][20]. Experimentally the area-per-lipid is 64 Å^2 and volume-per-lipid 1232 Å^3 at 50 °C [205].

Variation of the area-per-lipid and volume-per-lipid with temperature in the range 10 °C to 90 °C at 10 °C intervals is given in Figures 5.14 - 5.15 respectively and is in line with experimental trends. The area per lipid in fairly constant in both sets of parameters between 10 °C to 40 °C then increases quite sharply between 40 °C to 50 °C; this is particularly clear in the case of parameter set I. This again indicates a possible phase transition between 40 °C to 50 °C. The volume per lipid remains fairly constant in the range 10 °C to 90 °C in line with experimental and simulation observations [98].

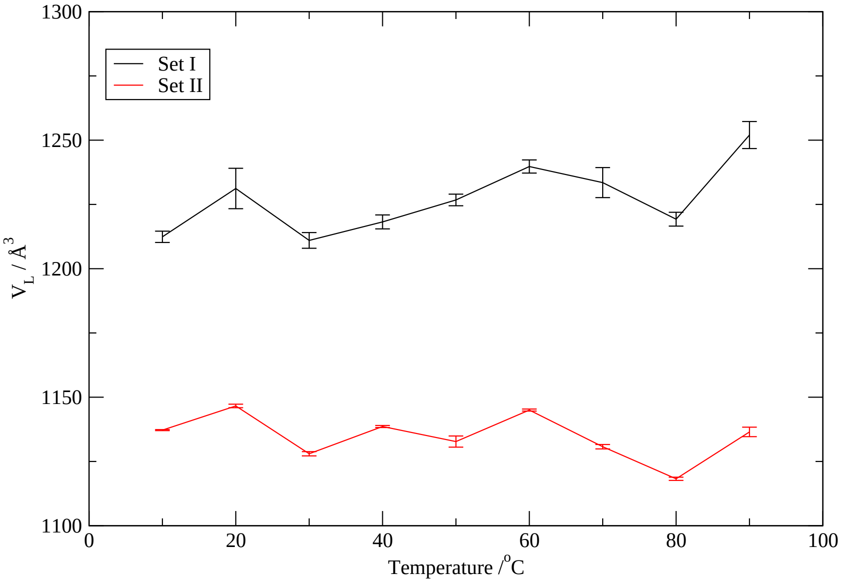

Figure 5.15: Volume-per-Lipid variation with Temperature-standard errors are indicated in the plot

Bilayer Thickness

Experimentally the bilayer thickness is measured using the electron density profile maps with the two peaks corresponding to the location of the phosphate negative charge. The bilayer thickness in the model was measured using the location of the head dipole. The bilayer thickness was measured by identifying the location of the head dipole of each lipid and binning them in 0.45 Å bins. From the plot the measured bilayer thickness is around 32.3 ± 0.93 Å for parameter set I and 33.84 ± 0.20 Å for parameter set II, in agreement with the bilayer thickness of DPPC bilayers at 50 °C measured experimentally [138] and by simulations, both atomistic and coarse grain methods [20]. Given that the ellipsoid of the lipid tail is 19.2 Å in length the bilayer thickness indicates that there is interdigitation, expected given the ellipsoidal shape of the tails. If the average result from the two parameter sets for the bilayer thickness (end-to-end ellipsoid length) is taken as 33.1 Å then the layer spacing corresponds to 0.862 l where l is the ellipsoid length , this is in good agreement with other Gay Berne liquid crystal simulations [29][31].

Variation of the bilayer thickness with temperature in the range 10 °C to 90 °C at 10 °C intervals is given in Figure 5.16. The bilayer thickness should decrease with increasing temperature (increasing temperatures increases splaying of the hydrocarbon tails); this is not so clearly visible in both sets of parameters due to the discussed non-splaying tails.

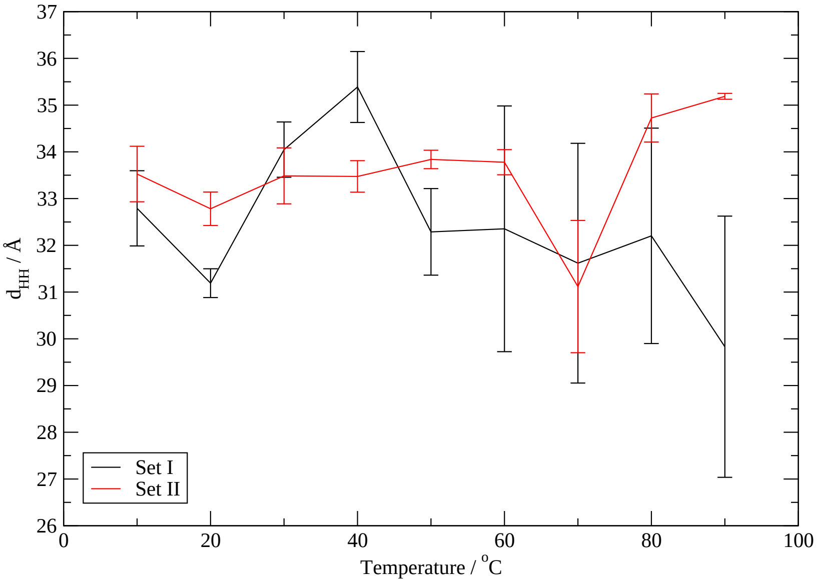

Figure 5.16: Bilayer thickness variation with Temperature-standard errors are indicated in the plot

Orientational Order Parameter

The second rank order parameter serves as a useful tool for quantifying the orientational order of uniaxial ellipsoidal tail and has seen extensive use in liquid crystal simulation during the past few decades [29][30][31][209][210]. The second rank order parameter can be calculated for a particular configuration using the expression given by Equation 5.4.

$$P_2^{(k)} = 1/N \sum_{i=1}^{N} P_2 \left(\hat{\mathbf{e}}_i \cdot \hat{\mathbf{n}} \right) \tag{5.4}$$

The orientational order parameter calculation requires the determination of the director orientation. Self assembly runs produced bilayers not aligned in any of the principal axes (x, y or z axis of the simulation box) meaning the director in the cartesian coordinates needs to be computed using the method outlined by Vieillard-Baron [207][208][209]. The method requires building of a traceless symmetric tensor **Q** for that configuration whose elements are given by Equation 5.5 where $e_{i\alpha}e_{i\beta}$ is the direction cosine of the ith lipid tail and the αth box axis. Diagonalising **Q** gives the order parameter P_2 as the largest eigenvalue and the director is given by the corresponding eigenvector. The diagonalising of the symmetric matrix was carried out using the Jacobi transformation method [224]. This method was applied to the largest lipid bilayer fragment resulting from the self assembly runs. For preassembled bilayers the director orientation is taken as the z axis and then the P_2 parameter computation simplifies and can be computed using Equation 5.6; where θ is the angle between the director and the lipid tail orientation vector and <> denotes the time average.

$$Q_{\alpha\beta} = (2N)^{-1} \sum_i (3e_{i\alpha}e_{i\beta} - \delta_{\alpha\beta}) \tag{5.5}$$

$$P_2^{(k)} =< 3\cos^2\theta - 1 > /2 \tag{5.6}$$

The order parameters computed for the self assembled bilayers and preassembled bilayers were compared to ensure consistency. Once the P_2 order parameter is calculated the it can be compared to the experimental S_{mol} of phopholipids measured using NMR [225][226] and S_{zz} obtained by atomistic simulations [211][227].

The orientational order parameter was measured at around 0.88 for parameter set I and 0.77 for parameter set II at 50 °C (for both self assembled bilayers and preassembled bilayers). While these values are extremely large compared to experimental and other simulation studies, it is to be expected from a rigid lipid model. The order parameter however does decrease with increasing temperature similar to experimental observations, but does not show a sharp decrease in the range 40 °C to 50 °C where a possible phase transition was detected. The orientational order parameter may not be sensitive enough to detect this transition.

5.8.2 Lipid Lateral Diffusion

Lipids in the bilayer are in the L_α phase and are in a fluid like state constantly moving and diffusing within the bilayer. However, they do not diffuse as a simple 2D fluid in the L_α phase since they are in the smectic A liquid crystal phase (L_α phase is a type of smectic A phase). Several theories have been proposed to explain the diffusive behaviour of lipids; the most popular being a combination of 2D liquid diffusion and cage jumping between locally confined regions in the liquid crystal lattice [106]. The lateral diffusion coefficient is measured in simulations using the Einstein relation [24][40] that measures the "true" displacement of sites. The diffusion is measured both in short (< 1 ns) and long time scales (> 50 ns) to verify the theory that lipid motion consists of local rattling at sub nanosecond regimes classified as microscopic diffusion and slower diffusive motion at longer time scales also known as macroscopic diffusion [217]. Experimental techniques can be used to measure both short and long term diffusion, techniques such as NMR and ESR for macroscopic and techniques such as fluorescence quenching and neutron scattering for microscopic diffusion [72]. A recent atomistic simulation [228] however has cast doubt on the 2D fluid diffusion and cage hopping mechanism though the run extended only for 10 ns and may be too short to observe the cage hopping mechanism, which is a conclusion borne out by the large diffusion coefficient of

$30 \pm 6 \times 10^{-8}$ cm^2 s^{-1} measured in this work which is more in line with experimental [229] and simulation [230] measurements for microscopic diffusion. The microscopic diffusion arises due to the flux in the density and movement into regions of high free volumes; this occurs at a relatively faster pace than the true displacement. It is the collective average of such movement that gives rise to the macroscopic diffusion in lipids [72]. The diffusion coefficient of preassembled bilayers at 50 °C over 80 ns was $4.0 \pm 0.1 \times 10^{-8}$ cm^2 s^{-1} for parameter set I and $2.00 \pm 0.05 \times 10^{-8}$ cm^2 s^{-1} for parameter set II, with the experimentally measured value in the range 2-8 $\times 10^{-8}$ cm^2 s^{-1} measured using fluorescence recovery after photobleaching (FRAP) and NMR [106][231], indicating that the model is able to reproduce the diffusive behaviour of lipids in a bilayer well.

To get a better idea of the overall motion of lipid in the bilayer a "stroboscopic" plot was developed as shown in several previous atomistic and coarse grain models [51][106][107][228] [230]. This is done by plotting the x, y coordinates of the COM of the lipids in the bilayer; the coordinates are the "true" positions without the application of periodic boundary conditions. In line with other simulations of both atomistic and coarse, grain models the motion occurs in two different temporal scales. The fast motion occurs as lipids move into free volumes created by fluctuations, that are smeared out over long time scales with the slower jumping motion that indicates the diffusive behaviour of lipids. The "stroboscopic" plot of three selected lipids over an 80 ns simulation are given in Figure 5.17, it clearly shows the cage rattling and subsequent hopping mechanism.

5.8.3 Lipid Packing in Bilayers

While the lipid lateral diffusion confirms the fluid nature of the bilayers the rigid nature of the single site lipid tail raises the real possibility of the lipid molecules packing together to closely giving rise to long range structural order. To confirm that the bilayers do not exhibit long range structural order along the plane the RDF of the lipids was plotted, the plot shown in Figure 5.18 confirms that the lipid bilayer is not only a fluid at 50 °C (see Section 5.8.2) but it also does not exhibit any long range structural order as shown by real lipid bilayers.

5.8.4 Summary

The developed lipid model has been demonstrated to reproduce many of the bilayer properties expected of a PC lipid having approximately 14-18 C in the hydrocarbon tail. The model is able to reproduce correctly structural parameters such as the area-per-lipid, volume-per-lipid, bilayer thickness and order parameters. It is also able to show overall agreement with

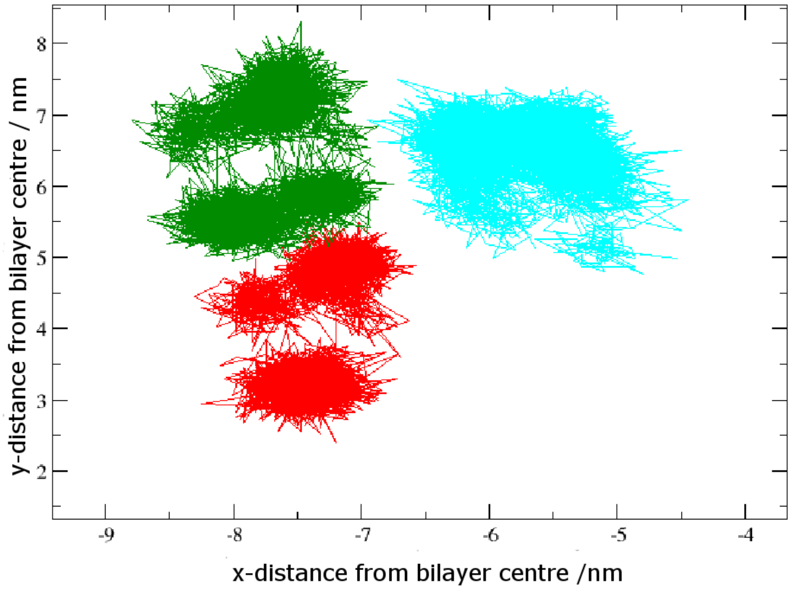

Figure 5.17: Diffusion on the X-Y place (Stroboscopic Plot)

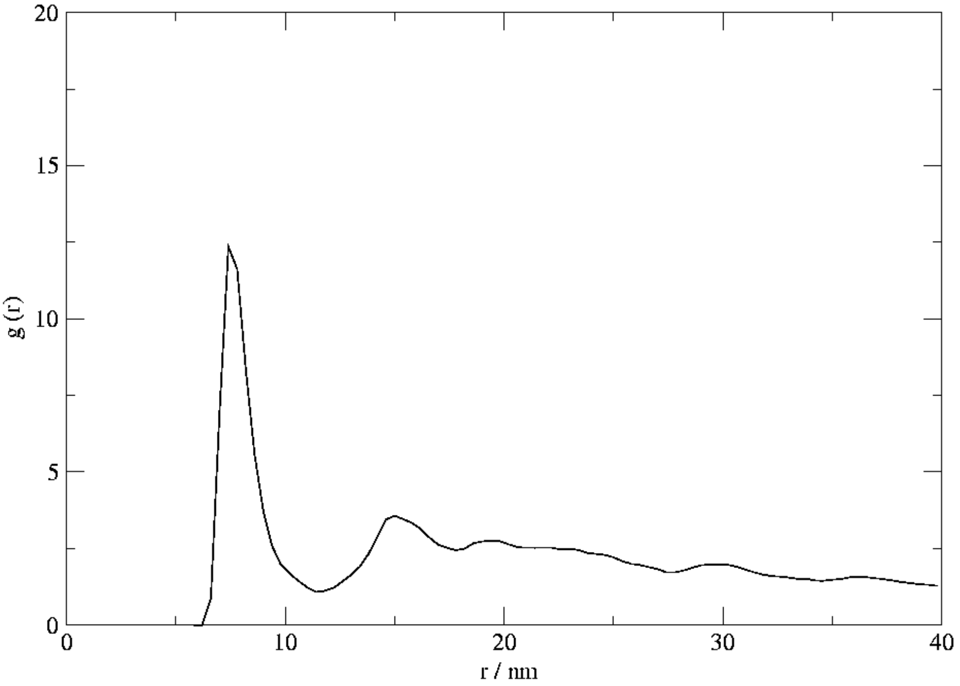

Figure 5.18: Packing of Lipids showing the Smectic A phase

the expected thermal behaviour of bilayers. The mechanical properties of the bilayer such as the compressibility and bending modulus in good agreement with both experimental and simulation studies.

5.8.5 Lipid Micelle

Micelles are formed by lipids with a packing parameter <1/3 [33][34][35]. In order to capture the correct packing parameter value, the diameter of the ellipsoid was reduced. According to the theory of packing parameters if the diameter of the cylinder enclosing the lipid is taken as the diameter of the head group then the diameter of the enclosing cylinder would be 7.4 Å (see Table 5.2) where the volume of the ellipsoid tail is computed as 1/3 the volume of the cylinder and then from it the diameter calculated. The calculated diameter was 5.2Å for the ellipsoid. All other parameters were unchanged from the lamellar forming system. Both parameter set I and II (see above) were run with the smaller ellipsoid diameter of 5.2Å. However, it was noted that the shape anisotropy introduced via the changing of the diameter of the tails did not prove to be sufficient to drive the full formation of micelles. This was thought to be due to the Gay Berne parameter $\kappa^{'}_{GB}$. For the systems forming lamellar systems, the $\kappa^{'}_{GB}$ value was set to 0.3 and 0.9 in parameter set I and parameter set II respectively. The $\kappa^{'}_{GB}$ gives the ratio between the energy minimum for the side-by-side to end-to-end ratio and a value of < 1.0 indicates that the potential favours the end to end configuration of the ellipsoid. As indicated in the lamellar section above this parameter set was chosen to form stable but liquid bilayer phases. However as seen here this leads to micelle like clusters but they remain bilayer like at the edges without curving to form perfect spherical micelles, as seen in Figure 5.19. To overcome this problem the value of $\kappa^{'}_{GB}$ was increased to 5, 10 and 15. By favouring the side-by-side interaction it was hoped that the formation of monolayer like structures could be favoured, but by having the shape anisotropy it was hoped that it would drive the system towards micelle formation. With the $\kappa^{'}_{GB}$ value at 15 the system did form micelle like structures. What is interesting though is that a series of simulations were run by varying only the $\kappa^{'}_{GB}$ and/or $scaleFac$ (a parameter which scales the interactions (Lennard Jones) between the head and solvent), all other parameters were set to the lamellar forming lipid parameter values (See Table 5.2). Increasing the $\kappa^{'}$ for the GB tails improves the formation of the micelle structures being formed. Another interesting observation is the effect of $scaleFac$ on the phase behaviour of these lipids. Increasing the $scaleFac$ to 4.0 with no increase to $\kappa^{'}$ also results in improved formation of micelles. This can be clearly seen in Figure 5.20. The resulting structure with $\kappa^{'}$=15.0 and $scaleFac$ =1.0 is given in 5.21 while increasing the $scaleFac$ to 4.0 results in micelles that look like Figure 5.22. What

$\kappa^{'}$	$scaleFac$
0.3	1.0
0.3	4.0
15.0	1.0
15.0	4.0
40.0	1.0
40.0	4.0

Table 5.8: List of varying Parameters for the Micelle

is apparent from these simulations is that while the shape anisotropy is necessary for the formation of micelles it is not sufficient for the proper formation of spherical micelles; the variation of either $\kappa^{'}$ and $scaleFac$ was necessary, with $scaleFac$ having a greater effect on the formation of the more spherical micelles. As further confirmation of these observations two sets of simulations with $\kappa^{'}$ 40.0 with $scaleFac$ 1.0 and 4.0 respectively also point to the same trends with the effect of increasing $\kappa^{'}$ less significant with further increases. A list of various micelle forming parameter sets is given in Table 5.8. Since lipids form micelles at high water concentrations these simulations were run in a system consisting of 385 lipids and 6417 solvent (38502 waters)giving a lipid to solvent ratio of 1:100.

