MOLECULAR DYNAMICS OF A RESPIRATORY PROTEIN FRAGMENT IN PHOSPHOLIPID MONOMOLECULAR FILMS

SANTOSH K. GUPTA

Molecular Dynamics of a Respiratory Protein Fragment in Phospholipid Monomolecular Films

by

© Santosh K. Gupta

A thesis submitted to the

School of Graduate Studies

in partial fulfilment of the

requirements for the degree of

Master of Science

Department of Physics

Memorial University of Newfoundland

October, 2009

St. John's   Newfoundland

Abstract

Molecular dynamics simulations of the fragment Mini-B of Surfactant Protein B were carried out in the NVT ensemble for simulation times ranging from 20 ns to 80 ns. These simulations were performed in DPPC:POPG monolayers and POPG monolayers at different cross-sectional areas per lipid, varying the initial depth and orientation of the peptide. Analysis was carried out to determine what depth the peptide can attain in the monolayers towards the end of the simulations. Secondary structure of the peptide was also examined, and it was found that in the majority of cases both alpha helices remained alpha helical to a reasonable degree in the monolayers. The tilt angles of the alpha helices of Mini-B in the DPPC:POPG monolayers seemed to show relatively small deviations from those of the corresponding structure determined in SDS micelles by NMR, provided that the peptide was placed at a low enough depth. A study was also performed to determine the effect of sodium ions on the interaction between the cationic Arginine and Lysine residues of Mini-B and the anionic phosphate headgroups of POPG.

Acknowledgements

This work was compiled under the supervision of Dr. Valerie Booth. I would like to thank Jennifer Rendell for help in general and useful information during various stages of this work. I would also like to acknowledge both her and John Bartlett for procedures and perlscripts used for the simulations. Finally, I would like to thank Muzaddid Sarker for helpful information.

These simulations were performed on ACE-net and Westgrid clusters, and I would also like to acknowledge the support teams for their aid.

Contents

v

List of Tables

List of Figures

Chapter 1

Introduction to Lung Surfactant

1.1 Overview of Lung Surfactant

Lung surfactant, which lines the inner wall of the alveolus, reduces the surface tension at the interface separating the air within the alveolus and the aqueous zone (or "subphase") on the other side of the film [1]. It is a mixture of phospholipids, neutral lipids, and proteins in the form of a thin monomolecular film, ie. a monolayer. In the absence of lung surfactant, the air-water interface has a surface tension of approximately 72 mN/m, and a substantial amount of work is needed to re-expand the alveolus after expiration. Lack of surfactant may lead to conditions such as RDS (Respiratory Distress Syndrome), which occurs most frequently in neo-nates [2, 3].

Surface tension forces increase the amount of work needed to expand a

1

surface (here the terms "surface", 'film", and "monolayer" will be used inter-changeably where it is convenient since only surfaces in the form of monomole-cular films are dealt with in the following). The surface tension in lung sur-factant reduces to nearly 0 mN/m when the film is fully compressed at the end of expiration [4]. This reduction in surface tension makes it possible for the alveoli to re-expand with minimal effort upon air intake. Without this decrease in surface tension, the alveoli would undergo a state of collapse and the breathing process would not continue normally [5].

The surface tension of a film at the air-water interface is denoted by the parameter γ, and it can also be expressed in terms of the surface pressure π:

$$\pi = \gamma_{water} - \gamma. \tag{1.1}$$

Here γ_{water} is the surface tension of pure water. Hence a reduction of surface tension can be interpreted as an increase in surface pressure. The surface pressure is often used in the context of describing experimental results on a broad class of films, and has other uses from a theoretical perspective. Lung surfactant is part of one class of films referred to as surfactants [6], and the surface pressure characteristics of surfactants will be described in further detail in Section 1.2.

Lung surfactant is secreted by type II pneumocytes, which are part of the epithelial cells that make up the outer wall of the alveolus [7]. The phospho-

2

lipids that make up the surfactant first form bilayer structures in the subphase that serve as a source or "reservoir" from which lipid components are added to the interface [8, 9]. At the interface, the lipid components arrange themselves in a side-by-side fashion with phosphate headgroups pointing into the subphase and lipid tails pointing into the air medium. This is illustrated in Fig. 1.1 below.
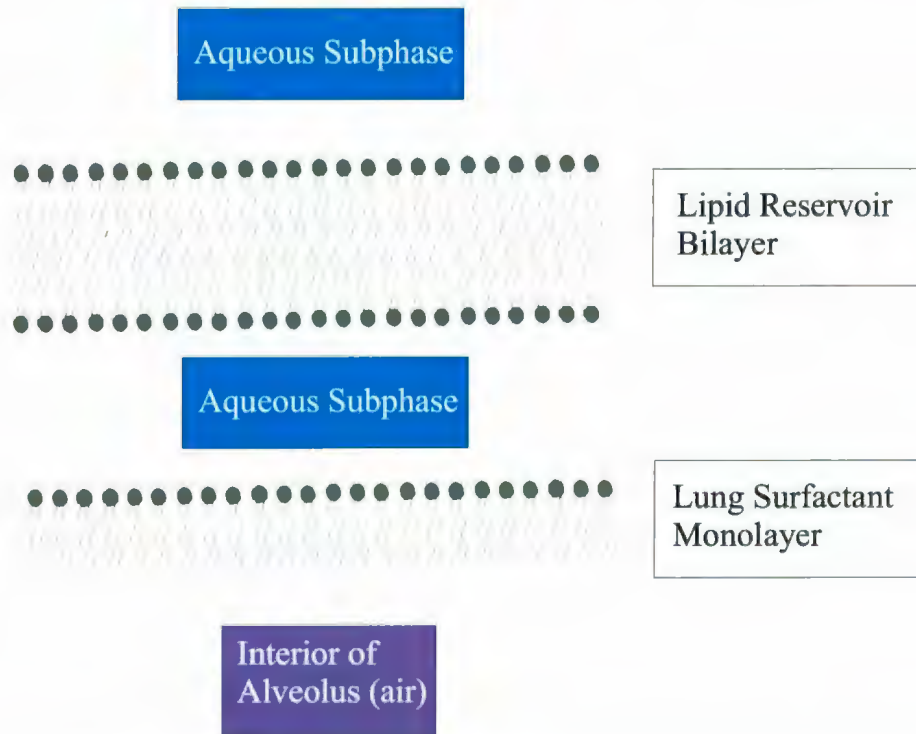
Figure 1.1: Lung surfactant at air-water interface with nearby lipid reservoir bilayer.

Lung surfactant is made up primarily of phospholipids of the PC (phos-

phatidylcholine) class. The PC class alone makes up approximately 80% of the film [10], which is a very high amount in comparision to the amount present in a variety of cell membranes [11]. These phospholipids are zwitterionic, ie. their phosphate headgroups possess an equal amount of positive and negative charge. Approximately 50% of the PC class phospholipids by weight is DPPC (dipalmitoylphosphatidylcholine), which possesses unbranched, fully saturated hydrocarbon chains [12]. One key property of DPPC is that a pure monolayer of DPPC can attain low surface tension values (similar to those of lung surfactant) upon compression [13]. Other significant phospholipid classes are PE (Phosphatidylethanolamine), PG (Phosphatidylglycerol), PI (Phosphatidylinositol), and PG (Phosphatidylglycerol).

Other noteworthy constituents of lung surfactant are the proteins SP-A, SP-B, SP-C, and SP-D [14] (Here SP stands for "Surfactant Protein"). Of these, SP-B and SP-C are most critical for the ability of lung surfactant to adsorb efficiently onto the air-water interface [15]. The surfactant proteins, and their roles in the function of lung surfactant, are described further in section 1.3. Emphasis, however, will be laid on the structure and activity of SP-B, as it is most relevant to the simulations discussed in Chapters 3-5.

1.2   Surfactants and Surface Pressure

Isotherms

Before describing surfactant films and their surface pressure characteristics,
it is useful to illustrate the concepts of surface tension and surface pressure.
Surface tension forces arise due to an asymmetry in intermolecular interactions
at an interface [16]. For example, if one considers pure water adjacent to a
medium of air, the water molecules in the bulk part of the liquid experience
attractive polar forces from all sides. However, these forces cancel out. On
the other hand, water molecules at the interface only experience these polar
forces from one side due to the fact there is only air on the other side (Fig.
1.2 below). This leads to a net force pulling the water molecules inward and
a consequent decrease in the surface area of the liquid. At equilibrium, the
pressure difference across the interface will balance this net force leading to a
total force of zero on each of the molecules, but for the present we will simply
speak of the net force that results due to the asymmetry in intermolecular
forces before the system comes to equilibrium. The net force per unit length
(typically expressed in mN/m), is equivalent to the surface tension of the
substance.

Thermodynamically, the surface tension is defined as the change in Gibbs
free energy with respect to surface area:
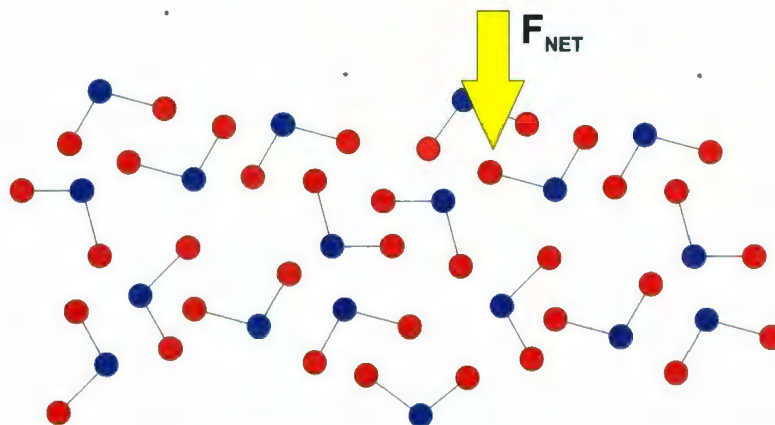
Figure 1.2: A diagram depicting the origins of surface tension.

$$\gamma = \left(\frac{\partial G}{\partial A}\right)_{T,P}, \tag{1.2}$$

where the subscripts denote that temperature and pressure are held constant. Hence the surface tension represents the amount of work needed to expand the surface area. A positive surface tension implies that a surface tends to adopt a configuration which has the lowest surface area attainable given the intermolecular forces present within the surface. This relationship between surface tension and work leads to a relationship between surface tension and the pressure difference across a surface, referred to as the Young-Laplace law.

For a closed surface (the case of an open surface is omitted for simplicity):

$$P = \gamma \left(1/R_1 + 1/R_2 \right),$$
(1.3)

where R_1 and R_2 are the radii of curvature of the surface (in the case of a sphere these would be identical). This implies that if the surface tension of the interface is very high, a sufficiently large pressure would be needed to maintain the shape of the surface, and an even larger pressure would be needed to expand it. Lack of sufficient pressure can lead to collapse, as mentioned earlier with regard to alveoli that lack lung surfactant. In this case, the internal pressure after air intake is not large enough to re-expand the alveolus due to the presence of large surface tension forces.