Increasing the concentration of lipids (hence decreasing the concentration of water) was carried out to check if the system showed any lyotropic behaviour. For a system with $\kappa^{'}$ with 0.3 and with only the dimensions of the lipid altered the system consisting of 900 lipids and 4050 solvent (24300 waters) giving a lipid to solvent ratio of 1:27 gave rise to bilayer fragments. This same set of parameters with a higher water concentration gave micelle type lipids (See above system with lipid to solvent ratio of 1:100), at higher concentrations these micelle structures have merged together to form bilayer fragments as clearly visible in Figure 5.23 While the same system with $\kappa^{'}$ with 15.0 and $scaleFac$ 1.0 in a system with a lipid to solvent ratio of 1:60 results in the formation of cubic like phases as seen in Figure 5.24. All micelle self-assembly simulations started from a random configuration (confirmed by P_2 order parameter and visual observation) were run initially in the canonical ensemble for over 100,000 steps then switched to isobaric-isothermal conditions under isotropic pressure coupling for 100,000 steps then finally under isobaric-isothermal conditions with full anisotropic pressure coupling for over 300 ns. All simulations were run at 50 °C with a time-step of 40 fs.

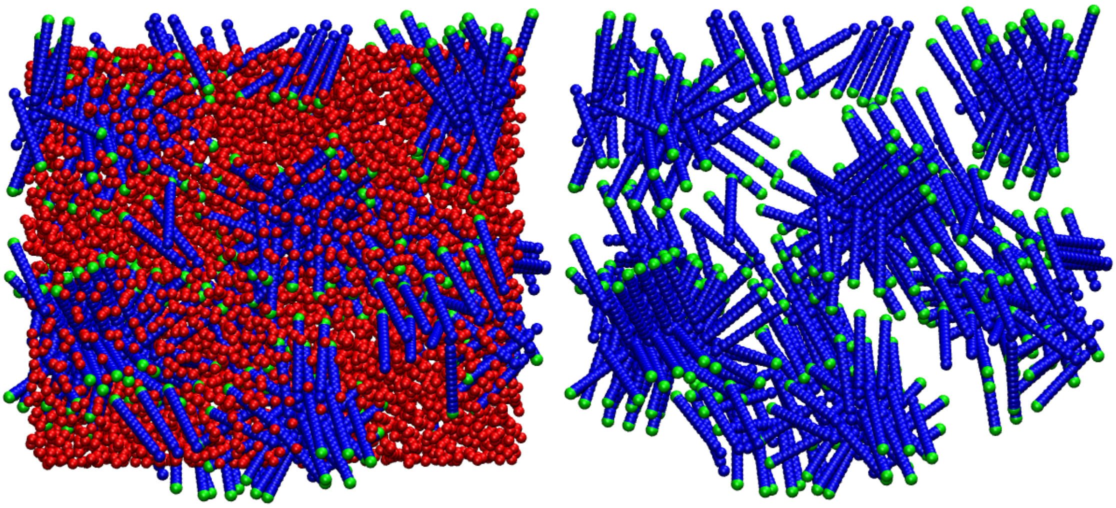

Figure 5.19: A snapshot of a micelle like structures (lipid to solvent ratio of 1:100) with $\kappa^{'}$=0.3 and scaleFac=1.0 after 200ns with and without the solvent

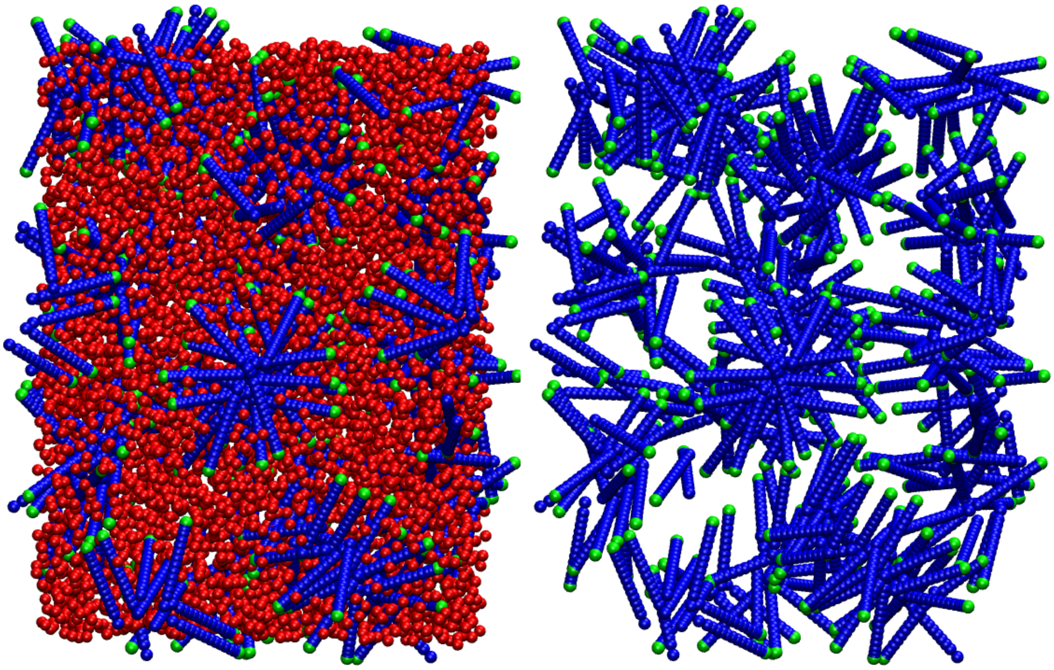

Figure 5.20: A snapshot of a micelle like structures (lipid to solvent ratio of 1:100) with $\kappa^{'}$=0.3 and scaleFac=4.0 after 200 ns with and without the solvent.

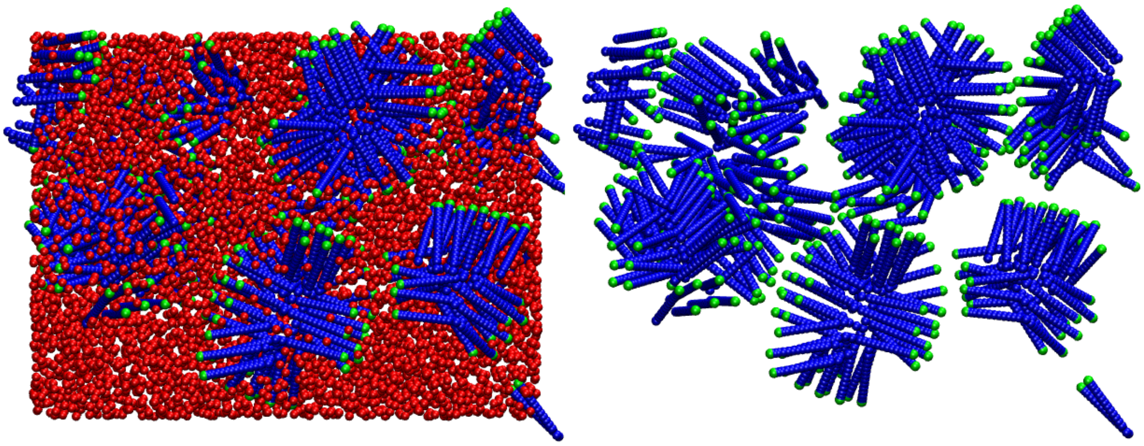

Figure 5.21: A snapshot of a micelle like structures (lipid to solvent ratio of 1:100) with $\kappa^{'}$=15.0 and scaleFac=1.0 after 200 ns with and without the solvent.

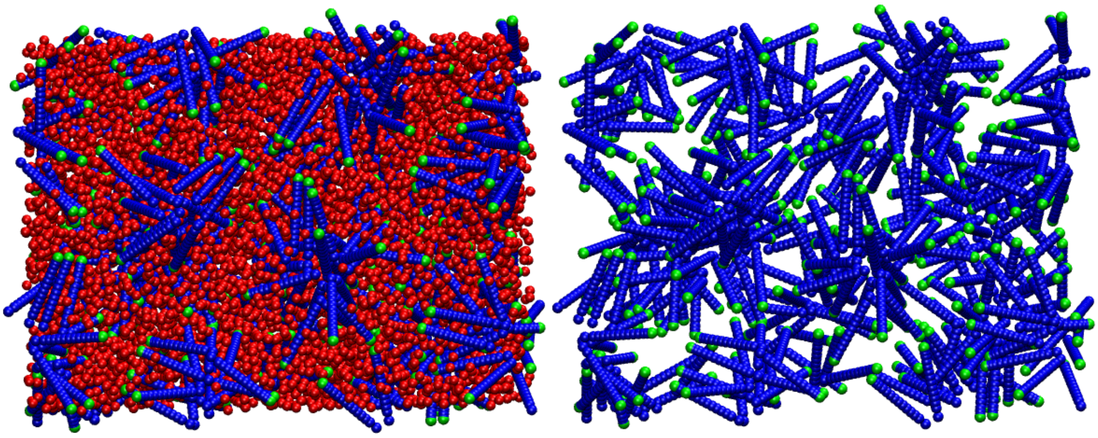

Figure 5.22: A snapshot of a micelle like structures (lipid to solvent ratio of 1:100) with $\kappa^{'}$=15.0 and scaleFac=4.0 after 200 ns with and without the solvent.

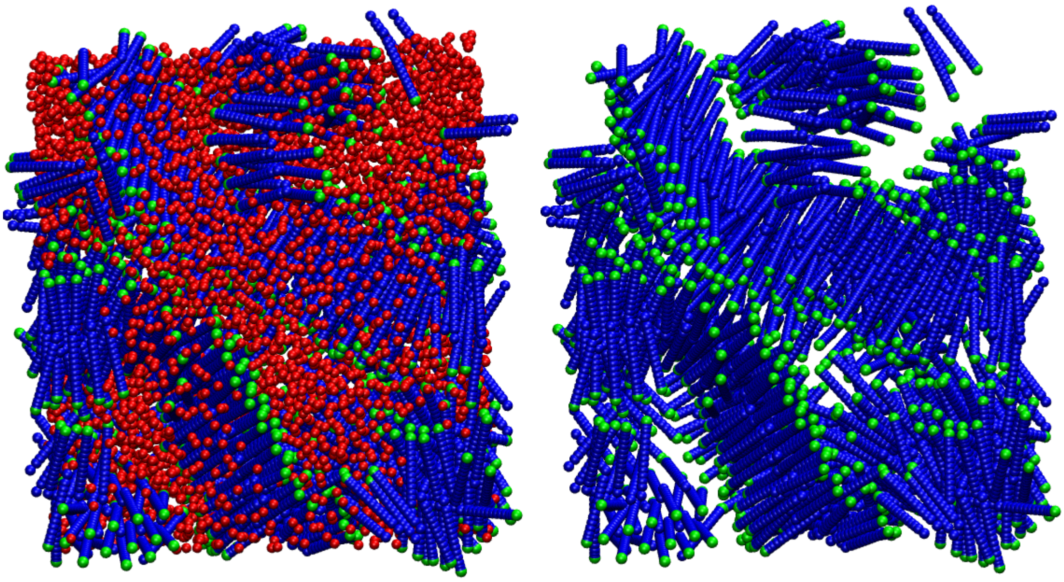

Figure 5.23: A snapshot of a bilayer like structures (lipid to solvent ratio of 1:27) with $\kappa^{'}$=0.3 and scaleFac=1.0 after 200 ns with and without the solvent.

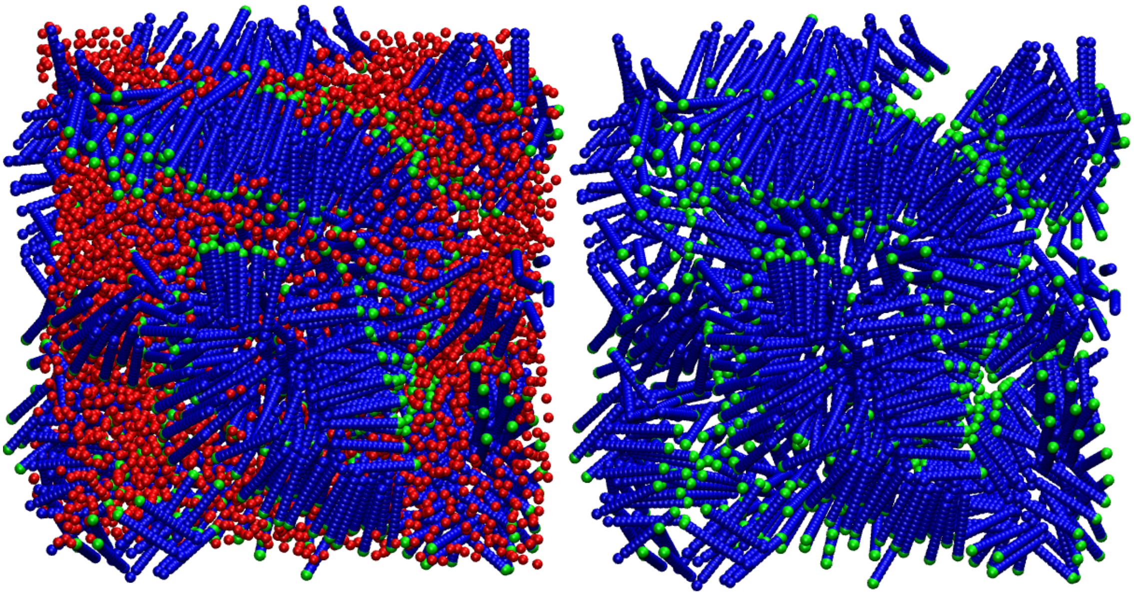

Figure 5.24: A snapshot of a Cubic structures (lipid to solvent ratio of 1:60) with $\kappa^{'}$=15.0 and scaleFac=1.0 after 200 ns with and without the solvent.

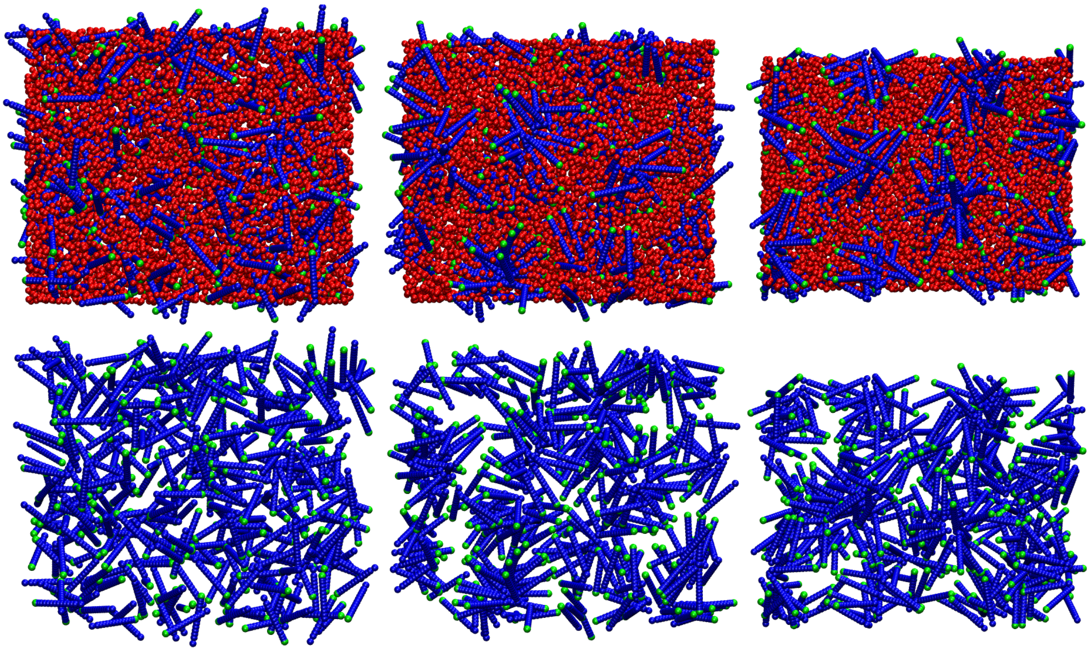

Figure 5.25: A snapshot of a 385 lipid with 38502 (6417 sites)(lipid to solvent ratio of 1:100) solvent system after 0 ns , 100 ns and 200 ns with and without the solvent showing the self-assembly of micelles.

Lipid Micelle Self Assembly

The formation of micelle-like structures are shown in Figure 5.25. The formation of micelles usually occurs at low lipid concentrations (solvent:lipid ratio 100:1) and this was simulated in this work. The system shown has a $\kappa^{'}$ of 15.0 and $scaleFac$ of 1.0. Other micelle forming parameter sets also showed similar behaviour (see Table 5.8). A clustering algorithm [40] that counts the number of lipids within a specified cutoff to determine quantitatively the size of the micelles formed was applied to the system, the average size of the micelles with $\kappa^{'}$ of 15.0 and $scaleFac$ of 1.0 was 44±4, the other micelle forming parameters also showed similar clustering behaviour. Experimental observations of dodecylphosphocholine (DPD) micelles suggest an aggregation number of about 56 [232], DPD lipids are single tailed PC lipids with twelve carbons in its hydrocarbon tail. While direct correlation between the model and DPD is not possible the aggregation number of around 44 is encouraging.

5.8.6 Lipid Hexagonal Phase

Hexagonal phase formation is quite rare in nature; it is usually formed by lipids with highly polar head groups such as lysolipids[36]. The formation of hexagonal phases is usually facilitated by the packing parameter of around 1/2 [33][34][35]. As a first step the diameter of the lipid tail was reduced to 6.8Å in accordance with the packing theory. However the system did not form hexagonal structures but lamellar structures for both parameter set I and parameter set II. This is similar to the issues faced in the micelle simulations. In addition to the $\kappa^{'}$ parameter favouring the end-to-end arrangement the small anisotropy in the diameters of the heads and tails means that the formation of bilayer structures was not surprising. All simulations were run under full anisotropic coupling as in other simulations but it is particularly important in simulations where non lamallar phases are expected to ensure that the phase formed is not forced into a particular phase in order to conform to box dimensions. All hexagonal self-assembly simulations started from a random configuration (confirmed by P_2 order parameter and visual observation) were run initially in the canonical ensemble for over 100,000 steps then switched to isobaric-isothermal conditions under isotropic pressure coupling for 100,000 steps then finally under isobaric-isothermal conditions with full anisotropic pressure coupling for over 300 ns. All simulations were run at 50 °C with a time-step of 40 fs.

Changes Introduced for Hexagonal Phase Formation

As mentioned above hexagonal phases are usually formed by lipids with highly polar head groups. To mimic this behaviour the interactions between the heads of the lipids and solvent Lennard Jones interactions were enhanced. They were increased by an input parameter called $scaleFac$ that scaled the interactions. A value of 4.0 was used for hexagonal forming systems. This was done based on the observations for micelle systems above, where enhancement of the solvent-lipid head interactions favoured the formation of micelles (see above). Increasing the $scaleFac$ alone was not sufficient for the formation of hexagonal structures; the system continued to form bilayer like structures but with increased curvature.

In order to favour the further enhance the formation of hexagonal structures the $\kappa^{'}$ value was increased to 15.0 as in the micelle systems. This resulted in the hexagonal structures as shown in Figure 5.26. These lipid parameters differ from the lamellar forming lipids in the diameter of the tail, value of $scaleFac$ and $\kappa^{'}$; all other parameters were set to the lipid bilayer forming values as given in Table 5.2.