A surfactant in general is a film of *amphiphillic* molecules which exists at an interface such as one separating a liquid (eg. water) and a gas (eg. air). An amphiphillic molecule possesses both a hydrophillic region and a hydrophobic region [17]. Lung surfactant phospholipids have hydrophillic regions corresponding to the polar phosphate headgroups and hydrophobic regions corresponding to the lipid tailgroups. The amphiphillic character of surfactant molecules affects how the molecules orient themselves at the interface, and also plays a key role in how they interact with each other and with other molecules that may be present at the interface. At an air-water interface, the hydrophillic region will point towards the water while the hydrophobic region

7

will point towards the air, as already mentioned in the case of lung surfactant phospholipids.

A surfactant can exhibit different two-dimensional phases upon undergoing different amounts of compression, similar to how three-dimensional substances exhibit the phases of gas, liquid, and solid. The two-dimensional phases for a surfactant are gas, expanded (or *liquid-expanded*), and condensed. There are also phases at points of compression between these previously mentioned phases which coexist between gas and expanded or expanded and condensed respectively.

A two-dimensional gas consists of molecules which are too far apart to exhibit any significant interaction, much like a three-dimensional gas. In this case the surface pressure of the film obeys an ideal gas type of law [18]

$$\pi A = NkT \qquad (1.4)$$

where π is the film surface pressure, A is the surface area, N is the number of surfactant molecules, k is Boltzmann's constant, and T is the temperature. The relationship between surface pressure and area is referred to as the *surface pressure versus area isotherm* of the film. For the two-dimensional gas it would be a straight line increasing as the surface area became smaller.

When the film is compressed enough, the molecules are brought close enough together for a significant interaction to occur. However, the state

8

of the surfactant molecules is still relatively disordered. This is referred to as the liquid-expanded phase. Upon further compression the molecules tend to organize themselves into the condensed phase, in which the molecules are arranged in a more periodic fashion. The surface pressure-area isotherm exhibits a plateau known as the *collapse zone*, and further compression does not lead to any further changes in surface pressure. In the case of lung surfactant, the hydrophobic tails of the phospholipids may be tilted at varying angles, but become completely perpendicular to the interface at the collapse zone.

While a high collapse pressure close to 72 mN/m is necessary for the proper functioning of lung surfactant, a film of DPPC by itself can exhibit a plateau in its surface-pressure isotherm at a similar magnitude [19]. Hence DPPC by itself would be sufficient for the reduction of surface tension that is necessary at the alveolar air-water interface at the end of expiration.

A film of pure DPPC, however, does not adsorb rapidly enough to the interface [20]. Despite this deficiency in pure DPPC films, the fact that the surface pressure isotherm of DPPC is similar to that of lung surfactant suggests a means by which the surface pressure of lung surfactant can attain high values. This is described by the "classical model" of lung surfactant function, which proposes that constituents of lung surfactant other than DPPC (POPG, cholesterol, etc.) undergo "squeeze-out" upon compression [21]. This would leave behind a monolayer abundant in DPPC and result in higher surface pressure values.

The squeeze-out of lipids from lung surfactant upon compression has been observed by fluoresence microscopy when surfactant proteins SP-B and SP-C are absent, but does not occur when these proteins are present [22]. Instead the monolayer is observed to undergo a buckling transition upon collapse. This results in an effective enhancement of DPPC in the unfolded layer of the film, and hence may be an alternative mechanism for lung surfactant function. The model of squeeze-out of lipids has also been argued against on the basis of kinetic experiments which suggest that squeeze-out is not necessary for the film to exhibit the phase behaviour observed upon compression provided the film is compressed at a high enough rate [23].

1.3   Surfactant Proteins

The surfactant proteins SP-A, SP-B, SP-C, and SP-D fall into two types. The larger proteins SP-A and SP-D are believed to be associated more with immunological functions [24]. The smaller proteins SP-B and SP-C however are believed to play a key role in lung surfactant function, for example with regard to its ability to adsorb new lipid components onto the film.

As already mentioned previously in Section 1.2, DPPC by itself adsorbs too slowly onto the interface to play an adequate role of a surfactant for the lungs. Experiments have shown, however, that the addition of SP-B or SP-C to a film of DPPC can alter its adsorption characteristics drastically [25]. In either

case the surface pressures of the protein-lipid systems rise rapidly to values of approximately 50 mN/m when a lipid vesicle is inserted near the interface for a large range of starting surface pressures. This confirms that these proteins play an important role in the adsorption behaviour of lung surfactant. However, the means by which these proteins carry out this role is not well understood. In the following, further details will be provided on SP-B due to its relevance to Chapters 3-5.

SP-B is a small protein of roughly 8.7 kDa in weight and consists of 79 residues [26, 27, 28]. Its amino acid sequence is provided in Table 1.1. The color-coding scheme according to residue type is provided at the bottom of the table. Note that Glycine is given a special category simply because it does not fall under any of the other classifications.