Lipid Hexagonal Self Assembly

The formation of hexagonal phases is given in Figure 5.26 with clear hexagonal arrangement seen in the final structure. Initially the hexagonal arrangement was not seen in the system at 50 °C, instead large spherical objects with large internal spaces formed. It was not apparent if these were a result of the model/force field or if it was simply an intermediate state in

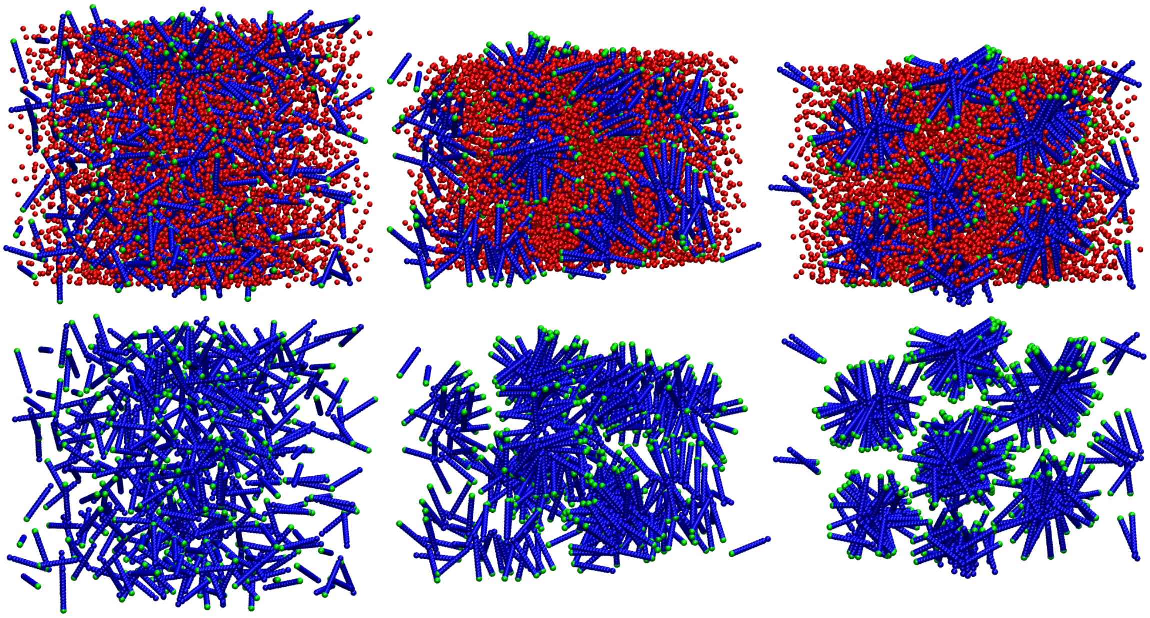

Figure 5.26: A snapshot of a 450 lipid with 27000 (4500 sites) (lipid to solvent ratio of 1:60) solvent system after 0 ns , 130 ns and 200 ns with and without the solvent showing the self-assembly of hexagonal structures.

the formation of the hexagonal phase. The simulations were continued and eventually the structures collapsed to form hexagonal phases.

A system of 385 lipids and 6417 solvent sites (38502 waters) was also run with identical parameters (a lipid to water ratio of 1:100). It is well known that that increasing hydration can favour the formation of micelle type phases. This simulation quickly formed micelle phases starting from a random start up as shown in Figure 5.27. This clearly indicates the model is capable of showing lyotropic behaviour and is in agreement with *real* lipid phase behaviour that predicts micelle structures at increased solvent concentrations. A snapshot of the micelle structures formed for hexagonal parameters at increased hydration can be seen in Figure 5.27. Since there was a possibility that the formed micelles may merge to form hexagonal phases the simulations were continued for 1000 ns , however the micelles remained stable and showed no signs of merging together.

5.8.7 Lipid Inverse Micelle

Inverse structures are formed by lipids a relatively small head group and a large tail group or at higher temperatures due to the increased thermal energy driving the splaying of the lipid tails. The effective size of the tail group may therefore drive a lamellar structure to adopt an inverse phase. Lipids show a propensity to form inverse phases at low hydration levels.

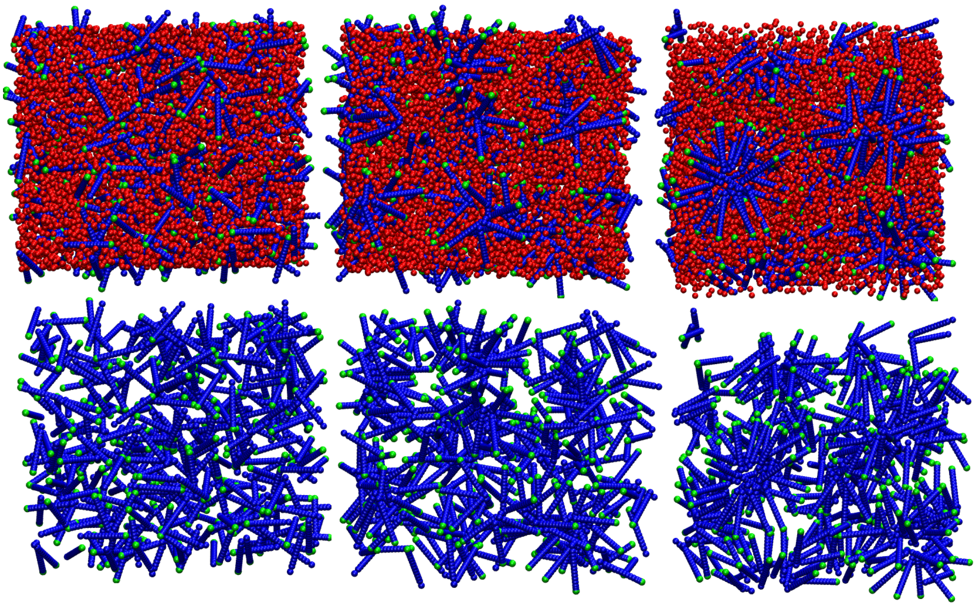

Figure 5.27: A snapshot of a 385 lipid with 38502 (6417 sites) (lipid to solvent ratio of 1:100) solvent system after 0 ns , 300 ns and 450 ns with and without the solvent showing the formation of micelle like structures.

To see if varying the packing parameter to >1 [33][34][35] would drive the lipid towards an inverse phase the size of the lipid tails were increased (with a slight reduction of the head group), while it is possible to keep the size of the tail group constant and to decrease the size of the head group , the large dipole moment embedded in the head group meant that the ability to reduce the head group diameter was limited. The lipid tail diameter was increased to 9.0 Å while the head diameter was reduced to 7.0 Å relative to bilayer forming system (see Table 5.2). Like the micelle and hexagonal systems, initial runs did not modify the κ' or the $scaleFac$ but only the diameters of the heads and tails. However unlike the micelle and hexagonal systems the shape anisotropy was sufficient to drive the formation of inverse phases.

A low solvent density system containing 500 lipids and 7200 water molecules (1200 sites) with a solvent to lipid ratio of 14:1 produced inverse miceller systems with water entrapped in the central core of the sphere under full anisotropic pressure coupling as seen in Figure 5.28. However the low density of the system resulted in the simulation becoming unstable (uncontrolled box expansion under full anisotropic pressure coupling), this instability may be caused by the highly coarse grain nature of water.

The same set of parameters were run using a higher solvent density system (solvent : lipid ration 60:1) producing systems resembling cubic phases as seen in Figure 5.29. Further

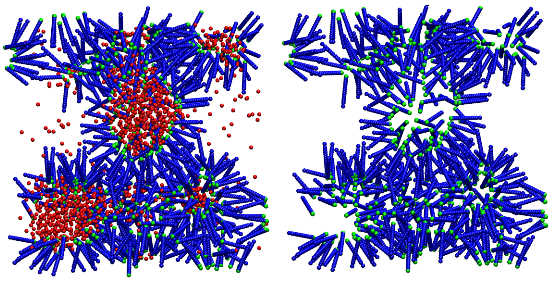

Figure 5.28: A snapshot of a 500 lipid with 7200 (1200 sites) (lipid to solvent ratio of 1:14) solvent system after 25 ns with and without the solvent showing inverse micelle like structures.

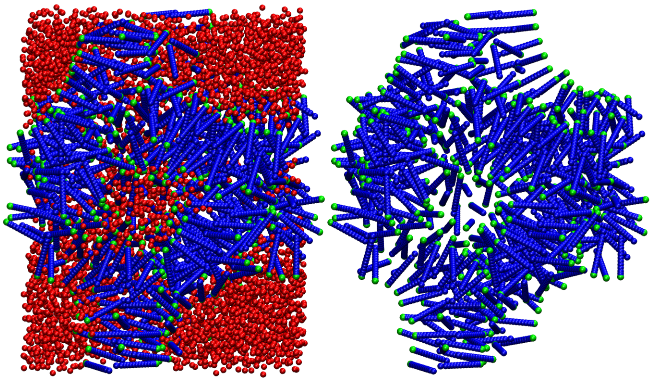

Figure 5.29: A snapshot of a 450 lipid with 4050 (1200 sites) (lipid to solvent ratio of 1:60) solvent system after 175 ns with and without the solvent showing cubic phases.

increase in the solvent concentration leading to a system of 385 lipids and 38502 waters (6417 sites) (lipid:solvent ratio of 1:100) leads to bilayer like structures forming as seen in Figure 5.30. All inverse phase simulations started from a random configuration (confirmed by P_2 order parameter and visual observation) were run under isobaric-isothermal conditions with full anisotropic pressure coupling at 50 °C with a time-step of 40 fs.

5.8.8 Summary

The model developed for simulation of lipid bilayers has been extended to capture some of the other mesomeric phases of lipids in agreement with the packing theory of lipids [33][34][35]. The model has also been to show lyotropic phase behaviour; again in agreement with experimental trends.

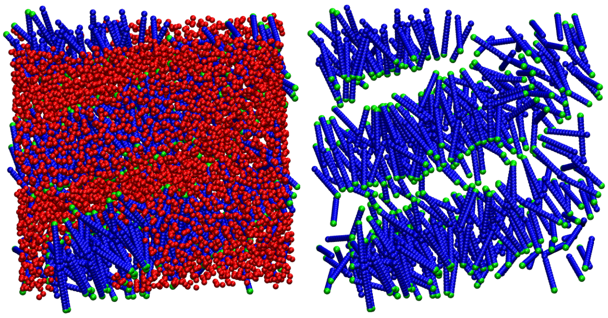

Figure 5.30: A snapshot of a 385 lipid with 38502 (6417 sites) (lipid to solvent ratio of 1:100) solvent system after 140 ns with and without the solvent showing bilayer like structures.

5.9 DNA Model Results

The developed DNA model was tested for its ability to reproduce several structural properties. The simulation set up included not only the DNA but the surrounding solvent and counter ions for charge neutrality. All these simulations were carried out with full anisotropic pressure coupling under NPT conditions at 30 °C.

5.9.1 Persistence Length

The persistence length of the DNA was calculated using Equation 5.7 where L_p is calculated persistence length, l_o the average distance between adjacent nucleotide pairs [9][233][234] in this model the distances were measured between the centre of the disk to the adjacent disk, and θ the angle between triplets of disks in the strand as shown in Figure 5.31. The persistence length was calculated using a DNA strand containing 196 base pairs (196 disks) with 68815 solvent sites(412890 solvent molecules) and 560 cations. The calculated persistence length was 61.5±0.13 nm, the experimental persistence length is around 50.0± 5 nm [96].

$$L_p = l_o/(1 - < cos\theta >) \tag{5.7}$$

The calculation of persistence length is beyond the capabilities of atomistic simulations and most coarse grain models have persistence lengths on the order of tens of angstroms while the model by Knotts et al. [19] has a persistence length of 20 nm, while the model by Khalid et al. [9] has a persistence length of ≈50 nm with external restraints required to maintain the helix structure and to reproduce the correct persistence length. Given that the persistence length did not enter the parametrisation process and the model does not require additional restraints and relies solely on the stability provided by the stacking interactions

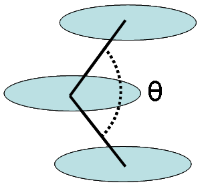

Figure 5.31: Definition of theta angle for the calculation of persistence length.

for the rigidity of the model a value of 61.5±0.13 nm is encouraging.

The simulation was run under full anisotropic pressure coupling under NPT conditions in a box of 200 Å× 200 Å× 800 Å box. Initially a box of the water model with a density of 1.0 g cm^{-3} was run under full anisotropic pressure coupling and then a DNA strand consisting of 196 base pairs inserted in the center of the box and any overlapping solvents removed. To ensure charge neutrality 560 solvent sites were replaced by cations. To prevent any problems due to possible overlap between the solvent and DNA base pairs the simulation was started with the water dipole removed and the size of the solvent sphere reduced and then grown in stages as in previous lipid simulations (See above). The replacement of 560 solvent sites by ions was done in a random fashion.

Initial attempts consisted of inserting a straight ladder like DNA strand, however due to its longer uncoiled length a larger initial simulation box was required. In addition the large uncoiled DNA tended to become entangled and the large system size prevented its possible complete untangling and subsequent coiling. To overcome this problem a DNA strand with the correct helix structure was inserted into the solvent box. The long DNA structure with the correct helix structure was constructed from a smaller strand of 22 base pairs by replicating the structure (with z coordinate displacement) to obtain a strand of 196 base pairs.

5.9.2 Helix Repeat

The helix repeat was calculated by finding the dot product between the first nucleotide vector $\boldsymbol{\mu}$ and the subsequent ones along the strand, the ellipsoid director e and the vector $\boldsymbol{\mu}$ are placed perpendicular to each other as shown in Figure 5.32, the dot product calculated between the vectors $\boldsymbol{\mu}$ along the strand. The $\boldsymbol{\mu}$ vector was defined as the vector extending from the center of the disk to the phosphate backbone. The dot product was taken along the stand on one side of the the vector extending from the center of the disk to one of the phosphate backbones. The dot product will become 1.0 when the DNA strand makes a complete turn. The calculated value for a complete helix turn in the model is 34.45 ± 0.69 Å in good agreement with experimental values from X-ray crystallography that gives values around 34 Å. The number

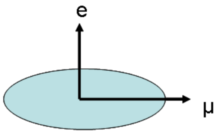

Figure 5.32: The definition of the orientation vectors in the DNA disk.

of base pairs in one complete turn of the DNA helix gives the helix repeat of the model, the value obtained for the model was 9.4 ± 0.16 while experimentally it is 10 in the canonical B DNA [37]. Initial simulations with no parameterisation (solely based on structural factors) of the helix repeat gave a value of around 11-12, the value was improved by slight adjustment of the height of the disk (reduction of height from 3.5 Åto 3.3 Å) and the bond length between backbone sites(bond length increased from 6 Å to 7 Å).

The simulation set up included a DNA strand of 22 base pairs (22 disks) 60 positive counter ions and surrounding solvent in a 100 Å× 100 Å× 100 Å box. The DNA strand was inserted as a straight ladder in a box of water with a density of 1.0 g cm^{-3} and any overlapping solvent removed. For the insertion of the counter ions, 60 solvent sites were replaced by counter ions in a random fashion. To prevent any overlap clashes the DNA was restrained (both translational and rotational motion) and the solvent dipole switched off and their diameter reduced and the system run in the canonical ensemble. This allowed the system to relax and any overlaps to be fixed. Then the restraints on the DNA removed and the system run unconstrained. This allowed for the DNA inserted as a straight ladder to coil. The coiling was observed to be initiated at both ends of the strand as in the model by Voth et al. [22]. The lack of chirality saw the formation kinks when the coiling at both ends was initiated in opposite directions; which was overcome with the introduction of a torsion potential (See above Section 5.7.4).

5.9.3 Ion Environment

The negative charges around the DNA backbone ensures that it is surrounded by positively charged ions. The ion environment around the DNA has been investigated using several experimental and modelling techniques. These indicate that while the ions are around the DNA they do not penetrate the grooves often and their residence times are very low [80][82][83]. That is a very diffusive ion cloud exists around the DNA. The diffusion of ions around the developed model was calculated using the Einstein relation obtaining the "true" displacement of ions [40]. During this process the DNA translational motion was restrained to prevent the motion of the DNA contributing towards the net diffusion of ions. The calculated value for the diffusion coefficient is 0.7×10^{-9} m^2 s^{-1} while all atom simulations have given a value in the range 1.3-2.1$\times$ 10^{-9} m^2 s^{-1} [83].

To better understand the ion environment around the DNA the DNA translational diffusion was removed allowing for rotational motion of the disks and then the ion motion over a 10 ns time period was superimposed around the DNA strand and displayed in Figure 5.33. What is clearly indicated in Figure 5.33 is that while the ions preferentially surround the DNA they also move away from the DNA periodically, that is the ions are in a diffusive state which confirms earlier simulation studies and experimental work [80][82][83]. This diffusive nature of ions is further confirmed by the "stoboscopic plots" of all the ions given in Figure 5.34 showing the XY diffusion on the XY plane of the simulation box.

5.10 Conclusions

The solvent model has been developed to capture the correct density, diffusive behaviour and dielectric constant. The lipid model developed and parameterised has been shown to be realistic and able to reproduce many of the structural, mechanical and dynamic properties of a phosphatidylcholine lipid bilayers and by simple adjustment of parameters capture several other lipid phases. The lipid model was also demonstrated to show lyotropic behaviour expected of lipids with variation of solvent concentration. The DNA model is able to reproduce many of the structural properties of the most common canonical form of DNA, the B form, including its persistence length. The ions in the models are a crude representation of cations in nature and have been shown to have realistic dynamic behaviour around the DNA strand. The developed models are robust and realistic enough and have provided a good starting point to mix these models together and study lipid mixtures and lipid-DNA mixtures.