As can be seen from the figure, the positively charged residues correspond to Arginine and Lysine while the negatively charged residues correspond to Glutamic acid and Aspartic acid (however, residues with net charge are also typically considered to be polar) [29]. The polar (but electrostatically neutral) residues correspond to Cysteine, Serine, and Glutamine. The hydrophobic residues correspond to Valine, Leucine, Isoleucine, Methionine, Alanine, Proline, and Tryptophan. The abundance of hydrophobic residues in SP-B is apparent. However, the distribution of charged and polar residues with respect to hydrophobic residues in space gives it an amphiphillic character. This plays a key role in how it interacts with the phospholipids in lung surfactant.

11

Table 1.1: Amino acid sequence of SP-B

Phe	Pro	Ile	Pro	Leu	Pro	Tyr	Cys	Trp	Leu
1	2	3	4	5	6	7	8	9	10
Cys	Arg	Ala	Leu	Ile	Lys	Arg	Ile	Gln	Ala
11	12	13	14	15	16	17	18	19	20
Met	Ile	Pro	Lys	Gly	Ala	Leu	Ala	Val	Ala
21	22	23	24	25	26	27	28	29	30
Val	Ala	Gln	Val	Cys	Arg	Val	Val	Pro	Leu
31	32	33	34	35	36	37	38	39	40
Val	Ala	Gly	Gly	Ile	Cys	Gln	Cys	Leu	Ala
41	42	43	44	45	46	47	48	49	50
Glu	Arg	Tyr	Ser	Val	Ile	Leu	Leu	Asp	Thr
51	52	53	54	55	56	57	58	59	60
Leu	Leu	Gly	Arg	Met	Leu	Pro	Gln	Leu	Val
61	62	63	64	65	66	67	68	69	70
Cys	Arg	Leu	Val	Leu	Arg	Cys	Ser	Met	
71	72	73	74	75	76	77	78	79	

Color-coding scheme: yellow = hydrophobic, grey = negatively charged, cyan = positively charged, green = polar (neutral), tan = Glycine (other)

Although the full tertiary structure of SP-B has not yet been determined, its homology to related proteins had to led to the belief that it consists of four alpha helices stabilized by three disulfide bonds. On the other hand, the structure of several different fragments of SP-B have been determined in different mediums, in most cases by using one or more alpha helices of the full structure. For example, SP-B$_{CTERM}$ consists of one terminal alpha helix of the full structure, and its structure has been determined by NMR in HFIP (hexafluoroisopropanol) and SDS (sodium dodecyl sulfate) micelles [30]. When both terminal helices are retained, as in the case of Mini-B, the helices can be linked together and stabilized by disulfide bridges. This structure has also been determined by NMR in HFIP and SDS [31]. Fig. 1.3 illustrates how Mini-B is derived from SP-B. Here a rough topology of the different structures are drawn, with alpha helices depicted as coils. In addition, Mini-B is seen to retain two of the disulfide bridges of full length SP-B. It also retains an overall charge of +7 at neutral pH.

Mini-B, which is used in the simulations in Chapters 3-5, shows a similar level of function as full-length SP-B in rodent models with respect to increases in blood oxygen levels [32]. It consists of residues 8-25 and 63-78 of SP-B and is also amphiphillic. The amino acid sequence of Mini-B is provided separately in Table 1.2, using the same color coding scheme as in Table 1.1.

Graphical representations of Mini-B in SDS generated by the program Visual Molecular Dynamics (VMD) [33] are provided below in Fig. 1.4. In Fig.

13
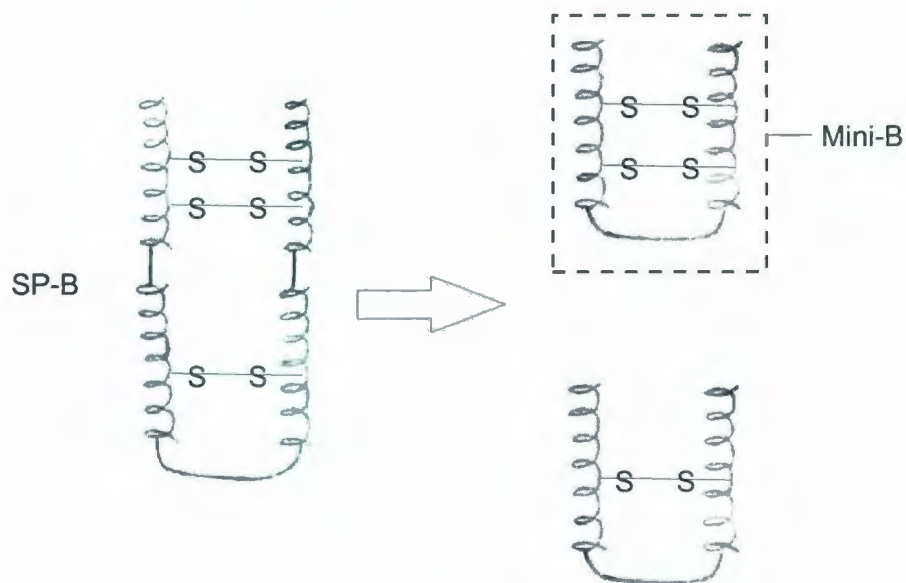
Figure 1.3: Mini-B as it is derived from SP-B.

1.4(A) an atomic structure is shown from a side view, with a slightly different color coding scheme. This color coding scheme will be used in remaining graphical representations of Mini-B. In Fig. 1.4(B) a top view backbone ribbon structure is shown, showing the alpha helices. The bottom helix is in the N-terminal region of the peptide, while the top helix is in the C-terminal region.