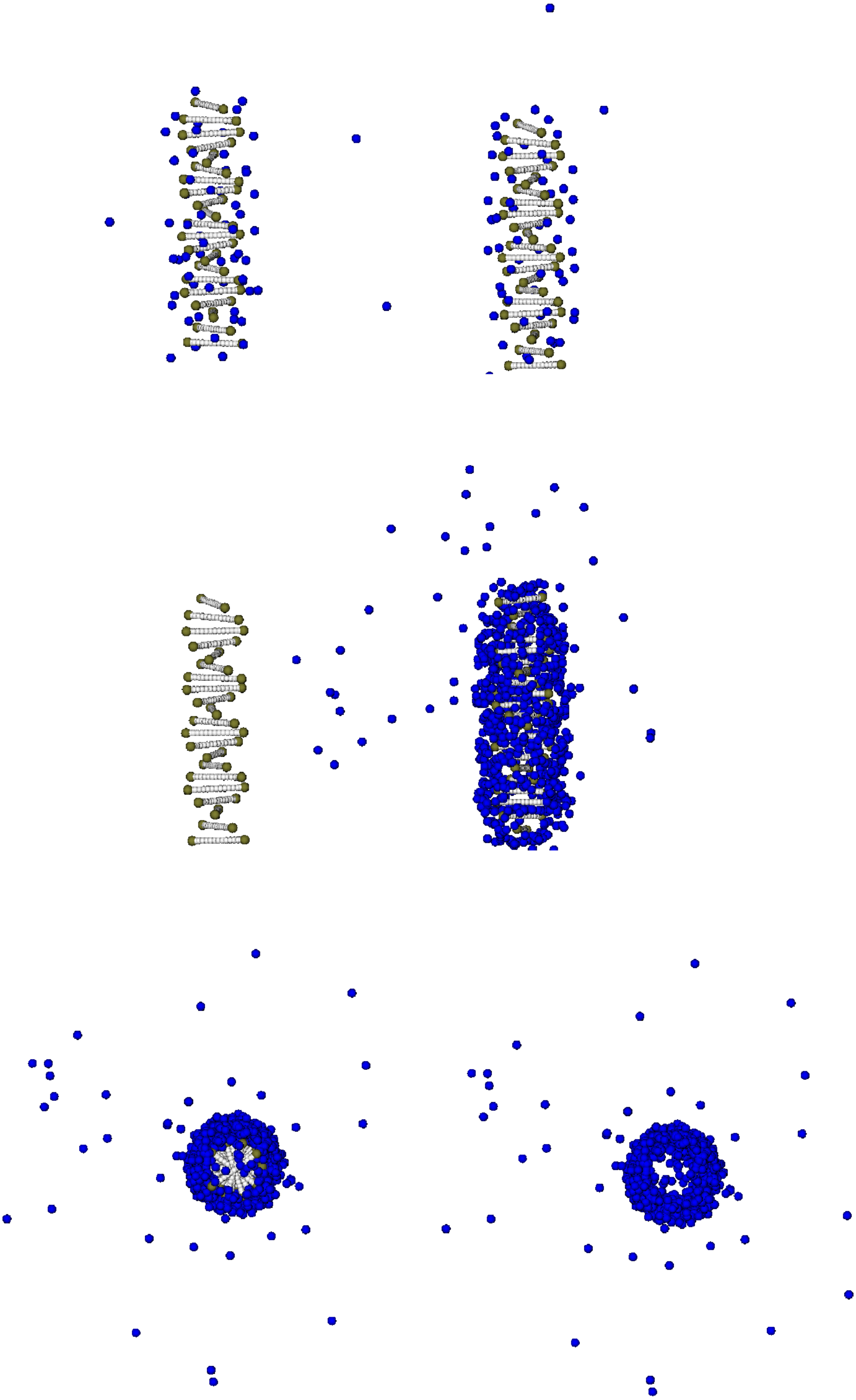

Figure 5.33: Superimposed ion positions over 10ns

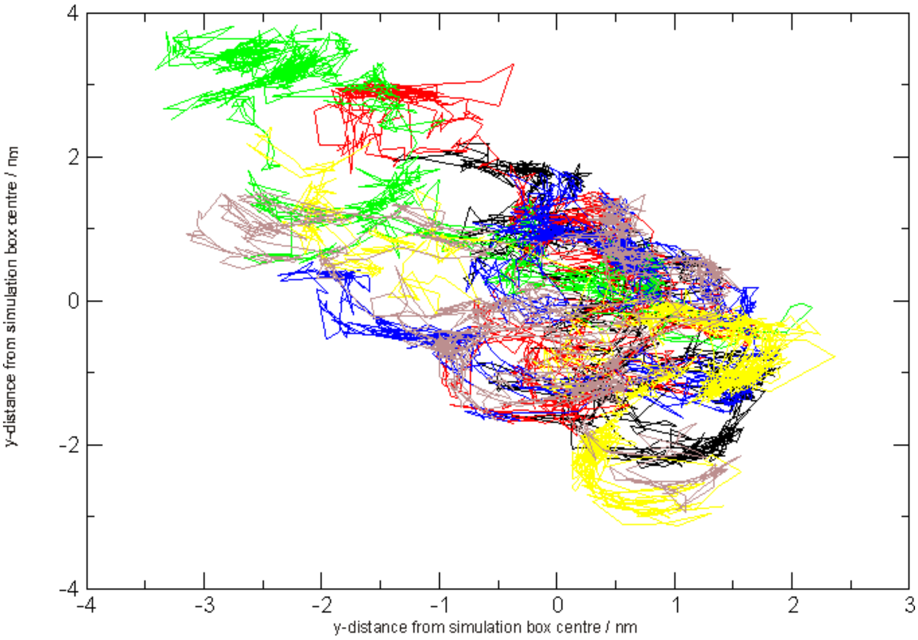

Figure 5.34: Stroboscopic plot of ion diffusion

Chapter 6

Mixtures of Different Types of Lipids

6.1 Introduction

Lipids are primarily found in the cell membrane, in addition to the intracellular organelles of a cell. In addition, as mentioned in Chapter 1 (Introduction), several types of lipids, in particular PC lipids are to be found in the nuclear matrix. These encompass a wide range of lipids with varying structures and functions at different concentrations. The existence of different types of lipids in this organelle is presumably no mere accident but nature uses this to control the structure and function these lipids play in the cell. Given the important role a mixture of lipids play in the cell membrane it is unfortunate that many of the lipid simulation models have been limited to single component lipid models. This is primarily due to the complexity of including multicomponent species in the simulation in sufficient numbers to see their effect.

One of the advantages of simplified models that has been alluded to before is the ability of using such models to look at effects that a second, or third component have on the behaviour of a lipid model. This is by no means a new approach, lipid mixtures have been simulated previously and has provided valuable insights on the effects of the second/third lipid component on lipid properties.

6.2 Mixing of Unlike Lipids

To simulate different types of lipids the simulation software needed to be extended to handle Generalised Gay Berne potential. The original Gay Berne potential is only able to handle interactions between identical ellipsoids (or identical LJ spheres under limiting conditions). It has been extended to handle interactions between unlike ellipsoids by Cleaver et al. [42]. However, there are several concerns with regard to the generalised Gay Berne potential, the primary being that the generalised potential does not tend to the *original* version under limiting conditions of identical ellipsoids. This was overcome by having two potentials in the systems and depending on whether the interaction was between like and dislike particles different parts of the code were executed. Explicit formulae for the generalised Gay Berne potential has been listed in Chapter 2.

6.3 Micelle Forming Lipids

The micelle forming lipids with a $\kappa^{'}$ of 15.0 and $scaleFac$ of 1.0 which were introduced in the last chapter (See Chapter 5 Section 5.8.5 Table 5.8) were mixed as the second component in system consisting of bilayer forming lipids (Parameter Set I/II - See Chapter 5 Section 5.7.3). A comparison of the two types of lipid parameters is given in Table 6.1. Two sets of simulations were carried out, one on a preassambled bilayer and the other self assembly simulation runs. All simulations were run at 50 °C under NPT conditions with full anisotropic pressure coupling using a timestep of 40 fs.

6.3.1 Comparison of the two types of lipids

6.3.2 Preassembled Bilayers

A set of preassembled bilayers simulations were run with varying concentration of micelle forming lipids. The system was set up by replacing 10% of and 30% of the bilayer forming lipids (Parameter Set I and Parameter Set II-See Chapter 5 Section 5.7.3) with the micelle forming lipids. The set up was done by obtaining the coordinates of the preassembled bilayer used for the calculation of bilayer properties (See Chapter 5). Each monolayers contained a random number of the second component lipid. This was done by replacing randomly some of the bilayer forming lipids with the micelle forming lipids until the system contained 10% and 30% lipids. The system consisited of 450 lipids and 24300 waters (4050 sites) with a lipid to solvent ratio of 1:60. The simulation box was initially 120 Å× 120 Å× 120 Å as in the

$Parameter$	$Bilayer\ Value$	$Micelle\ Value$
σ_{Tail}/Å	7.0	5.2
ϵ_{Tail}/kcal mol^{-1}	0.5	0.5
κ_{Tail}	2.74	3.692
$\kappa^{'}_{Tail}$	0.3/0.9	15.0
μ_{Tail}	2.0	2.0
ν_{Tail}	1.0	1.0
$\kappa^{'}_{GBLJ}$	0.2	0.2
σ_{Head}/Å	7.4	7.4
ϵ_{Head}/kcal mol $^{-1}$	0.9	0.9
$dipole_{Head}$/D	18.0	18.0
$mass$/amu	800.0	800.0
$HeadDistance$/Å	9.6	9.6

Table 6.1: Bilayer and Micelle Lipid Head and Tail Parameters

bilayer lipid simulation setup.

The system was equilibrated under the canonical ensemble and then switched to constant isobaric isothermal conditions (isotropic) before measurements were made under full anisotropic isobaric isothermal conditions as in the bilayer forming systems (see Chapter 5).

A comparison of some properties of the bilayer with the bilayer made up of a single component lipid are given in Table 6.2 and Table 6.3. One of the most interesting observations is the significant drop in the volume compressibility. This indicates that the bilayer is becoming significantly softer with the introduction of the second component. This is seen with both Parameter Set I and Parameter Set II but also quite significantly the effect of the second substituent is similar in both sets. A further increase in micelle type lipids to 30% has no effect on the bending modulus.

A possible qualitative explanation of this effect could be that, the thin tails in the micelle lipids creates voids in the hydrocarbon region of the membrane and as a results allows for greater fluctuations in the tail region thereby softening the bilayer. The addition of a second substituent to a bilayer has been shown to decrease the bending modulus and soften the bilayer in experiments where short surfactants (pentanol) addition to a DMPC bilayer decreased the bending rigidity [235] where the authors assign the decrease to membrane thinning.

Parameter	100%	90%	70%
K_A/dyn cm^{-1}	1372.73 ± 45.73	1047.27 ± 53.64	1111.33 ± 261.64
K_V/k bar	6.4 ± 0.11	0.2 ± 0.00	0.167 ± 0.03
κ^B/J	8.79 ± 1.4	23.33 ± 1.66	27.54 ± 0.01
V_L/Å^3	1132.75 ± 2.17	1027.67 ± 1.85	983.67 ± 3.38
A_L/Å^3	59.40 ± 0.21	56.49 ± 0.04	53.34 ± 0.58
d_{HH}/Å	33.84 ± 0.20	32.67 ± 0.43	34.58 ± 0.34

Table 6.2: Variation of bilayer properties with varying concentration of Parameter Set I lipid

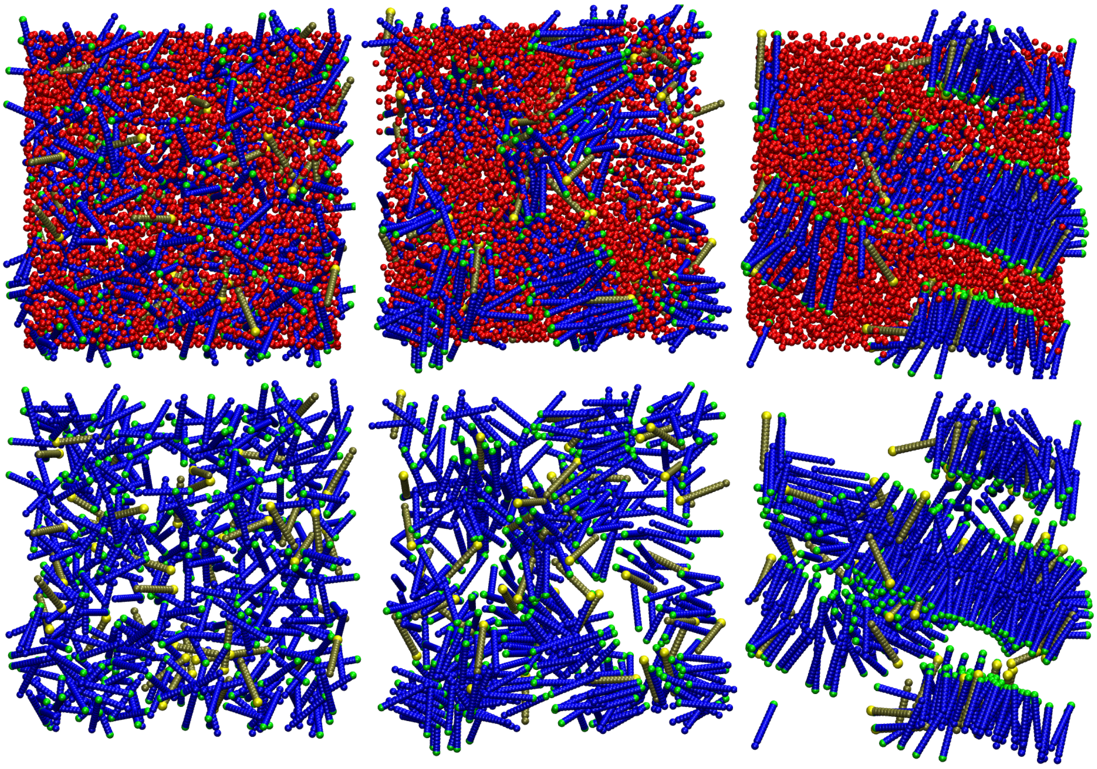

Figure 6.1: A snapshot of a 45 micelle forming lipids and 405 Parameter Set I bilayer lipids with 27000 (4500 sites) solvent system after 0 ns , 6 ns and 60 ns with and without the solvent

Parameter	100%	90%	70%
K_A/dyn cm^{-1}	406.23 ± 67.92	611.67 ± 82.23	822.42 ± 197.24
K_V/k bar	3.12 ± 0.21	0.2 ± 0.03	0.2 ± 0.03
κ^B/J	32.76 ± 1.3	15.47 ± 2.15	21.14 ± 5.00
V_L/Å^3	1226.75 ± 2.25	1122.00 ± 3.32	1071 ± 8.29
A_L/Å^3	62.84 ± 0.93	57.60 ± 0.08	57.35 ± 0.45
d_{HH}/Å	32.29 ± 0.93	34.83 ± 0.27	34.26 ± 0.22

Table 6.3: Variation of bilayer properties with varying concentration of Parameter Set II lipid

6.3.3 Self Assembly of Mixtures

Self assembly runs of mixtures containing 10% and 30% percent of micelle forming lipids were carried out. These simulations, as in other self assembly runs, started from a random configuration. This was confirmed by the P_2 order parameter and visual inspection. One observation, though not quantified, was a possible faster self assembly as compared to the single component bilayer forming lipids. The self assembly of these lipids is seen in Figure 6.1 where Parameter Set I lipid ellipsoids are shown in dark blue and micelle forming lipids in gold. The system consists of 405 Parameter Set I lipids and 45 micelle lipids. The results for other systems (30% micelle with Parameter Set I and 10% and 30% micelle with Parameter Set II) show similar trends. However, the bilayer is more fragmented with Parameter Set II, an observation seen in pure lipid systems containing Parameter Set II (Chapter 5) in comparision to Parameter Set I lipids, possibly due to the stronger affinity in Parameter Set I ellipsoids to arrange end-to-end than the Parameter Set II.

6.4 Inverse-Phase Forming Lipids

The inverse phase forming lipids which were introduced in the last chapter (See Chapter 5 Section 5.8.7) were mixed as the second component in the systems. A comparison of the parameters for the two lipid types (inverse and bilayer forming lipids) is given in Table 6.4. All simulations were run at 50 °C under NPT conditions with full anisotropic pressure coupling using a timestep of 40 fs.

6.4.1 Comparison of the two types of lipids

$Parameter$	$Bilayer\ Value$	$Inverse\ Lipid\ Value$
$\sigma_{Tail}/\text{Å}$	7.0	9.0
$\epsilon_{Tail}/\text{kcal mol}^{-1}$	0.5	0.5
κ_{Tail}	2.74	2.1333
κ'_{Tail}	0.3/0.9	0.3/0.9
μ_{Tail}	2.0	2.0
ν_{Tail}	1.0	1.0
κ'_{GBLJ}	0.2	0.2
$\sigma_{Head}/\text{Å}$	7.4	7.0
$\epsilon_{Head}/\text{kcal mol}^{-1}$	0.9	0.9
$dipole_{Head}/\text{D}$	18.0	18.0
$mass/\text{amu}$	800.0	800.0
$HeadDistance/\text{Å}$	9.6	9.6

Table 6.4: Bilayer and Inverse Lipid Head and Tail Parameters

6.4.2 Self Assembly of Mixtures

Self assembly runs of mixtures containing 10% and 30% percent of micelle forming lipids were carried out. These simulations as in other self assembly runs started from a random start up. This was confirmed by the P_2 order parameter and visual inspection. The self assembly of these lipids is seen in Figure 6.2 where Parameter Set I lipid ellipsoids are shown in dark blue and inverse phase forming lipids in gold. The system consists of 405 Parameter Set I lipids and 45 inverse phase lipids. The results for other systems (30% micelle with Parameter Set I and 10% and 30% micelle with Parameter Set II) show similar trends.

6.5 Bicelles and Rafts

The simplicity of the models developed and the ease in which different types of lipids can be mixed in this model have raised the possibility of using it to study bicelles and rafts. Bicelles (refer Chapter 3 Section 3.15 for details of bicelle structure and use) are now increasingly used in experimental study of lipid-protein systems though there is still some controversy as

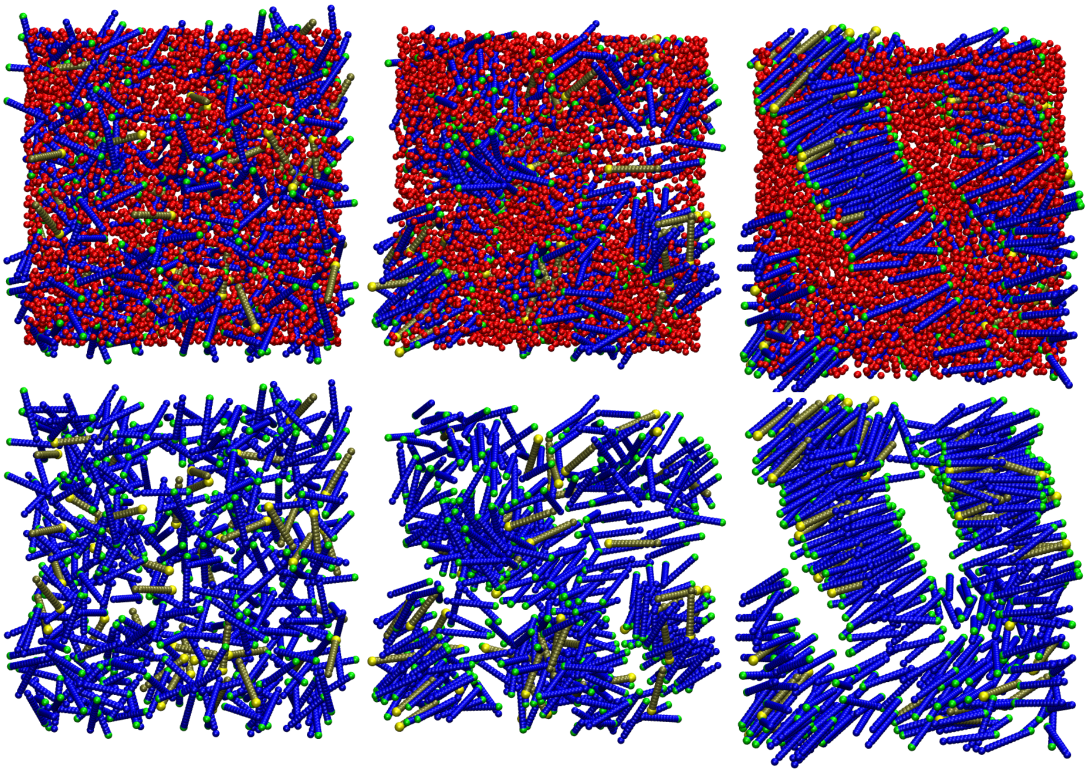

Figure 6.2: A snapshot of a 45 inverse phase forming lipids and 405 Parameter Set I bilayer lipids with 27000 (4500 sites) solvent system after 0ns , 5ns and 50 ns

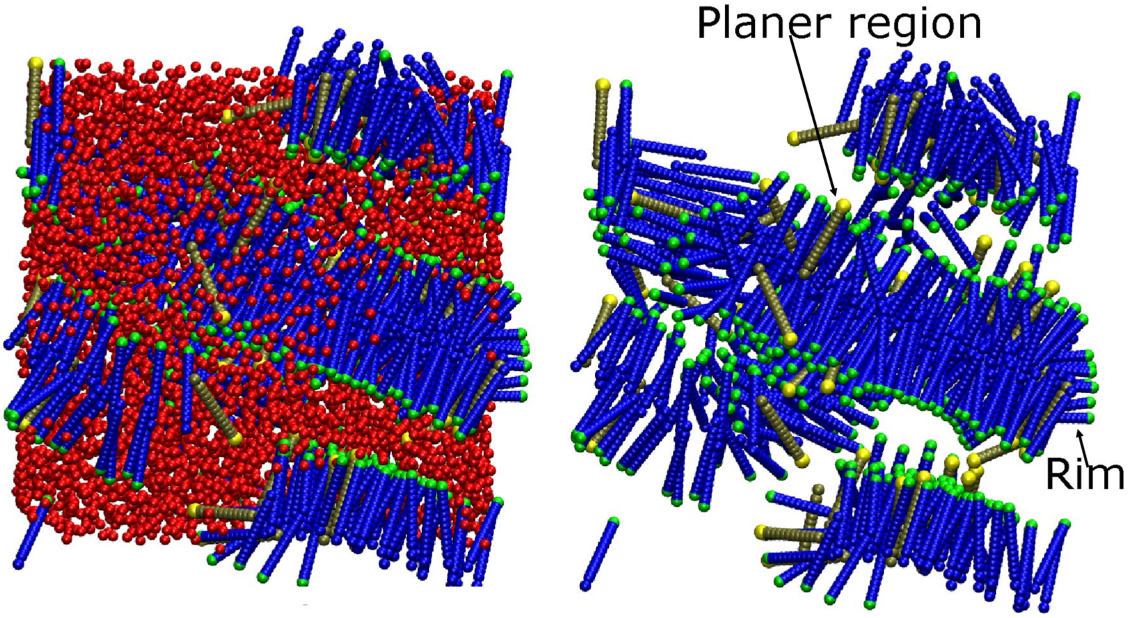

Figure 6.3: A snapshot of a 45 micelle forming lipids and 405 Parameter Set I bilayer lipids with 27000 (4500 sites) solvent system after 60 ns with and without the solvent showing (possibly) initial formation of bicelle like structures

to their structure. The mixing of micelle forming lipids and lamellar type lipids in this study have already shown structures which possibly could be precursors to bicelles as highlighted in Figure 6.3. Unfortunately the simulation was only extended for 60 ns; far too short to see the possible full self-assembly of bicelles. Future work in this area would involve extended runs of atleast 1μs to see if this model is able to form bicell like structures.

The model is easily extendable to represent molecules such as cholesterol (small head-group large bulky tail group) and could be incorporated with existing lipid models to look at the possibility of raft and domain formation as highlighted in Chapter 7 under future extensions of this model. For structural and functional details of rafts, refer Chapter 3 Section 3.14.

6.6 Conclusions

It has been demonstrated in this chapter that using simple models of lipids it is possible to study mixtures of various lipids and better understand the effect of concentration and the type of second substituent (or several substituent) have on the single component lipid behaviour. The simplicity of the models used allows for systematic study of lipid mixtures to better understand more complex biological systems found in nature. These models also give rise to the possibility of studying the formation of domains and rafts.

Chapter 7

Conclusions and Future Possibilities

7.1 Conclusions

Highly coarse-grain lipid and DNA models have been developed from scratch using simple physical principles that includes the incorporation of the correct electrostatics of the solvent and solute. The developed lipid model has been demonstrated to have the capability to form bilayer phases spontaneously and has through simple realistic parameter changes based on the packing parameter [33][34][35], been shown to form many other phases that lipids in nature show such, as the hexagonal, cubic, micelle, and inverse phases. Many of these phases are yet to be captured using a single model such as the one developed here. The model is also remarkable in that it is able show the lyotropic phase behaviour expected for lipids with changing concentration of solvent.

The spontaneously formed bilayer has been shown to posses many of the mechanical, dynamic and structural properties applicable to neutral phosphatidylcholine lipids in nature with around 14-20 carbons in the hydrocarbon tails. This is remarkable considering the simplicity of the model developed. This was made possible by carefully considering the most important aspects of the lipid model that are relevant to the problem at hand followed by careful and repetitive parameterisation.

One of the most difficult problems that were faced in parameterising this model was the need to have a system that self assembles into bilayers but one that avoids the formation of static gel like or solid structures with long range structural order. The path taken in discovering a set of suitable parameters that satisfied certain caveats such as the above mentioned dynamic and liquid structure, the ability to self assemble, and to reproduce certain fundamental structural and mechanical properties applicable to phosphatidylcholines is, outlined in Chapter 5.

Chapter 5 also details the behaviour of the lipid model with varying temperatures. While the model was able to show overall trends expected for a phosphatidylcholine lipid, such as the decrease in bending modulus with increasing temperature, the absence of two tails in the hydrocarbon region limits the quantitative capabilities of such models. Possible improvements to the model are suggested below.

A highly coarse-grain DNA model, designed on the same simple physical principles based on identifying the very essential features required to capture some of the structural properties of the DNA such as helix repeat and persistence length, have been discussed in Chapter 5, including the path taken in its development.

The possibility of using such simple models in the study of lipid mixtures, and the importance of such lipid mixtures in nature has been highlighted in Chapter 6.

7.2 Future Work

The model developed has very many future applications, in its current form as well as with further extensions and improvements. There are also a number of challenges that need to be overcome in extending the models as discussed below.

7.2.1 Current Applications

The list below gives the possible applications using the models developed.

- Effect of lipid length on the phase behaviour of lipids;
- Effect of head group dipole magnitude on phase behaviour;
- Determination of the critical micelle concentration;
- Effect of charges on bilayer stability and phase behaviour;
- Effect of lipid composition on lipid properties;
- Formation of lipid rafts and domains;
- Formation of bicelle like structures using lipid mixtures;
- Mechanical behaviour of DNA strand under stretching forces;
- Simulation of lipid-DNA mixtures to better understand the interactions between lipids and DNA.

7.2.2 Future Extensions

The list below gives the possible extensions that can be implemented in the lipid and DNA models.

- Allowing for the head dipole vector to be freely rotating, better capturing the dynamic behaviour of the head group;
- Extension of the hydrophobic tail site to be represented by two ellipsoids, better reflecting the thermal behaviour of phosphatidylcholine lipids;
- Representing the DNA base pairs as two GB disks, allowing the study of DNA melting and base pairing.

7.2.3 Challenges

The most difficult aspect of simulating lipids, developing a model that self assembles spontaneously while retaining the physical properties of lipids has been the mixed interactions between the solvent and the lipid head and tails. These mixed interaction scaling values were discovered through time consuming and repetitive simulations using a trial and error method. Extending the model to study lipid-DNA mixtures would be a greater challenge because of the complex interplay expected between the various species in the simulation setup such as ions, DNA, lipids and the solvent. The fact that the model contains particles of non-spherical shape makes this process more complicated. As a first step in simulating lipid-DNA mixtures (and done with lipid mixtures) simple mixing rules such as the Lorentz Berthelot mixing rules would be applied to pre-assembled lipid-DNA mixtures. The system could be setup with the lipids forming bilayers and the DNA intercalated between the bilayers. The simulation would be run, the stability of the system checked and the process continued as for the lipids simulations (Chapter 5). While this process is expected to be time consuming it should be feasible based on the experience gained in parameterising the lipid model. Further work would be required to setup a system of non-lamellar forming lipids along with DNA, ions, solvent as well as the lamellar forming lipids to make the models realistic.

An alternative strategy would entail simply mixing the different components in the system using standard Lorentz Berthelot mixing rules and observing the structures that arise and then parameterising the interactions to capture certain benchmark properties of the mixtures. While this alternative strategy in theory should work, past experience in using such a strategy in simulating lipids means that it would be potentially more likely to fail.

7.2.4 End Summary

The model developed has proven its capability as a valuable tool in discerning the properties of the many lipid phases that it is able to reproduce. This capability is beyond the scope of atomistic models and other coarse grain models at the same level of granularity. Even in instances where several of the other phases have been reported by other coarse grain lipid models [116][117][118][126][125], they have generally simplified the representation of electrostatics and solvent properties, both of which the current model has rectified. In terms of the coarse grain DNA models, the currently available models again have the shortcomings seen in coarse grain lipid models such as the neglect or over simplification of electrostatics and solvent properties, which play a significant role in DNA [167][168][168][170][23][21][22] [19][183][176][177][178] [179]. The model presented here has been developed to solve these issues, while maintaining its overall simplicity. The developed DNA model has also been shown to reproduce the correct persistence length of DNA. Overall, the developed lipid and DNA models have been shown to be generally superior in reproducing many of the mechanical and physical properties of lipids and DNA, compared to models of the same granularity. Simulations performed using the developed lipid and DNA models have shown that these simple models have the capability to be combined to study the structures that these complex mixtures may adopt, in the context of better understanding the role of lipids within the nucleus [1][2][3][4], as well as the use of such mixtures in gene therapy [6][7][8][181][185][186][236][237][238][239][240][241][242][243].

Appendix A

Dimensionless Units

A.1 Dimensionless Reduced Units

The use of dimensionless units is undertaken in MD simulations for several reasons. One of the primary reasons being able to work with values not too distant from unity and this avoids the possible problems of rounding off errors that may result in using extremely small values associated with atomistic systems. The use of dimensionless units also allow for equations of motion to be simplified as most of the parameters defining the system are absorbed into the units. This also means that the model once developed may be used to study a whole range of problems using these units and then the units scaled to the the appropriate physical units for each problem [40].

Choosing the LJ parameters σ and ϵ and the particle mass m to be the fundamental units of length, energy and mass respectively the list of the definitions of some reduced quantities are listed below [24].

length: $L^* = L/\sigma$

energy: $E^* = E/\epsilon$

mass: $M^* = M/m$

number density: $\rho^* = \rho\sigma^3$

time: $t^* = t\sqrt{\epsilon/m\sigma^2}$

temperature: $T^* = Tk_B/\epsilon$

pressure: $P^* = P\sigma^3/\epsilon$

charge: $z^* = z/\sqrt{4\pi\varepsilon_0\sigma\epsilon}$

dipole moment: $\mu^* = \mu/\sqrt{4\pi\varepsilon_0\sigma^3\epsilon}$

force: $f^* = f\sigma/\epsilon$

torque: $\tau^* = \tau/\epsilon$

A.2 Definition of Dimensionless Units in *LANKA*

The fundamental units in *LANKA* of length, mass and energy are defined as follows.

length: $\sigma_{unit} = 3.035\mathring{A}$

mass: $m_{unit} = 18.01\ amu/2.99 \times 10^{-29}\ kg$

energy: $\epsilon_{unit} = 0.152\ kcal\ mol^{-1}$

Appendix B

Interaction Potentials, Forces and Torques

B.1 Interaction Potentials

The potential energy $U(r)$ is in principle obtainable from quantum mechanics though this intractable for all but the simplest of problems. One can show from perturbation theory, however that asymptotically it takes the form as given by B.1. Often simple analytical equations with adjustable parameters for $U(r)$ that go asymptotically as r^{-6}. The most well used form of such equations is given in B.2 where σ is the distance at which $u(r)$ is zero and ϵ is the depth of the well. The exponent n may take values from 9 to 15 though 12 is the most popular and widely used and is known as the Lennard Jones 6-12 potential is given in B.3.

$$\mathcal{U}(r) \rightarrow -C_6 r^{-6} \tag{B.1}$$

$$\mathcal{U}(r) = \frac{n\epsilon}{n-6}\left(\frac{n}{6}\right)^{6/(n-6)}\left[\left(\frac{\sigma}{r}\right)^{n} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{B.2}$$

$$\mathcal{U}(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{B.3}$$

B.2 Lennard-Jones Interaction

The Lennard Jones pair potential is commonly used to model interactions between spherical particles. It has both an attractive element and a repulsive element. The attractive part comes into play at large intermolecular separations and used to model dispersive interactions due to correlation between electron clouds surrounding the molecules. At shorter distances a negative well represents the cohesion in condensed phases and finally a sharply rising repulsive element at shorter distances corresponding to non-bonded overlap between electron clouds. The LJ pair-potential is characterised by a long-range attractive tail of the form $-1/r^6$, a negative well of depth ϵ, and a steeply rising repulsive wall at distances less than $r \sim \sigma$:

For a pair of atoms i and j located at $\mathbf{r_i}$ and $\mathbf{r_j}$ the potential energy is given by equation A.1, where $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ The parameter ϵ governs the strength of the interactions while σ defines a length scale.

$$\mathcal{U}(r_{ij}) = \begin{cases} 4\epsilon_{\mathrm{LJ}}[(\sigma_{\mathrm{LJ}}/r_{ij})^{12} - (\sigma_{\mathrm{LJ}}/r_{ij})^6] & \text{if } r_{ij} < r_c \\ 0 & \text{if } r_{ij} \geq r_c \end{cases} \tag{B.4}$$

B.2.1 Lennard-Jones Forces

The force corresponding to $u(r_{ij})$ is given by equation B.5 and equation B.6, where B.6 gives the force exerted by particle j on i. The force on j due to i, according to Newton's third law will be $\mathbf{f}_{ij} = -\mathbf{f}_{ji}$

$$\mathbf{f} = -\nabla u(r) \tag{B.5}$$

$$\mathbf{f}_{ij} = \left(\frac{48\epsilon}{\sigma^2}\right)\left[\left(\frac{\sigma}{r_{ij}}\right)^{14} - \frac{1}{2}\left(\frac{\sigma}{r_{ij}}\right)^{8}\right]\mathbf{r}_{ij} \tag{B.6}$$

B.3 Corner Potential

The Lennard Jones potential is only suitable for molecules that posses spherical symmetry and this is a major drawback of this potential energy function, to overcome this shortcoming several potential energy formulations have been proposed one of the earliest being the Corner potential. The potentials proposed not only depend on the position of the particles but also on their orientation. The importance of the Corner potential is that is a single site potential where the parameters σ and ϵ depend on both the position and orientation. In Corner's formalism, the potential takes a Lennard Jones 12-6 form as given by B.7. Where $\epsilon(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}})$ is the strength function and is taken as the well depth and is given by B.8 with $\sigma(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}_{ij}})$ is taken as the contact distance and is given in B.9.

$$\mathcal{U}(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}_{ij}}) = 4\epsilon(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}})\left[\left(\frac{\sigma(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}_{ij}})}{r_{ij}}\right)^{12} - \left(\frac{\sigma(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}_{ij}})}{r_{ij}}\right)^{6}\right] \tag{B.7}$$

$$\left(\frac{\epsilon(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}_{ij}})}{\epsilon_o}\right)^{1/2} = 1 + \left[\frac{10}{9}\left(\frac{l}{\sigma_o}\right)^2 - \frac{9}{8}\frac{l}{\sigma_o}\right] F_1 + \frac{9}{8}\frac{l}{\sigma_o} F_3 - \frac{l}{7\sigma_o} F_4 \tag{B.8}$$

$$\left(\frac{\sigma(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}_{ij}})}{\sigma_o}\right)^{1/3} = 1 - \frac{2}{5}\frac{l}{\sigma_o} + \frac{7}{2}\frac{l}{\sigma_o} F_1 + 21\left(\frac{l}{\sigma_o}\right)^3 F_2 - 14\left(\frac{l}{\sigma_o}\right) + 16\left(\frac{l}{\sigma_o}\right)^3 F_1 F_4 - 25\left(\frac{l}{\sigma_o}\right)^3 F_2 F_4 \tag{B.9}$$

The short hand notations used in B.7- B.9 are expanded as

$F1 = f_1 + f_2$,

$F2 = f_1^2 f_2^2$,

$F3 = f_0 f_1 f_2 - f_1^2 f_2^2$,

$F4 = f_0^2 - 2 f_0 f_1 f_2 + f_1^2 f_2$,

where $f_0 = \hat{\mathbf{u}}_i \cdot \hat{\mathbf{u}}_j, f_1 = \hat{\mathbf{u}}_i \cdot \hat{\mathbf{r}}$ and $f_2 = \hat{\mathbf{u}}_j \cdot \hat{\mathbf{r}}$. For site-site models the pair potential was determined as the sum of individual site-site interactions of atoms in differant molecules. Although the corner potential is a vast improvement and is computationally simple it is unable to simulate any molecules other than the simplest ones containing anisotropy. The Gay Berne potential based on the corner potential overcomes this and has been used in developing our own model.

B.4 Gay-Berne Interaction

B.4.1 Berne Pechukas Potential

The Berne Pechukas potential [41] was a forerunner to the Gay Berne potential and is essentially a symmetric Gaussian overlap model generalised to a Lennard Jones form. The potential contains the basic 12-6 part of the interactions with angular dependence σ determined by the overlap of two Gaussian overlap ellipsoidal functions.

Berne and Pechukas treated the molecules as ellipsoidal in shape with a Gaussian charge distribution to determine the well depth and contact distance. Evaluating the overlap between two such distributions gives the orientation dependence of contact distance, σ as given by B.10.

$$\sigma(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}}) = \sigma_{\mathrm{GB}} \left\{1 - \frac{\chi}{2}\left[\frac{(\hat{\mathbf{r}} \cdot \hat{\mathbf{u}}_i + \hat{\mathbf{r}} \cdot \hat{\mathbf{u}}_j)^2}{1 + \chi\, \hat{\mathbf{u}}_i \cdot \hat{\mathbf{u}}_j} + \frac{(\hat{\mathbf{u}}_i \cdot \hat{\mathbf{r}} - \hat{\mathbf{u}}_j \cdot \hat{\mathbf{r}})^2}{1 - \chi\, \hat{\mathbf{u}}_i \cdot \hat{\mathbf{u}}_j}\right]\right\}^{-\frac{1}{2}}, \tag{B.10}$$

where χ is the measure of anisotropy which is determined by the length of major and minor axes of the anisotropic particle and is given by 2.10.

$$\chi = \frac{(\sigma_{\|}/\sigma_{\perp})^2 - 1}{(\sigma_{\|}/\sigma_{\perp})^2 + 1} = \frac{\kappa^2 - 1}{\kappa^2 + 1}, \tag{B.11}$$

Where $\sigma_{\|}$ and $\sigma_{\perp}$ are proportional to the length and cross section diameter of the ellipsoidal particle.The energy parameter is also dependant upon χ and is given by B.12.

$$\epsilon(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j) = \epsilon_o[1 - \chi^2(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j)^2]^{-\frac{1}{2}} \tag{B.12}$$

The parameters ϵ_0 and σ_0 are the well depth and contact distance for the cross configuration and $\hat{\mathbf{u}}_i$ and $\hat{\mathbf{u}}_j$ are unit vectors describing the orientations of the molecule. The resultant Berne Pechukas potential is given by B.13

$$\mathcal{U}(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}_{ij}}) = 4\epsilon(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j)\left[\left(\frac{\sigma(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}_{ij}})}{r_{ij}}\right)^{12} - \left(\frac{\sigma(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}_{ij}})}{r_{ij}}\right)^{6}\right] \tag{B.13}$$

The Berne-Pechukas potential was used by Kushick and Berne [244] to simulate a model mesogen, however, it was found to suffer from several unrealistic features. Firstly the strength parameter does not depend on the intermolecular separation vector $\mathbf{r}_{ij}$ resulting in equal interaction potential for both side-by-side and end-to-end placements, while it would be expected to be stronger for side-by-side configuration. The other undesirable feature was that the well width depended on the intermolecular separation $\mathbf{r}_{ij}$ and was larger for end-to-end than side-by-side placements where as they would have been expected to be the same. The two unrealistic features mentioned were investigated and rectified by the Gay-Berne potential [25].