Table 1.2: Amino acid sequence of Mini-B

Cys	Trp	Leu	Cys	Arg	Ala	Leu	Ile	Lys	Arg
1	2	3	4	5	6	7	8	9	10
Ile	Gln	Ala	Met	Ile	Pro	Lys	Gly	Gly	Arg
11	12	13	14	15	16	17	18	19	20
Met	Leu	Pro	Gln	Leu	Val	Cys	Arg	Leu	Val
21	22	23	24	25	26	27	28	29	30
Leu	Arg	Cys	Ser						
31	32	33	34						

Color coding scheme for atomic structure (A):
yellow = hydrophobic, blue = positively charged,
green = polar (neutral), tan = Glycine(other).
This color coding scheme is used in all remaining
atomic representations of Mini-B.

(A)
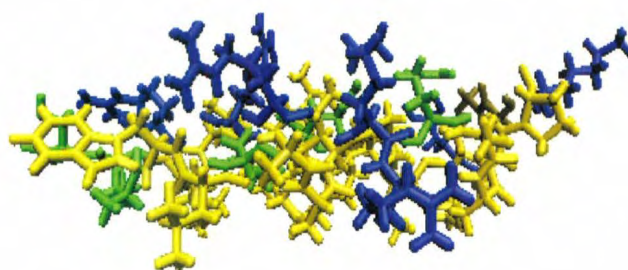
(B)
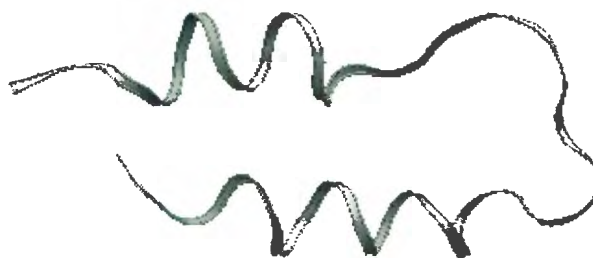
Figure 1.4: Graphics of Mini-B in SDS. (A) Side view atomic structure, (B) Top view backbone ribbon structure.

Chapter 2

Molecular Dynamics Simulations

Molecular Dynamics Simulation is a technique for obtaining the trajectories of a large group of atoms or molecules from starting configurations, using an empirical force-field to represent the forces between atoms. The coordinates of the molecules are loaded into a structure file, along with initial velocities (typically generated from a Maxwell-Boltzmann distribution), and the evolution of the system follows Newton's equations of motion given a proper force field. The GROMACS (Groningen Machine for Chemical Simulations) software package is a program for performing molecular dynamics simulations (henceforth referred to as MD simulations) [34, 35, 36, 37]. The simulations in Chapters 3-5 used different sub-versions of GROMACS version 3.3, depending

on which sub-version was available on a given machine.

2.1 Basic Algorithms

The preliminary step in determining the trajectories of the atoms is to integrate
Newton's equations of motion and apply periodic boundary conditions to avoid
edge effects. In this case the force-field is derived from the negative gradient of
a given potential function (In GROMACS, the terms "potential function" and
"force-field" are sometimes used interchangeably). The potential functions
themselves have numerical parameters that are provided by experimental data
and/or quantum mechanical considerations, such as the $\epsilon - \sigma$ parameters for
the Lennard-Jones interaction potential described in Section 2.2.1. This set
of parameters is referred to as a molecular topology. Since these parameters
determine the strengths of the interactions between atoms in a given molecule,
in addition to how these atoms interact with other atoms outside the molecule,
these parameters also determine the configurations of different molecules in
space.

The system is bound by a region of space defined by a box, and periodic
boundary conditions result in multiple images of the system on each side of the
box. In many cases a rectangular box can be used, especially when monolayers
are involved in the form of planar structures.

Newton's equations of motion are then integrated using the leap-frog algo-

rithm:

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \vec{v}_i\left(t + \frac{\Delta t}{2}\right)\Delta t \tag{2.1}$$

$$\vec{v}_i(t + \frac{\Delta t}{2}) = \vec{v}_i(t - \frac{\Delta t}{2}) + \frac{\vec{F}_i(t)}{m}\Delta t \tag{2.2}$$

where $\vec{r}_i(t)$ and $\vec{v}_i(t)$ represent the coordinates and velocities of the ith atom, t is the time at a particular step in the trajectory, and Δt is the time interval between steps in the trajectory.

Often it is desired that the system is kept at a fixed temperature (and sometimes also fixed pressure) if this matches well with empirical conditions. However, a direct integration of Newton's equations of motion through the leap-frog algorithm could lead to large changes in total kinetic energy, and consequently lead to a drastic change in the temperature. Therefore additional constraints on the equation of motion are required if the system is to be kept at a fixed temperature. One method, referred to as Berendsen temperature coupling, causes deviations from a fixed temperature value to decay exponentially according to a decay time constant τ [38],

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau} \tag{2.3}$$

where T_0 is the desired fixed tempature, and T is the actual temperature

19

of the system as a function of time as determined by the total kinetic energy E_{kin} of the atoms in the system,

$$T = \frac{2E_{kin}}{k_B N_{df}}.$$ (2.4)

Here N_{df} is the number of degrees of freedom of the system, k_B is Boltzmann's constant, and E_{kin} is defined as simply

$$E_{kin} = \sum_{i=1}^{N} \frac{1}{2} m_i v_i^2.$$ (2.5)

The m_i are the masses of individual atoms in the system, and N represents the total number of atoms in the system.

The exponential decay in temperature deviations is carried out by scaling the velocities at each time step by a constant factor λ given by

$$\lambda = \left(1 + \frac{\Delta t}{\tau'} \left(\frac{T_0}{T(t - \frac{\Delta t}{2})} - 1\right)\right)^{1/2}$$ (2.6)

where

$$\tau' = \frac{N_{df} k_B \tau}{2 C_V}$$ (2.7)

and C_V is the heat capacity of the system.

Pressure coupling can be obtained in a similar manner, except that one must take into account that the pressure can have different components in

20

different directions. Further description of pressure coupling will not be discussed, however, since it is not used in the simulations in later chapters.

The approach of integrating Newton's equations by itself is unable to treat the motions of atoms connected by bonds that vibrate at frequencies greater than that due to thermal fluctuations (ie. $k_B T/\hbar$, which is roughly 10^{12} Hz). In this case quantum effects would need to be considered. Full-scale quantum dynamics is avoided in GROMACS by placing constraints on bond lengths. These constraints on the bond lengths are applied after use of the leap-frog algorithm (along with the application of the temperature coupling constraint in Eqn. 2.3). Although different methods exist for applying these constraints, one popular method is the LINCS (LINear Constraint Solver) algorithm. This algorithm works by resetting bond lengths to their correct values after they undergo rotation and stretching according to Newton's equations of motion [39].

Energy Minimization

In addition to determining the trajectories of atoms in a system, the GROMACS program can also perform energy minimization of the system (ie. modify the positions of the atoms to obtain a lower energy configuration). Energy minimization can be performed through one of the following methods: steepest descents, conjugate gradients, or l-bfgs (limited memory Broyden-Fletcher-Goldfarb-Shanno).

21

It is also possible to use a combination of different methods. Here a more detailed description of the steepest descents method will be given, though other methods are based on similar underlying concepts.

The steepest descents method works by displacing the coordinates of the system in a direction that is given by the (negative) gradient in the potential at the original point. In other words, if $\vec{r}_0$ is the original set of coordinates of the system, the final coordinates $\vec{r}_1$ for the first step in the algorithm are obtained by

$$\vec{r}_1 = \vec{r}_0 + \frac{\vec{F}}{\max\left(\left|\vec{F}\right|\right)} h_0. \tag{2.8}$$

Here $\max\left(\left|\vec{F}\right|\right)$ is the magnitude of the force which has the maximum value compared to the magnitudes of all other forces on the atoms within the system, and h_0 is some initial maximum displacement (with a scalar value) that is specified by the user. For each subsequent step the maximum displacement is determined by whether or not there was a decrease in the total potential energy. (There might not be a decrease in the energy, if, for example, there is a very sharp transition in the potential and the displacement carried the coordinates too far). If k is used to denote the current step in the algorithm, and U_k is the potential energy at step k, then $h_{k+1} = 1.2h_k$ if the potential energy decreases and $h_{k+1} = 0.2h_k$ if the potential energy does not decrease. Hence in general the coordinates at each step are determined by

$$\vec{r}_{k+1} = \vec{r}_k + \frac{\vec{F}_k}{\max(|\vec{F}_k|)} h_k. \tag{2.9}$$

The user typically must specify a condition for which the energy minimization algorithm will decide that the potential has been reduced enough and come to a stop. Often this condition is in terms of a tolerance value for $\max\left(|\vec{F}_k|\right)$, and the algorithm will stop once $\max\left(|\vec{F}_k|\right)$ is below this tolerance value.

A useful means for choosing this tolerance value is obtained by considering the force which causes a bond to vibrate at a particular frequency ν. Near equilibrium, the only source of energy by which the bond will stretch is due to thermal fluctuations. Hence the maximum force on a bonded atom should give an energy that is roughly on the order of $k_B T$. It is also assumed that the fluctuations are small enough so that the bonded atoms can be considered to be in a harmonic oscillator well with angular frequency $\omega = 2\pi\nu$. If ϵ is used to denote the tolerance value, then the maximum displacement of the system from equilibrium is roughly $\epsilon/m\omega^2$, and the energy corresponding to this must be

$$\frac{1}{2}m\omega^2 \frac{\epsilon^2}{m^2\omega^4} = \frac{1}{2}\frac{\epsilon^2}{m\omega^2}. \tag{2.10}$$

Hence this leads to the condition

$$\epsilon^2 \approx 2mk_BT\omega^2 \Rightarrow \epsilon \approx 2\pi\nu\sqrt{2mk_BT}. \tag{2.11}$$

If we choose (as an example) a frequency of roughly $k_BT/\hbar$, a mass equivalent to that of Carbon (ie. 12 amu), and a temperature of 310 K, this leads to a tolerance value of $\epsilon \approx 3.9 \times 10^{-11}$ N. In GROMACS, forces are expressed in units of kJ mol^{-1} nm^{-1}, so this leads to a tolerance value of $\epsilon \approx 24$ kJ mol^{-1} nm^{-1}. In the preparation for the simulations described in the next chapter a tolerance value of 100 kJ mol^{-1} nm^{-1} is used in the energy minimization sequences, and this would seem to be reasonable from the above considerations.

2.2   Force-fields

Many different force-fields exist for the GROMACS package–a couple of examples are GROMOS-96 [40] and OPLS (Optimized Potential for Liquid Simulations) [41, 42]. Each has it's own useful function in different contexts, though more focus will be laid on the OPLS force field since this is the force-field chosen for this work. In some cases there are also different versions of these force fields, and typically they are classified as being coarse-grained, united atom, or all-atom. Coarse-grained force fields tend to treat specific groups of atoms as single units, and usually these groups are defined in terms of properties such as polarity. Coarse-grained force fields run much faster in comparison to the other types. United atom force fields integrate the hydrogens into the atoms to

which they are attached, and, while they are not as fast as the coarse-grained force fields, they are faster than the all-atom force field. The all-atom force field involves the interaction between each atom in the system. Although this leads to slower simulations, they are also in a sense more rigorous.