B.4.2 Gay Berne Potential

To rectify the anomalous features, Gay and Berne proposed the following modification, the energy parameter was modified to be explicitly dependant upon $\mathbf{r}_{ij}$ and given by B.14, which are a generalisation of the changes proposed by Tsykalo and Bagmet [245]. Where $\epsilon_1(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j)$ is the strength parameter given by the Berne-Pechukas potential and $\epsilon_2(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}})$ is a function taking the same form as the Berne-Pechukas $\sigma(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}_{ij}})$ term and given by B.15 . The term χ' represents an interaction strength ratio and is given by B.16. The exponents ν and μ are adjustable and were originally used to obtain a good agreement to an array of four Lennard Jones spheres. The values used in $LANKA$ are μ=2 and ν=1.

$$\epsilon(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}}) = \epsilon_0\epsilon_1^{\nu}(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j)\epsilon_2^{\mu}(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}}) \tag{B.14}$$

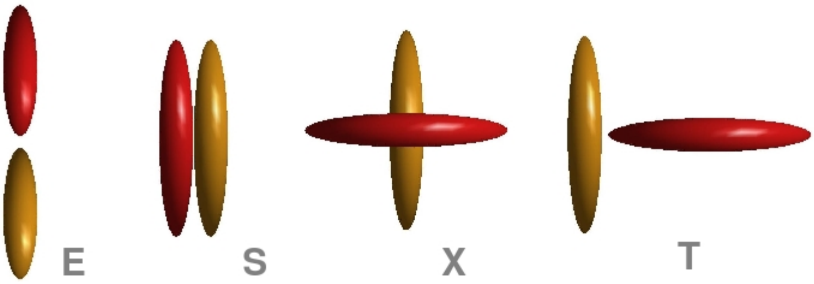

Figure B.1: The primary four configurations of GB ellipsoids

$$\epsilon_2(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}}) = 1 - \frac{\chi'}{2}\left[\frac{(\hat{\mathbf{u}}_i \cdot \hat{\mathbf{r}} + \hat{\mathbf{u}}_j \cdot \hat{\mathbf{r}})^2}{1 + \chi'\,\hat{\mathbf{u}}_i \cdot \hat{\mathbf{u}}_j} + \frac{(\hat{\mathbf{u}}_i \cdot \hat{\mathbf{r}} - \hat{\mathbf{u}}_j \cdot \hat{\mathbf{r}})^2}{1 - \chi'\,\hat{\mathbf{u}}_i \cdot \hat{\mathbf{u}}_j}\right] . \tag{B.15}$$

$$\chi' = \frac{(\epsilon_\perp/\epsilon_\parallel)^{\frac{1}{\mu}} - 1}{(\epsilon_\perp/\epsilon_\parallel)^{\frac{1}{\mu}} + 1} = \frac{\kappa'^{1/\mu} - 1}{\kappa'^{1/\mu} + 1} , \tag{B.16}$$

The second modification involves a few functional form of the exponent such that the potential is displaced rather than dilated as first suggested by Stone [246]. The resulting Gay-Berne potential is expressed by B.17 where $\sigma(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}})$ is defined as for the Berne-Pechukas model.

$$\mathcal{U}(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \mathbf{r}) = 4\epsilon(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}}) \left[\left(\frac{\sigma_{\mathrm{GB}}}{r - \sigma(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}}) + \sigma_{\mathrm{GB}}}\right)^{12} - \left(\frac{\sigma_{\mathrm{GB}}}{r - \sigma(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{r}}) + \sigma_{\mathrm{GB}}}\right)^{6}\right] \tag{B.17}$$

The $\kappa^{'}$ defined in B.16 gives the ratio of potential energy minimum for a particles placed side-by-side and end-to-end. The values of $\kappa, \kappa^{'}, \mu$ and ν are the adjustable parameters and are used to obtain a range of mesogenic structures using the Gay Berne potential. Depending on the orientation of the ellipsoids the interaction strength varies, the primary four arrangements are shown in Figure B.4.2 with the relevant interparticle separation and well depths in Table B.1 [26].

B.4.3 Gay Berne Forces and Torques

For an interaction between two sites via the Gay Berne potential the force on site i at (x_i, y_i, z_i) is given by Equation B.4.3.

$$\mathbf{f}_i = -\begin{pmatrix} \partial\mathcal{U}/\partial x_i \\ \partial\mathcal{U}/\partial y_i \\ \partial\mathcal{U}/\partial z_i \end{pmatrix}$$

Configuration	$\sigma(\hat{u}_i, \hat{u}_j, \hat{r})$	$\epsilon(\hat{u}_i, \hat{u}_j, \hat{r})$
E	$\sigma_0(\sigma_\parallel/\sigma_\perp)\ (=\sigma_\parallel)$	$\epsilon_0(\epsilon_\parallel/\epsilon_\perp)(1-\chi^2)^{\nu/2}$
S	$\sigma_0\ (=\sigma_\perp)$	$\epsilon_0(1-\chi^2)^{\nu/2}$
X	$\sigma_0\ (=\sigma_\perp)$	ϵ_0
T	$\sigma_0\sqrt{((\sigma_\parallel/\sigma_\perp)^2+1)/2}$	$\epsilon_0[2/[(\epsilon_\perp/\epsilon_\parallel)^{1/\mu}+1]]^\mu$

Table B.1: Interparticle separation $\sigma(\hat{u}_i, \hat{u}_j, \hat{r})$ and well depth $\epsilon(\hat{u}_i, \hat{u}_j, \hat{r})$ for end-to-end **E**, side-by-side **S**, cross **X** and t-shaped **T** arrangements.

The Gay Berne particle also experiences a torque to the angular dependence of the of the potential. The torque experienced by a Gay-Berne particle is equivalent to a force acting on a point separated by a unit distance from the centre of mass and acting in a direction orthogonal to the molecular symmetry axis. This equivalent force can be defined in terms of the derivative of the potential with respect to the coordinates of this point where the centre of mass is taken as the origin. These coordinates are just the components of the unit vector $\hat{\mathbf{u}}_1$ describing the molecular orientation and so the equivalent force is given by Equation B.4.3.

$$\mathbf{e} = -\begin{pmatrix} \partial\mathcal{U}/\partial\hat{u}_{1x} \\ \partial\mathcal{U}/\partial\hat{u}_{1y} \\ \partial\mathcal{U}/\partial\hat{u}_{1z} \end{pmatrix}$$

Details of deriving explicit formula for the forces and torques are presented are given by Luckhurst et al. [26].

B.5 Generalised Gay Berne Potential

For capturing the interactions between unlike ellipsoidal particles. First developed by Clever et al. and extended by Ren et al. [42][43]. The potential energy terms are listed in Chapter 2, the relevant equations for the forces and torques are given in the work by Ren et al. [43].

B.6 Dipole-Dipole Interactions

The configurations of a pair of molecules is shown in B.6 [24]. The centre's of the molecule are separated by the vector $\mathbf{r}_{ij}$. θ_i and θ_j are the angles between $\mathbf{r}_{ij}$ and the unit vectors directed along the molecular axes, $\mathbf{e}_i$ and $\mathbf{e}_j$. ϕ_{ij} is the angle between the plane containing

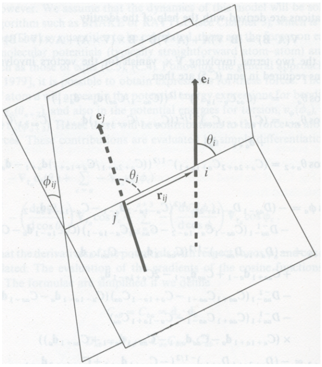

Figure B.2: The relative orientation of two linear rigid bodies [24, Figure C.2].

$\mathbf{e}_i$ and $\mathbf{r}_{ij}$. The angle γ_{ij} is defined as given in equation B.18. If the molecules i and j have a dipole moment of μ, then the interaction potential is given by equation B.19

$$\cos\gamma_{ij} = \mathbf{e}_i \cdot \mathbf{e}_i = \cos\theta_i \cos\theta_j + \sin\theta_i \sin\theta_j \cos\phi_{ij} \tag{B.18}$$

$$u_{ij}^{\mu\mu} = \frac{\mu^2}{r^3}(\cos\gamma_{ij} - 3\cos\theta_i\cos\theta_j) \tag{B.19}$$

The forces on i and j are given by equation B.20. If the dipole is found in the centre of mass from which the forces are computed then the torques on particles i and j are given by equations B.21 and B.22 respectively[24] pages [332-334].

However if as in the lipid model the dipoles are displaced along the ellipsoid and the force is computed from the centre of mass of the ellipsoid then the torques are computed as given below. The torque on molecule i due to molecule j is defined by equation B.23 with explicit formulae for the torques give by Equations B.24 and B.25. The equations for torques in molecules where the dipoles are displaced from the centre of mass were derived from work done by Price et al.[187]. The vectors $\mathbf{a}$ and $\mathbf{b}$ represent the displacement of the dipole from the centre of mass of the molecule. The orientation of the dipoles is as above given by $\mathbf{e}_i$ and $\mathbf{e}_j$.

$$\mathbf{f}_{ij} = -\mathbf{f}_{ji} = \frac{3\mu^2}{r^4}\left[(\cos\gamma_{ij} - 5\cos\theta_i\cos\theta_j)(\mathbf{r}_{ij}/r + \cos\theta_j\mathbf{e}_i + \cos\theta_i\mathbf{e}_j)\right] \tag{B.20}$$

$$\mathbf{T}_{ij} = -\frac{\mu^2}{r^3}\left[\mathbf{e}_i \times \mathbf{e}_j - 3\cos\theta_j(\mathbf{e}_i \times \mathbf{r}_{ij})/r\right] \tag{B.21}$$

$$\mathbf{T}_{ji} = -\frac{\mu^2}{r^3}\left[\mathbf{e}_j \times \mathbf{e}_i - 3\cos\theta_i(\mathbf{e}_j \times \mathbf{r}_{ij})/r\right] \tag{B.22}$$

$$\mathbf{T}_{ij} = -\mathbf{e}_i \times \nabla_{\mathbf{e}_i} u_{ij}^{\mu\mu} \tag{B.23}$$

$$\mathbf{T}_{ij} = \frac{\mu^2}{r^3}\left[-\frac{3}{r^2}\cos\gamma_{ij}(\mathbf{r}_{ij}\times\mathbf{a}) + \mathbf{e}_j\times\mathbf{e}_i + \frac{15}{r^2}\cos\theta_i\cos\theta_j(\mathbf{r}_{ij}\times\mathbf{a}) - \frac{3}{r}_{ij}\cos\theta_j\left[(\mathbf{r}_{ij}-\mathbf{a})\times\mathbf{e}_i\right] - \frac{3}{r}_{ij}\cos \right. \tag{B.24}$$

$$\mathbf{T}_{ji} = \frac{\mu^2}{r^3}\left[\frac{3}{r^2}\cos\gamma_{ij}(\mathbf{r}_{ij}\times\mathbf{b}) + \mathbf{e}_i\times\mathbf{e}_j - \frac{15}{r^2}\cos\theta_i\cos\theta_j(\mathbf{r}_{ij}\times\mathbf{b}) + \frac{3}{r}_{ij}\cos\theta_i\left[(\mathbf{r}_{ij}+\mathbf{b})\times\mathbf{e}_j\right] + \frac{3}{r}_{ij}\cos\theta_j \right. \tag{B.25}$$

B.7 Monopole-Monopole Interaction

Monopole-Monopole Potential

Monopole-monopole interactions are computed using the Coulomb's law and is given by Equation B.26.

$$u_{ij}(r) = \frac{Q_i Q_j}{4\pi\epsilon_0 r} \tag{B.26}$$

Monopole-Monopole Forces

$$\mathbf{f}_{ij}(r) = \frac{Q_i Q_j}{4\pi\epsilon_0 r^3}\mathbf{r} \tag{B.27}$$

with $\mathbf{r} = \mathbf{r}_i - \mathbf{r}_j$.

B.8 Monopole-Dipole Interaction

Interaction between a dipole μ with associated orientation vector $\mathbf{e}_\mu$ and a charge Q, assuming site j to be a point-dipole and site i to collapse to a point-charge.

Monopole-Dipole Potential

Considering site j a dipole μ, with corresponding orientation vector $\mathbf{e}_\mu$, and site i a charge Q, the electrostatic interaction energy is:

$$u^{Q\mu} = \frac{Q\mu}{4\pi\epsilon_0 r^3}\left(\mathbf{e}_\mu\cdot\mathbf{r}\right) = \frac{Q\mu}{4\pi\epsilon_0 r^2}\,\cos\theta_j \tag{B.28}$$

with $\mathbf{r} = \mathbf{r}_i - \mathbf{r}_j = \mathbf{r}_Q - \mathbf{r}_\mu$.

Monopole-Dipole Force

The force between a charge and a dipole is given by the Equation B.29.

$$\mathbf{f}_{ij} = -\mathbf{f}_{ji} = \frac{Q\mu}{4\pi\epsilon_0 r^5}\left[3(\hat{\mathbf{e}}_\mu \mathbf{r})\mathbf{r} - r^2\hat{\mathbf{e}}_\mu\right] \tag{B.29}$$

with $\mathbf{r} = \mathbf{r}_i - \mathbf{r}_j = \mathbf{r}_Q - \mathbf{r}_\mu$.

Monopole-Dipole Torque

If the dipole is in in the centre of mass then pair torque on site j:

$$\mathbf{T}_{ji} = \frac{Q\mu}{4\pi\epsilon_0 r^3}\left(\mathbf{r} \times \mathbf{e}_j\right) = \mathbf{T}_{\mu Q} = \frac{Q\mu}{4\pi\epsilon_0 r^3}(\mathbf{r} \times \mathbf{e}_\mu) \tag{B.30}$$

with $\mathbf{r} = \mathbf{r}_i - \mathbf{r}_j = \mathbf{r}_Q - \mathbf{r}_\mu$. If the dipole is displaced, then it is computed as in the displaced dipole-dipole interactions above.

B.9 Computing Torques

In addition to the torques mentioned above , additional torques may arise if the force is acting on the molecule is at a point away form its centre of mass. Then these additional torques maybe computed using the standard formula as given by Equation B.31 where $\mathbf{r}$ is the displacement vector from the centre of mass and $\mathbf{F}$ the force acting on that point.

$$\mathbf{T} = \mathbf{r} \times \mathbf{F} \tag{B.31}$$

References

[1] A.N. Hunt, G.T. Clark, G.S. Attard, and A.D. Postle. Highly saturated endonuclear phosphatidylcholine is synthesized in situ and colocated with CDP-choline pathway enzymes. *J.Biol.Chem*, 276:8492–8499, 2001.

[2] A.N. Hunt. Dynamic lipidomics of the nucleus. *J. Cell. Biol. Chem.*, 97:244–251, 2006.

[3] A.N. Hunt and A.D. Postle. Mass spectrometry determination of endonuclear phospholipid composition and dynamics. *Methods*, 39:104–111, 2006.

[4] S.M. Mat, R. R. Brenner, and A. Ves-Losada. Endonuclear lipids in liver cells. *Can.J.Physiol.Pharmacol.*, 84:459–468, 2006.

[5] T. Schlick. *Molecular Modeling and Simulation-An Interdisciplinary Guide.* Springer, 2002.

[6] S. May and A. Ben-Shaul. DNA-Lipid complexes: stability of honeycomb-like and spaghetti-like structures. *Biophys J*, 73:2427–2440, 1997.

[7] D. Harris, S. May, W. M. Gelbart, and A. Ben-Shaul. Structure, stabilty and thermodynamics of lamellar DNA-Lipid complexes. *Biophys J*, 75:159–173, 1998.

[8] S. May, D. Harris, and A. Ben-Shaul. The phase behaviour of cationic lipid-DNA complexes. *Biophys J*, 78:1681–1697, 2000.

[9] S. Khalid, PJ. Bond, J. Holyoake, R.W. Hawtin, and M.S.P. Sansom. DNA and lipid bilayers : self-assembly and insertion. *Interface*, 5:S241–S250, 2008.

[10] J. M. Berg, J. L. Tymoczko, and L. Stryer. *Biochemistry.* W. H. Freeman and Company, 2003.

[11] J.D. Watson and F.C.H. Crick. Molecular structure of nucelic acids - a structure for deoxyribose nucleic acids. *Nature*, 171:737–738, 1953.

[12] T Schlick, D.A. Beard, J. Huang, D. A. Strahs, and X Qian. Computational challenges in simulating large DNA over long times. *Computing in Science and Engineering*, 02:38–51, 2000.

[13] S. Louise-May, P. Auffinger, and E. Westhof. Calculations of nucleic acid conformations. *Curr Opin Struct Biol*, 06:289–298, 1996.

[14] M. Venturoli, M.M. Sperottob, M. Kranenburgc, and B. Smit. Mesoscopic models of biological membranes. *Physics Reports*, 437:1–54, 2006.

[15] M. Muller, K. Katsov, and M. Schick. Biological and synthetic membranes: What can be learned from a coarse-grained description? *Physics Reports*, 434:113–176, 2006.

[16] G. Brannigan, L.C.L. Lin, and F.L.H. Brown. Implicit solvent simulation models for biomembranes. *Euro. Biophys. J.*, 35:104–124, 2006.

[17] S. J. Marrink, E. Lindahl, O. Edholm, and A. E. Mark. Simulation of the spontaneous aggregation of phospholipids into bilayers. *J Am Chem Soc*, 123:8638–8639, 2001.

[18] T. E. Cheatham III and M.A. Young. Molecular dymanics simulation of nucleic acids: Successes, limitaions, and promise. *Biopolymers*, 56:232–256, 2001.

[19] TA. Knotts IV, N. Rathore, D.C. Schwartz, and J. J. de Pablo. A coarse grain model for DNA. *J. Chem. Phys.*, 126:084901–1–084901–12, 2007.

[20] S. J. Marrink, A. H. Vries, and A. E. Mark. Coarse grained model for semiquantitative lipid simulations. *J Phys Chem B*, 108:750–760, 2004.

[21] B. Mergell, M. R. Ejtehadi, and R. Everaers. Modeling DNA structure, elasticity, and deformations at the base-pair level. *Phys Rev E*, 68:art, 2003.

[22] H.L. Tepper and G.A. Voth. A coarse-grained model for double-helix molecules in solution: Spontaneous helix formation and equilbrium properties. *J Chem Phys*, 122:art, 2005.

[23] K. Drukker, G. Wu, and G. C. Schatz. Model simulations of DNA denaturation dynamics. *J Chem Phys*, 114:579–590, 2001.

[24] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids.* Oxford Science Publications, 1987.

[25] J. G. Gay and B. J. Berne. Modification of the overlap potential to mimic a linear site-site potential. *J Chem Phys*, 74:3316–3319, 1981.

[26] G. R. Luckhurst, R. A. Stephens, and R. W. Phippen. Computer-simulation studies of anisotropic systems. XIX. mesophases formed by the Gay-Berne model mesogen. *Liq Cryst*, 8:451–464, 1990.

[27] G. R. Luckhurst and P. S. J. Simmonds. Computer-simulation studies of anisotropic systems .XXI. parametrization of the Gay-Berne potential for model mesogens. *Mol Phys*, 80:233–252, 1993.

[28] A. P. J. Emerson, G.R. Luckhurst, and S.G. Whatling. Computer simulation studies of anisotropic systems XXIII. the Gay-Berne discogen. *Mol. Phys.*, 82:113–124, 1994.

[29] R. Hashim, G.R. Luckhurst, and S. Romano. Computer-simulation studies of anisotropic systems part XXIV- constant-pressure investigations of the Smectic B phase of the Gay-Berne mesogen. *J.Chem.Soc.Faraday Trans*, 91:2141–2148, 1995.

[30] M.A. Bates and G.R. Luckhurst. Computer simulation studies of anisotropic systems XXVI. Monte Carlo investigations of a Gay-Berne discotic at constant pressure. *J.Chem.Phys.*, 104:6696–6709, 1996.