The potentials giving rise to the force-fields can be divided between non-bonded potentials, bonded potentials, and restraining potentials (or simply referred to as restraints). The non-bonded potentials act over a range of distances and includes interactions such as the Coulomb interaction and Van der Waals interaction. The bonded potentials are the potentials that involve specifically the interactions between atoms that are part of the same bond or which exist in nearest-neighbour bonds. Restraining potentials are imposed in a variety of contexts, such as in the case where one wants to keep one molecule fixed in place while allowing the rest of the system to evolve in the usual manner.

The OPLS force-field in particular includes the Coulomb potential and Lennard-Jones potential for non-bonded potentials. For bonded potentials, it consists of the harmonic distance and angle potentials as well as the Ryckaert-Bellemans dihedral angle potential. These will be described in more detail in the following subsections.

2.2.1 Non-bonded potentials (OPLS)

The OPLS force field, as mentioned above, includes both the Coulomb interaction and Lennard-Jones interaction for non-bonded potentials. The effects of the Coulomb interaction, $kq_1q_2/|\vec{r}_{12}|$ (where k is approximately $9 \times 10^9 \frac{\mathrm{Nm}^2}{\mathrm{C}^2}$, q_1 and q_2 are the charges on two separate atoms, and $|\vec{r}_{12}|$ is the distance between them), can be calculated readily at very short distances. Although determining the force due to this potential is straightforward for a pair of atoms, it becomes cumbersome if one considers the interactions between all pairs of atoms in the system as well as between atoms in different image boxes. For this reason it is useful to impose a cut-off radius R_c. For interatomic distances that are less than this cut-off radius the Coulomb interaction is calculated by its normal formula and is incorporated as a contribution to the overall force in the leap-frog algorithm. Outside of the cut-off radius a different algorithm is used to determine the effects of the Coulomb interaction.

One algorithm for performing this kind of task is the Ewald summation method [43], which essentially Fourier transforms the radial dependence of the Coulomb interactions (for atoms that are separated by distances beyond the cut-off radius), and then inverse Fourier transforms the final sum. In other words, the total potential becomes

$$V_C = \frac{k}{2} \sum_{i,j}^{N} \frac{q_i q_j}{|\vec{r}_{ij}|}$$

26

$$= \frac{k}{2\pi\Omega} \sum_{i,j}^{N} q_i q_j \sum_{m_x,m_y,m_z} \frac{\exp\left(-\left(\frac{\pi\vec{m}}{\beta}\right)^2 + 2\pi i\vec{m}\cdot(\vec{r}_i - \vec{r}_j)\right)}{m^2} \tag{2.12}$$

where $\vec{m}$ refers to a reciprocal lattice vector and Ω is the unit volume of Fourier space. Here β determines at which point the reciprocal space terms become more significant (and hence is a function of R_c).

In practice even the Ewald method is not quite as efficient as might be desired. In this case other methods such as the Particle Mesh Ewald method can be used [44, 45]. This is based on the Ewald method, but it has a special means of reducing the number of reciprocal space terms that need to be included in the final sum.

The other non-bonded potential is the Lennard-Jones potential. This has the form

$$V_{LJ}(r_{ij}) = 4\epsilon_{ij}\left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^6\right). \tag{2.13}$$

Here the parameters ϵ_{ij} and σ_{ij} must be provided for different pairs of atom types from experimental data–this information is stored as data in the topology files. The first term in the interaction is a repulsive term that arises from the Pauli exclusion principle and acts over very short distances. The second term is an attractive term that arises from the Van der Waals interaction. The sum of these terms leads to a potential of the form depicted in Fig. 2.1, with $\epsilon_{ij} = 0.92621$ kJ/mol and $\sigma_{ij} = 0.29126$ nm.
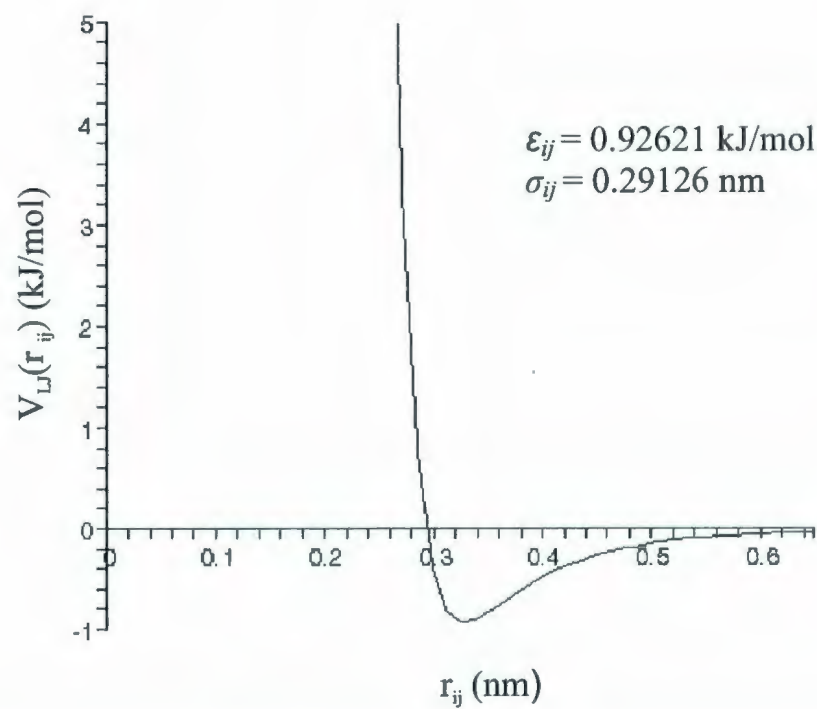
Figure 2.1: Lennard-Jones interaction potential versus radial distance between oxygens in -COOH molecular group.