[31] M.A. Bates and G.R. Luckhurst. Computer simulation studies of anisotropic systems part xxx. the phase behaviour and structure of a gay-berne mesogen. *J.Chem.Phys.*, 110:7087–7108, 1999.

[32] M. A. Bates and G.R. Luckhurst. Studies of translational diffusion in the Smectic A phase of a Gay-Berne mesogen using molecular dynamics computer simulation. *J Chem Phys*, 120:394–403, 2004.

[33] J.N. Israelachvili, D.J. Mitchell, and B.W. Ninham. Theory of self-assembly of hydrocarbon amphiphiles into micelles and bilayers. *J. Chem. Soc., Faraday Trans. 2*, 72:1525 – 1568, 1976.

[34] D.J. Mitchell and B.W. Ninham. Micelles, vesicles and microemulsions. *J. Chem. Soc., Faraday Trans. 2*, 72:601–629, 1981.

[35] D.F. Evans and B.W. Ninham. Molecular forces in the self-organization of amphiphiles. *J. Phys.Chem*, 90:226–234, 1986.

[36] O. G. Mouritsen. *Life - As a Matter of Fat. The Emerging Science of Lipidomics.* Springer, 2005.

[37] C.R. Calladine and H.R. Drew. *Understanding DNA the molecule and how it works.* Academic Press, San Diego, 1997.

[38] J. M. Haile. *Molecular Dynamics Simulation Elementary Methods.* John Wiley and Sons, 1992.

[39] A. R. Leach. *Molecular Modelling: Principles and Applications.* Prentice Hall, 2001.

[40] D. C. Rapaport. *The Art of Molecular Dynamics Simulation.* Cambridge University Press, 2004.

[41] B. J. Berne and P. Pechukas. Gaussian model potentials for molecular interactions. *J Chem Phys*, 56:4213–4216, 1972.

[42] D. J. Cleaver, C. M. Care, M. P. Allen, and M. P. Neal. Extension and generalization of the Gay-Berne potential. *Phys Rev E*, 54:559–567, 1996.

[43] P. A. Golubkov and P. Ren. Generalized coarse-grained model based on point multipole and gay-berne potentials. *J Chem Phys*, 125:064103, 2006.

[44] J. P. Ryckaert and A. Bellemans. Molecular dynamics of liquid alkanes. *Chem. Soc. Faraday. Discuss*, 66:95–106, 1978.

[45] H. Goldstein, C. Poole, and J. Safko. *Classical Mechanics - third edition.* Pearson, 2002.

[46] D.A. McQuarrie. *Statistical Mechanics.* Harper and Row, 1973.

[47] A. Dullweber, B. Leimkuhler, and R. McLachlan. Symplectic splitting methods for rigid body molecular dynamics. *J Chem Phys*, 107:5840–5851, 1997.

[48] D. Fincham. An algorithm for the rotational motion of rigid molecules. *CCP5 Quarterly*, 2:6–10, 1981.

[49] C. J. Fennell and J. D. Gezelter. On the structural and transport properties of the soft sticky dipole and related single-point water models. *J Chem Phys*, 120:9175–9184, 2004.

[50] X. Sun and J. D. Gezelter. Dipolar ordering in the ripple phases of molecular-scale models of lipid membranes. *J. Phys. Chem*, 112:1968–1975, 2008.

[51] M. Orsi, D.Y. Haubertin, W.E. Sanderson, and J. W. Essex. A quantitative coarse-grain model for lipid bilayers. *J. Phys. Chem. B.*, 112:802–815, 2007.

[52] D. J. Adams, E.M. Adams, and G.J. Hills. The computer simulation of polar liquids. *Mol. Phys*, 38:387–400, 1979.

[53] M. Neumann. Dipole moment fluctuation formulas in computer simulations of polar systems. *Mol. Phys.*, 50:841–858, 1983.

[54] Y. Liu and T. Ichiye. The static dielectric constant of the soft sticky dipole model of liquid water: Monte carlo simulation. *Chem Phys Lett*, 256:334–340, 1996.

[55] J.W. Essex. The application of the reaction-field method to the calculation of dielectric constants. *Mol. Sim*, 20:159–178, 1998.

[56] R.O. Watts. Monte Carlo studies of liquid water. *Mol. Phys.*, 28:1069–1083, 1974.

[57] R.O. Watts. Electric polarisation of water: Monte Carlo studies. *Chem. Phys.*, 57:185–195, 1981.

[58] H.E. Alper and R.M. Levy. Dielectric and thermodynamic response of a generalized reaction field model for liquid state simulations. *J. Chem Phys.*, 99:9847–9852, 1993.

[59] P. E. Smith and W.F van Gunsteren. Consistent dielectric properties of the simple point charge and extended simple point charge water models at 277 and 300 K. *J. Chem Phys.*, 100:3169–3174, 1994.

[60] M.R. Reddy and M. Berkowitz. The dielectric constant of SPC/E water. *Chem Phys. Lett*, 155:173–176, 1989.

[61] O. Steinhauser. Reaction field simulation of water. *Mol. Phys.*, 45:335–348, 1982.

[62] M. Neumann. Dielectric relaxation in water. computer simulations with the TIP4P potential. *J. Chem Phys.*, 85:1567–1580, 1986.

[63] M. Neumann. The dielectric constant of water. computer simulations with the MCY potential. *J. Chem Phys.*, 82:5663–5672, 1985.

[64] S.B. Zhu and C.F. Wong. Sensitivity analysis of water thermodynamics. *J. Chem Phys.*, 98:8892–8899, 1993.

[65] M. Levitt, M. Hirshberg, R. Sharon, and V. Daggett. Potential-energy function and parameters for simulations of the molecular-dynamics of proteins and nucleic-acids in solution. *Comput Phys Commun*, 91:215–231, 1995.

[66] A. Chandra and T. Ichiye. Dynamical properties of the soft sticky dipole model of water: Molecular dynamics simulations. *J Chem Phys*, 111:2701–2709, 1999.

[67] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. Dinola, and J.R. Haak. Molecular-dynamics with coupling to an external bath. *J Chem Phys*, 81:3684–3690, 1984.

[68] J. W. Essex, M. M. Hann, and W. G. Richards. Molecular dynamics simulation of a hydrated phospholipid bilayer. *Phil Trans R Soc Lond B*, 344:239–260, 1994.

[69] Phospholipids chemical structure. `http://tonga.usip.edu/gmoyna/biochem341/lecture28.ht`

[70] G. Cevc and D. Marsh. *Phospholipid Bilayers: Physical Principles and Models.* Wiley-Interscience, 1987.

[71] J.M. Seddon and R.H. Templer. *Handbook of Biological Physics.* Elsevier, 1995.

[72] P.F.F. Almeida and W.L.C. Vaz. *Lateral Diffusion in Membranes.* Elsevier, 1995.

[73] D.A Brown and E. London. Functions of lipid rafts in biological membranes. *Annu. Rev. Cell Dev. Biol.*, 14:111–136, 1998.

[74] L.J. Pike. Lipid rafts: bringing order to chaos. *J. Lipid Res.*, 44:655–667, 2003.

[75] I. Marcotte and M. Auger. Bicelles as model membranes for solidand solution-state nmr studies of membrane peptides and proteins. *Concepts Magn. Reson.*, 24A:17–37, 2005.

[76] J.A Whiles, R. Deems, R.R Vold, and E.A. Dennis. Bicelles in structurefunction studies of membrane-associated proteins. *Bioorg. Chem.*, 30:431–442, 2002.

[77] M. Egli. DNA-Cation Interactions Quo Vadis? *Chem.Biol.*, 9:277–286, 2002.

[78] G.S. Manning. Limiting laws and counterion condensation in polyelectrolyte solutions I. colligative properties. *J.Chem.Phys.*, 51:924–933, 1969.

[79] G.S. Manning. Limiting laws and counterion condensation in polyelectrolyte solutions II. self-diffusion of the small ions. *J.Chem.Phys.*, 51:934–938, 1969.

[80] M. Rueda, E. Cubero, C.A. Laughton, and M. Orozco. Exploring the counterion atmosphere around DNA: What can be learned from molecular dynamics simulations? *Biophysical J*, 87:800–811, 2004.

[81] N.V. Hud and M. Polak. DNA-cation interactions: The major and minor grooves are flexible ionophores. *Curr Opin Struct Biol.*, 11:293–301, 2001.

[82] S.Y. Ponomarev, K.M. Thayer, and D.L. Beveridge. Ion motions in molecular dynamics simulations on DNA. *Proc Natl Acad Sci*, 101:14771–14775, 2004.

[83] P. Vrnai and K. Zakrzewska. DNA and its counterions: a molecular dynamics study. *Nucleic Acids Res*, 32:4269–4280, 2004.

[84] V.P. Denisov and B. Halle. Sequence-specific binding of counterions to B-DNA. *Proc. Natl. Acad. Sci.*, 97:629–633, 2000.

[85] M.A. Young, B. Jayaram, and D.L. Beveridge. Intrusion of counterions into the spine of hydration in the minor groove of B-DNA: Fractional occupancy of electronegative pockets. *J.Am.Chem.Soc.*, 119:59–69, 1997.

[86] H.R. Drew, R.M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura, and R.E. Dickerson. Structure of a B-DNA dodecamer: conformation and dynamics. *Proc.Natl.Acad.Sci.*, 78:2179–2183, 1981.

[87] N.C. Stellwagen, S. Magnusdottir, C. Gelfi, and P.G. Righetti. Preferential counterion binding to A-tract DNA oligomers. *J. Mol. Biol.*, 305:1025–1033, 2001.

[88] D. Hamelberg, L. McFail-Isom, L.D. Williams, and W. D. Wilson. Flexible structure of DNA: Ion dependence of minor-groove structure and dynamics. *J. Am.Chem.Soc*, 122:10513–10520, 2000.

[89] B.M Feig, M. Pettitt. Sodium and chlorine ions as part of the DNA solvation shell. *Biophys. J*, 77:1769–1781, 1999.

[90] K.J. McConnell and D.L. Beveridge. DNA structure: What's in charge? *J.Mol.Biol.*, 304:803–820, 2000.

[91] C.C. Sines, L. McFail-Isom, S.B. Howerton, D. VanDerveer, and L.D. Williams. Cations mediate B-DNA conformational hetrogeneity. *J. Am.Chem.Soc*, 122:10048–11056, 2000.

[92] T.K. Chiu and R.E. Dickerson. 1 crystal structures of B-DNA reveal sequence-specific binding and groove-specific bending of DNA by magnesium and calcium. *J.Mol.Biol.*, 301:915–945, 2000.

[93] I. Rouzina and V. Bloomfield. DNA bending by small, mobile multivalent cations. *Biophys. J.*, 74:3152–3164, 1998.

[94] B. Schneider, K. Patel, and H.M. Berman. Hydration of the Phosphate Group in Double-Helical DNA. *Biophys. J*, 75:2422–2434, 1998.

[95] B. Schneidera, D.M. Cohena, L. Schleifera, A.R. Srinivasana, W.K. Olsona, and H.M. Berman. A systematic method for studying the spatial distribution of water molecules around nucleic acid bases. *Biophys. J*, 65:2291–2303, 1993.

[96] P.J. Hagerman. Fliexibilty of DNA. *Ann. Rev.Biophys.Biophys. Chem.*, 17:265–286, 1988.

[97] Y. Lu, B. Weers, and N.C. Stellwagen. DNA persistence length revisited. *Biopolymers.*, 61:261–275, 2002.

[98] M. J. Stevens. Coarse-grained simulations of lipid bilayers. *J Chem Phys*, 121:11942–11948, 2004.

[99] S. O. Nielsen, C. F. Lopez, G. Srinivas, and M. L. Klein. Coarse grain models and the computer simulation of soft materials. *J Phys-condens Mat*, 16:R481–R512, 2004.

[100] S. Izvekov, A. Violi, and G.A. Voth. Systematic coarse-graining of nanoparticle interactions in molecular dynamics simulation. *J Phys Chem B*, 109:17019–17024, 2005.

[101] A. P. Lyubartsev. Multiscale modeling of lipids and lipid bilayers. *Eur Biophys J*, 35:53–61, 2005.

[102] J. C. Shelley, M. Y. Shelley, R. C. Reeder, S. Bandyopadhyay, and M. L. Klein. A coarse grain model for phospholipid simulations. *J Phys Chem B*, 105:4464–4470, 2001.

[103] D. P. Tieleman, S. J. Marrink, and H. J. C. Berendsen. A computer perspective of membranes: molecular dynamics studies of lipid bilayer systems. *Biochim Biophys Acta*, 1331:235–270, 1997.

[104] S.E. Feller, R.M. Venable, and R.W. Pastor. Computer simulation of a DPPC phospholipid bilayer: Structural changes as a function of molecular surface area. *Langmuir*, 13:6555–6561, 1997.

[105] H. L. Scott. Modeling the lipid component of membranes. *Curr Opin Struc Biol*, 12:495–502, 2002.

[106] P. B. Moore, C. F. Lopez, and M. L. Klein. Dynamical properties of a hydrated lipid bilayer from a multinanosecond molecular dynamics simulation. *Biophys. J.*, 81:2484–2494, 2001.

[107] J.B. Klauda, B. R. Brooks, and R. W. Pastor. Dynamical motions of lipids and finite size effect in simulations of bilayers. *J. Chem. Phys.*, 125:144710, 2006.

[108] D. Bemporad, C. Luttmann, and J. W. Essex. Computer simulation of small molecule permeation across a lipid bilayer: Dependence on bilayer properties and solute volume, size, and cross-sectional area. *Biophys J*, 87:1–13, 2004.

[109] E. Lindahl and O. Edholm. Spatial and energetic-entropic decomposition of surface tension in lipid bilayers from molecular dynamics simulations. *J Chem Phys*, 113:3882–3893, 2000.

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

[111] G. Brannigan, L. C. L. Lin, and F. L. H. Brown. Implicit solvent simulation models for biomembranes. *Eur Biophys J Biophy*, 35:104–124, 2006.

[112] S. Izvekov and G.A. Voth. A multiscale coarse-graining method for biomolecular systems. *J Phys Chem B*, 109:2469–2473, 2005.

[113] M. Orsi, W.E. Sanderson, and J. W. Essex. *Molecular Interactions - Bringing Chemistry to Life.* Beilstein-Institut, Frankfurt, 2007.

[114] O. Farago. Water-free computer model for fluid bilayer membranes. *J Chem Phys*, 119:596–605, 2003.

[115] J.M. Drouffe, A.C. Maggs, and S. Leibler. Computer simulations of self-assembled membranes. *Science*, 254:1353–1356, 1991.

[116] G. Brannigan and F. L. H. Brown. Solvent-free simulations of fluid membrane bilayers. *J Chem Phys*, 120:1059–1071, 2004.

[117] G. Brannigan and F. L. H. Brown. Composition dependence of bilayer elasticity. *J Chem Phys*, 122:art., 2005.

[118] G. Brannigan, A. C. Tamboli, and F. L. H. Brown. The role of molecular shape in bilayer elasticity and phase behavior. *J Chem Phys*, 121:3259–3271, 2004.

[119] H. Noguchi and M. Takasu. Self-assembly of amphiphiles into vesicles:a brownian dynamics simulation. *Phys. Rev. E*, 64:041913, 2001.

[120] H. Noguchi and M. Takasu. Fusion pathways of vesicles:a brownian dynamics simulation. *J. Chem. Phys*, 115:9547–9551, 2001.

[121] H. Noguchi and M. Takasu. Structural changes of pulled vesicles:a brownian dynamics simulation. *Phys. Rev. E*, 65:051907, 2002.

[122] H. Noguchi and M. Takasu. Adhesion of nanoparticles to vesicles: A brownian dynamics simulation. *Biophys. J.*, 83:299–308, 2002.

[123] Z.J. Wang and D. Frenkel. Modeling flexible amphiphilic bilayers: A solvent-free off lattice Monte Carlo study. *J. Chem. Phys.*, 122:234711, 2005.

[124] Z.J. Wang and D. Frenkel. Pore nucleation in mechanically stretched bilayer membranes. *J. Chem. Phys.*, 123:154701, 2005.

[125] I.R. Cooke, K. Kremer, and M. Deserno. Tunable generic model for fluid bilayer membranes. *Phys. Rev. E.*, 72:011506, 2005.

[126] I.R. Cooke and M. Deserno. Solvent-free model for self-assembling fluid bilayer membranes: Stabilization of the fluid phase based on broad attractive tail potentials. *J. Chem. Phys.*, 123:224710, 2005.

[127] I.R. Cooke and M. Deserno. Coupling between lipid shape and membrane curvature. *Biophys. J*, 91:487–495, 2006.

[128] B.J. Reynwar, G. Illya, V.A. Harmandaris, M. M. Mller, K. Kremer, and M Deserno. Aggregation and vesiculation of membrane proteins by curvature-mediated interactions. *Nature*, 447:461–464, 2007.

[129] O. Farago and P Pincus. Statistical mechanics of bilayer membrane with a fixed projected area. *J Chem Phys*, 120:2934–2950, 2004.

[130] O. Farago, N. Grønbech-Jensen, and P. Pincus. Mesoscale computer modelling of lipid-DNA complexes for gene therapy. *Phys Rev Lett*, 96:018102, 2006.

[131] O. Farago and N. Grønbech-Jensen. Computational and analytical modeling of cationic lipid-DNA complexes. *Biophys J*, 92:3228–3240, 2007.

[132] R. Goetz and R. Lipowsky. Computer simulations of bilayer membranes: Self-assembly and interfacial tension. *J Chem Phys*, 108:7397–7409, 1998.

[133] G. Brannigan, P. F. Phillips, and F. L. H. Brown. Flexible lipid bilayers in implicit solvent. *Phys Rev E*, 72:art., 2005.

[134] T. Sintes and A. Baumgartner. Protein attraction in membranes induced by lipid fluctuations. *Biophys J*, 73:2251–2259, 1997.

[135] T. Sintes and A. Baumgartner. Membrane-mediated protein attraction. a monte-carlo study. *Physica A*, 249:571–575, 1998.

[136] T. Sintes and A. Baumgartner. Interactions of wedge-shaped proteins in flat bilayer membranes. *J. Phys. Chem. B*, 102:7050–7057, 1998.

[137] G. Ayton, S. G. Bardenhagen, P. Mcmurtry, D. Sulsky, and G.A. Voth. Interfacing continuum and molecular dynamics: An application to lipid bilayers. *J Chem Phys*, 114:6913–6924, 2001.

[138] J. F. Nagle and S. Tristram-Nagle. Structure of lipid bilayers. *Biochim Biophys Acta*, 1469:159–195, 2000.

[139] W. J. Sun, S. Tristram-Nagle, R.M. Suter, and J. F. Nagle. Structure of the ripple phase in lecithin bilayers. *Proc Natl Acad Sci*, 93:7008–7012, 1996.

[140] C. Loison, M. Mareschal, K. Kremer, and F. Schmid. Thermal fluctuations in a lamellar phase of a binary amphiphile-solvent mixture: A molecular-dynamics study. *J Chem Phys*, 119:13138–13148, 2003.

[141] C. Loison, M. Mareschal, and F. Schmid. Pores in bilayer membranes of amphiphilic molecules: Coarse-grained molecular dynamics simulations compared with simple mesoscopic models. *J Chem Phys*, 121:1890–1900, 2004.

[142] C. Loison, M. Mareschal, and F. Schmid. Fluctuations and defects in lamellar stacks of amphiphilic bilayers. *Comput Phys Commun*, 169:99–103, 2005.

[143] O. Lenz and F. Schmid. A simple computer model for liquid lipid bilayers. *J Mol Liq*, 117:147–152, 2005.

[144] B. Smit, P. A. J. Hilbers, K. Esselink, L. A. M. Rupert, N. M. van Os, and A. G. Schlijper. Computer simulations of a water oil interface in the presence of micelles. *Nature*, 348:624–625, 1990.

[145] B. Smit, P. A. J. Hilbers, K. Esselink, L. A. M. Rupert, N. M. van Os, and A. G. Schlijper. Structure of a water/oil interface in the presence of micelles: A computer simulation study. *J Phys Chem*, 95:6361–6368, 1991.

[146] B. Smit, K. Esselink, P.A.J. Hilbers, N.M. Van Os, L.A.M. Rupert, and I. Szleifer. Computer simulations of surfactant self-assembly. *Langmuir*, 9:9–11, 1993.

[147] R. Goetz, G. Gompper, and R. Lipowsky. Mobility and elasticity of self-assembled membranes. *Phys Rev Lett*, 82:221–224, 1999.

[148] W. K. Otter and W. J. Briels. The bending rigidity of an amphiphilic bilayer from equilibrium and nonequilibrium molecular dynamics. *J Chem Phys*, 118:4712–4720, 2003.

[149] W. K. Otter. Area compressibility and buckling of amphiphilic bilayers in molecular dynamics simulations. *J Chem Phys*, 123:214906, 2005.

[150] A. Imparato, J. C. Shillcock, and R. Lipowsky. Shape fluctuations and elastic properties of two-component bilayer membranes. *Europhys Lett*, 69:650–656, 2005.

[151] A. Imparato, J. C. Shillcock, and R. Lipowsky. Lateral and transverse diffusion in two-component bilayer membranes. *Eur Phys J E*, 11:21–28, 2003.

[152] M. J. Stevens, J. H. Hoh, and T. B. Woolf. Insights into the molecular mechanism of membrane fusion from simulation: Evidence for the association of splayed tails. *Phys Rev Lett*, 91:art., 2003.

[153] M. J. Stevens. Complementary matching in domain formation within lipid bilayers. *J. Am. Chem. Soc*, 127:15330–15331, 2005.

[154] J.C. Shelley, M.Y. Shelley, R.C. Reeder, S. Bandyopadhyay, P.B. Moore, and M.L. Klein. Simulations of phospholipids using a coarse grain model. *J Phys Chem B*, 105:9785–9792, 2001.

[155] C. F. Lopez, P. B. Moore, J. C. Shelley, M. Y. Shelley, and M. L. Klein. Computer simulation studies of biomembranes using a coarse grain model. *Comput Phys Commun*, 147:1–6, 2002.

[156] S. O. Nielsen, C. F. Lopez, P B. Moore, J. C. Shelley, and M. L. Klein. Molecular dynamics investigations of lipid Langmuir monolayers using a coarse grain model. *J Phys Chem B*, 107:13911–13917, 2003.

[157] Y. Liu and T. Ichiye. Soft sticky dipole potential for liquid water: A new model. *J Phys Chem*, 100:2723–2730, 1996.

[158] L. Whitehead, C. M. Edge, and J. W. Essex. Molecular dynamics simulation of the hydrocarbon region of a biomembrane using a reduced representation model. *J Comput Chem*, 22:1622–1633, 2001.

[159] M. Orsi, J. Michel, and J. W. Essex. Coarse-grain modelling of DMPC and DOPC lipid bilayers. *J. Phys.: Condens. Matter*, 22:155106, 2010.

[160] M. Orsi, W. E. Sanderson, and J. W. Essex. Permeability of small molecules through a lipid bilayer: a multiscale simulation study. *J. Phys. Chem. B*, 113:12019–12029, 2009.

[161] J. Michel, M. Orsi, and J. W. Essex. Prediction of partition coefficients by hybrid atomic level/coarse-grain simulations. *J. Phys. Chem. B*, 112:657–660, 2008.

[162] M. Rueda, S. G. Kalko, F. J. Luque, and M. Orozco. The structure and dynamics of DNA in the gas phase. *J Am Chem Soc*, 125:8007–8014, 2003.

[163] D. L. Beveridge and K.J. McConnell. Nucleic acids: theory and computer simulation, Y2K. *Curr Opin Struc Biol*, 10:182–196, 2000.

[164] T. E. Cheatham III. Simulation and modelling of nucleic acid structure, dynamics and interactions. *Curr Opin Struc Biol*, 14:360–367, 2004.

[165] A.K. Mazur. Accurate DNA dynamics without accurate long-range electrostatics. *J Am Chem Soc*, 120:10928–10937, 1998.

[166] A.K. Mazur. Molecular dynamics of minimal B-DNA. *J Comp Chem*, 22:457–467, 2001.

[167] M.L. Huertas, S. Navarro, M.C.L. Martines, and J. G de la Torre. Simulation of the conformation and dynamics of a double-helical model for DNA. *Biophys J*, 73:3142–3153, 1997.

[168] X.F. Wang, X.L. Lei, and H.P. Fang. What governs the unzipping process of double-stranded DNA. *Chin. Phys. Lett*, 23:1339–1342, 2006.

[169] A. Savelyev and G.A. Papoian. Molecular renormalization group coarse-graining of polymer chains: Application to double-stranded DNA. *Biophys. J*, 96:4044–4052, 2009.

[170] F. Trovato and V. Tozzini. Supercoiling and local denaturation of plasmids with a minimalist DNA model. *J. Phys. Chem. B,*, 112:13197–13200., 2008.

[171] M. Kenward and K.D. Dorfman. Brownian dynamics of single-stranded DNA hairpins. *J. Chem. Phys*, 130:095101, 2009.

[172] M. Daune. *Molecular Biophysics- structures in motion.* Oxford University Press, 1999.

[173] E.J. Sambriski, D.C. Schwartz, and J. J. de Pablo. A mesoscale model of DNA and its renaturation. *Biophys. J*, 96:1675–1690, 2009.

[174] E.J. Sambriski, D.C. Schwartz, and J. J. de Pablo. Sequence effects in the melting and renaturation of short DNA oligonucleotides: structure and mechanistic pathways. *J.Phys.: Condens. Matter*, 21:034105, 2009.

[175] G.A. Voth. *Coarse-Graining of Condensed Phase and Biomolecular Systems.* CRC Press, 2008.

[176] D. Sprous and S.C. Harvey. Action at a distance in supercoiled DNA: effects of sequence on slither, branching, and intramolecular concentration. *Biophys J*, 70:1893–1908, 1996.

[177] D. Sprous, K. Robert, Z. Tan, and S.C. Harvey. Molecular modelling of closed circular DNA thermodynamic ensembles. *Biopolymers*, 39:243–258, 1996.

[178] K. Robert, Z. Tan, D. Sprous, and S.C. Harvey. Molecular dynamics simulations of small DNA plasmids: Effects of sequence and supercoiling on intramolecular motions. *Biopolymers*, 39:259–278, 1996.

[179] M. Sales-Pardo, R. Guimera, A. Moreira, J. Widom, and L.A.N. Amaral. Mesoscopic modeling for nucleic acid chain dynamics. *Phys. Rev. E*, 71:051902, 2005.

[180] F. Zhang and M. A. Collins. Model simulations of DNA dynamics. *Phys Rev E*, 52:4217–4224., 1995.

[181] N. Dan. The structure of DNA complexes with cationic liposomes -cylindrical or flat bilayers? *Biochimica et Biophysica Acta*, 1369:34–38, 1998.

[182] S. Bandyopadhyay, M Tarek, and M.L. Klein. Molecular dynamics study of lipid-DNA complex. *J Phys Chem B*, 103:10075–10080, 1999.

[183] P. Poulain, A. Saladin, B. Hartmann, and C. Prvost. Insights on protein-dna recognition by coarse grain modelling. *J.Comp.Chem.*, 29:2582–2592, 2008.

[184] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. Comparison of simple potential functions for simulating liquid water. *J Chem Phys*, 79:926–935, 1983.

[185] F. Artzner, R. Zantl, G. Rapp, and J.O. Radler. Observation of a rectangular columnar phase in condensed lamellar cationic lipid-DNA complexes. *Phys Rev Lett*, 81:5015, 1998.

[186] R. Zantl, F. Artzner, G. Rapp, and J.O. Radler. Thermotropic structural changes of saturated-cationic-lipid-DNA complexes. *Europhys. Lett.*, 45:90, 1999.

[187] S. L. Price, A. J. Stone, and M. Alderton. Explicit formulas for the electrostatic energy, forces and torques between a pair of molecules of arbitrary symmetry. *Mol Phys*, 52:987–1001, 1984.

[188] R. Koynova and M. Caffrey. Phases and phase transitions of the phosphatidylcholines (vol 1376, pg 91, 1998). *Bba-biomembranes*, 1513:82–82, 2001.

[189] W. L. Jorgensen and J.D. Madura. Temperature and size dependence for monte carlo simulations of tip4p water. *Mol.Phys*, 56:1381–1392, 1985.

[190] P.W. Atkins. *Physical Chemistry.* Oxford University Press, 1984.

[191] G.S. Kell. Density, thermal expansivity, and compressibility of liquid water from 0 to 150C: Correlations and tables for atmospheric pressure and saturation reviewed and expressed on 1968 temperature scale. *J. Chem. Eng. Data*, 20:97–1105, 1975.

[192] A.K. Soper, F. Bruni, and M.A. Ricci. Site-site pair correlation functions of water from 25 to 400C: Revised analysis of new and old diffraction data. *J Chem Phys*, 106:247–254, 1997.

[193] A.K. Soper and M.G. Philips. A new determination of the structure of water at 25 C. *Chem Phys*, 107:47–60, 1986.

[194] G. Hura, J. M. Sorenson, R. M. Glaeser, and T. Head-Gordon. A high-quality x-ray scattering experiment on liquid water at ambient conditions. *J Chem Phys*, 113:9140–9148, 2000.

[195] T. Head-Gordon and G. Hura. Water structure from scattering experiments and simulation. *Chem Rev*, 102:2651–2669, 2002.

[196] A.H. Narten and H.A. Levy. Liquid water: molecular correlation functions from x-ray diffraction. *J Chem Phys*, 55:2263, 1997.

[197] M. Winger, D. Trzesniak, R. Baron, and W.F. van Gunsteren. On using a too large integration time step in molecular dynamics simulations of coarse-grained molecular models. *Phys. Chem. Chem. Phys*, 11:1934–1941, 2009.

[198] H. Akutsu and T. Nagamori. Conformational-analysis of the polar head group in phosphatidylcholine bilayers - a structural-change induced by cations. *Biochemistry*, 30:4510–4516, 1991.

[199] J.C.W. Shepherd and G. Buldt. Zwitterionic dipoles as a dielectric probe for investigating head group mobility in phospholipid membranes. *Biochim Biophys Acta*, 514:83–94, 1978.

[200] E.R. Maya, D.I. Kopelevicha, and A. Narang. Coarse-grained molecular dynamics simulations of phase transitions in mixed lipid systems containing lpa, dopa, and dope lipids. *Biophys. J.*, 94:878–890, 2008.

[201] J. Aquist. Ion-water interaction potentials derived from free eneregy pertubation simulations. *J.Phys.Chem.*, 94:8021–8024, 1990.

[202] L. Onsager. Electric moments of molecules in liquids. *J.Am. Chem. Soc.*, 58:1486–1493, 1936.

[203] J.A. Barker and R.O. Watts. Monte Carlo studies of the dielectric properties of water-like models. *Mol. Phys.*, 26:789–792, 1973.

[204] H.E. Alper and R.M. Levy. Computer simulations of the dielectric properties of water: Studies of the simple point charge and transferrable intermolecular potential models. *J. Chem Phys.*, 91:1242–1251, 1989.

[205] J. F. Nagle and S. Tristram-Nagle. Lipid bilayer structure. *Curr Opin Struc Biol*, 10:474–480, 2000.

[206] J. F. Nagle. Lipid bilayer structure. *Biophys J*, 80:135, 2001.

[207] J. Vieillard-Baron. Phase transitions of the classical hard-ellipse system. *J.Chem.Phys*, 56:4729–4744, 1972.

[208] J. Vieillard-Baron. The equation of state of a system of hard spherocylinders. *Mol.Phys*, 28:809–818, 1974.

[209] R. Eppenga and D. Frenkel. Monte carlo study of the isotropic and nematic phases of infinitely thin hard platelets. *Mol.Phys*, 52:1303–1334, 1984.

[210] G. R. Luckhurst and G. W.(eds.) Gray. *The Molecular Physics of Liquid Crystals.* Academic Press, 1979.

[211] A. M. Smondyrev and M. L. Berkowitz. United atom force field for phospholipid membranes: Constant pressure molecular dynamics simulation of dipalmitoylphosphatidicholine/water system. *J Comput Chem*, 20:531–545, 1999.

[212] D.E. Smith and L.X. Dang. Computer simulations of NaCl association inn polarizable water. *J.Chem.Phys*, 100:3757–3766, 1994.

[213] P. Auffinger, T.E. Cheatham III, and A.C. Vaiana. Spontaneous formation of KCl aggregates in biomolecular simulations: A force field issue? *J.Chem.Theory Comput.*, 3:1824–1859, 2007.

[214] A.P. Copestake, G.W. Neilson, and J.E. Enderby. The structure of a highly concentrated aqueous solution of lithium chloride. *J.Phys.C:Solid State*, 18:4211–4216, 1982.

[215] S. C. Harvey, R. K. Z. Tan, and T. E. Cheatham III. The flying ice cube: Velocity rescaling in molecular dynamics leads to violation of energy equipartition. *J Comput Chem*, 19:726–740, 1998.

[216] Z. Yi, M. Nagao, and D.P. Bossev. Bending elasticity of saturated and monounsaturated phospholipid membranes studied by the neutron spin echo technique. *J.Phys.Condens.Matter*, 21:155104, 2009.

[217] E. Sackmann. *Physical Basis of Self-Organization and Function of Membranes: Physics of Vesicles.* Elsevier, 1995.

[218] E. Evans and D. Needham. Physical properties of surfactant bilayer membranes: Thermal transitions , elasticity, rigidity, cohesion, and colloidal interactions. *J.Phys.Chem.*, 91:4219–4228, 1987.

[219] C.H. Lee, W.C. Lin, and J. Wang. All-optical measurements of the bending rigidity of lipid-vesicle membranes across structural phase transitions. *Phys. Rev. E.*, 64:020901, 2001.

[220] E. Lindahl and O. Edholm. Mesoscopic undulations and thickness fluctuations in lipid bilayers from molecular dynamics simulations. *Biophys J*, 79:426–433, 2000.

[221] W. Rawicz, K.C. Olbrich, T. McIntosh, D. Needham, and E. Evans. Effect of chain length and unsaturation on elasticity of lipid bilayers. *Biphys. J.*, 79:328–339, 2000.

[222] J. Daillant, E. Bellet-Amalric, A. Braslau, T. Charitat, G. Fragneto, F. Graner, S. Mora, F. Rieutord, and B. Stidder. Structure and fluctuations of a single floating lipid bilayer. *Proc.Natl.Acad.Sci.*, 102:11639–11644, 2005.

[223] I. Fernandez-Puente, L.and Bivas, M.D. Mitov, and P. Meleard. Temperature and chain length effects on bending elasticity of phosphatidylcholine bilayers. *Europhys. Lett.*, 28:181–186, 1994.

[224] W. H. Press. *Numerical recipes in C : the art of scientific computing.* Cambridge University Press,, 1992.

[225] A. Seelig and J. Seelig. Dynamic structure of fatty acyl chains in a phospholipid bilayer measured by deuterium magnetic-resonance. *Biochemistry*, 13:4839–4845, 1974.

[226] J.P. Douliez, A. Leonard, and E.J. Dufourc. Restatement of order parameters in biomembranes - calculation of C-C bond order parameters from C-D quadrupolar splittings. *Biophys J*, 68:1727–1739, 1995.

[227] W. Shinoda, N. Namiki, and S. Okazaki. Molecular dynamics study of a lipid bilayer: Convergence, structure, and long-time dynamics. *J.Chem. Phys.*, 106:5731–5743, 1997.

[228] U. Essmann and M.L. Berkowitz. Dynamical properties of phospholipid bilayers from computer simulation. *Biophys. J.*, 76:2081–2089, 1999.

[229] A. Blume, G. Cevc, and M. Dekker. *Dynamic Properties-Phospholipids handbook.* CRC Press, 1993.

[230] W. Shinodaa, M. Mikamia, T. Babab, and M Hato. Dynamics of a highly branched lipid bilayer: a molecular dynamics study. *Chem. Phys. lett.*, 76:35–40, 2004.

[231] G. Cevc and M. Dekker. *Phospholipids handbook.* CRC Press, 1993.

[232] T. Wymore, X. F. Gao, and T. C. Wong. Molecular dynamics simulation of the structure and dynamics of a dodecylphosphocholine micelle in aqueous solution. *J Mol Struct*, 485-486:195–210, 1999.

[233] P. Cifra. Differences and limits in estimates of persistence length for semi-flexible macromolecules. *Polymer*, 45:5995–6002, 2004.

[234] J.A.H. Cognet, C. Pakleza, D. Cherny, E. Delain, and E. Le Cam. Static curvature and flexibility measurements of DNA with microscopy. a simple renormalization method, its assessment by experiment and simulation. *J.Mol.Bio.*, 285:997–1009, 1999.

[235] E.B. Safinya, C.R.and Sirota, D. Roux, and G.S. Smith. Universality in interacting membranes: The effect of cosurfactants on the interfacial rigidity. *Phys.Rev.Lett.*, 62:1134–1137, 1989.

[236] E. Giudice and Y. Lavery. Simulations of nucleic acids and their complexes. *Acc Chem Res*, 35:350–357, 2002.

[237] C. S. O Hern and T.C. Lubensky. Sliding columner phase of DNA-Lipid complexes. *Phys Rev Lett*, 80:4345–4348, 1998.

[238] R. Bruinsma. Electrostatics of DNA -cationic lipid complexes: isoelectric instability. *Eur Phys J B*, 4:75–88, 1998.

[239] H. Gershon, R. Ghirlando, S.B. Guttman, and A. Minsky. Mode of formation and structural features of DNA-cationic liposome complexes used for transfection. *Biochem*, 32:7143–7151, 1993.

[240] B. Sternberg, F.L. Sorgi, and L. Huang. New structure formation between DNA and cationic liposomes visualized by freeze-fracture electron microscopy. *FEBS Lett*, 356:361–366, 1994.

[241] D.D. Lasic, H. Strey, M.C.A. Stuart, R. Podgornik, and P.M. Frederik. The structure of DNA liposome complexes. *J Am Chem Soc*, 119:832–833, 1997.

[242] J.O. Radler, I. Koltover, T. Salditt, and C.R Safinya. Structure of DNA cationic liposome complexes: DNA intercalation in multilamellar membranes in distinct interhelical packing regimes. *Science*, 275:810–814, 1997.

[243] D. Hirsch-Lerner and Y. Barenholz. Probing DNA cationic lipid interactions with flurophore trimethylamonium diphenyl- hexatriene (tmadph). *Biochimica et Biophysica Acta*, 1370:17–30, 1998.

[244] J. Kushick and B. J. Berne. Computer simulation of anisotropic molecular fluids. *J Chem Phys*, 64:1362–1367, 1976.

[245] A.L. Tsykalo and A.D. Bagmet. Molecular dynamics study of the nematic liquid crystals. *Mol.Cryst.Liq.Cryst*, 46:111–119, 1978.

[246] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans. *Intermolecular Forces.* Reidel, Dordrecht, 1981.