2.2.2 Bonded potentials (OPLS)

Bonded potentials generally deal with covalent bonds and the means by which atoms in nearest neighbour covalent bonds interact with each other. Generally the length of a covalent bond is stable (since chemical reactions are not to be dealt with here) as well as the angle between nearest neighbour covalent bonds. This leads naturally to the use of harmonic potentials. For example, the harmonic distance potential keeps the bond length stable:

$$V_d = \frac{1}{2} C \left(r - b^0 \right)^2 \tag{2.14}$$

where b^0 is the equilibrium bond length, r is the bond length at a particular time, and the spring constant C must be provided. An illustration of the harmonic distance potential is provided in Fig. 2.2 along with an example graph. Similarly the harmonic angle potential is

$$V_\theta = \frac{1}{2} D \left(\theta - \theta^0 \right)^2. \tag{2.15}$$

Here θ_0 is the equilbrium angle separation, θ is the angular separation at a particular time, and again D is a spring constant that must be provided. This is illustrated in Fig. 2.3.

Finally, there is an interaction between bonded atoms according to their dihedral angle separation. The dihedral angle ϕ is the angle between planes which each contain a pair of adjacent bonds. This is made more clear in Fig.
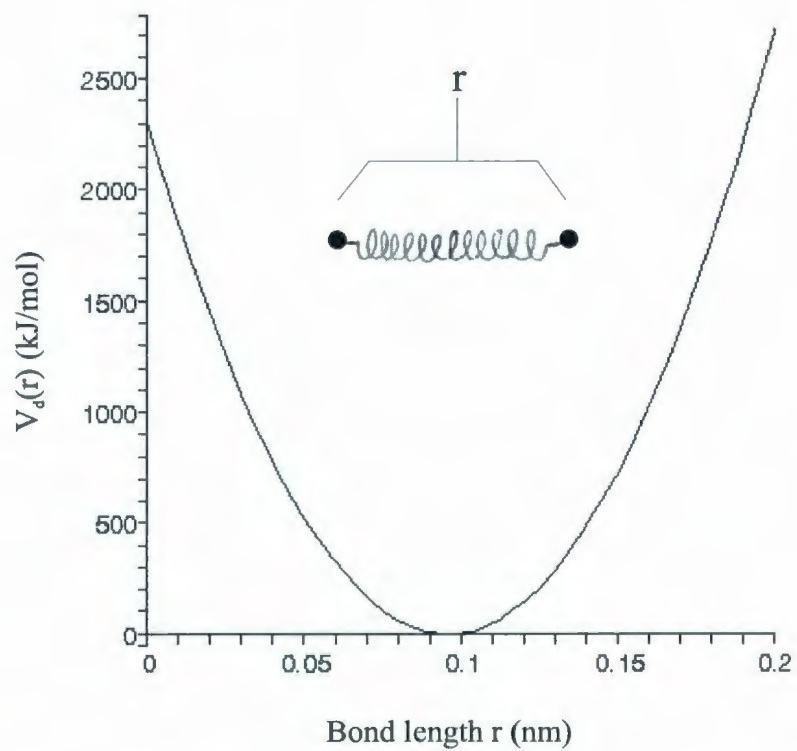
29

Figure 2.2: The harmonic distance potential.

2.4.

The potential for the dihedral angle separation, referred to as the Ryckaert-Bellemans potential [46], has the following form:

$$V_\phi = \sum_{n=0}^{5} C_n (\cos(\phi))^n \qquad (2.16)$$

where the C_n are parameters to be provided in the topology. A plot showing

30
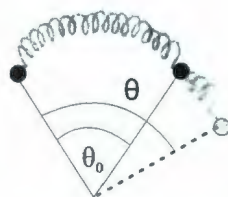
Figure 2.3: The harmonic angle potential.
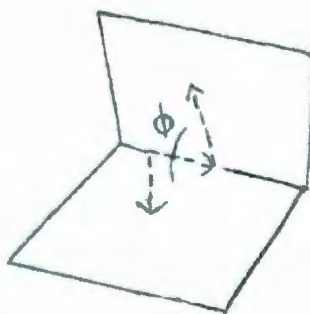
Figure 2.4: Illustration of dihedral angle.

an example of the Ryckaert-Bellemans potential as a function of ϕ is given below in Fig. 2.5. These parameters are borrowed from the Berger lipid parameter set [47], and in this case $C_0 = 9.2789$ kJ/mol, $C_1 = 12.156$ kJ/mol, $C_2 = -13.120$ kJ/mol, $C_3 = -3.0597$ kJ/mol, $C_4 = 26.240$ kJ/mol, and $C_5 = -31.495$ kJ/mol.

In general the Ryckaert-Bellemans dihedral potential is designed in a way so as to keep hydrocarbon chains in a more planar configuration.
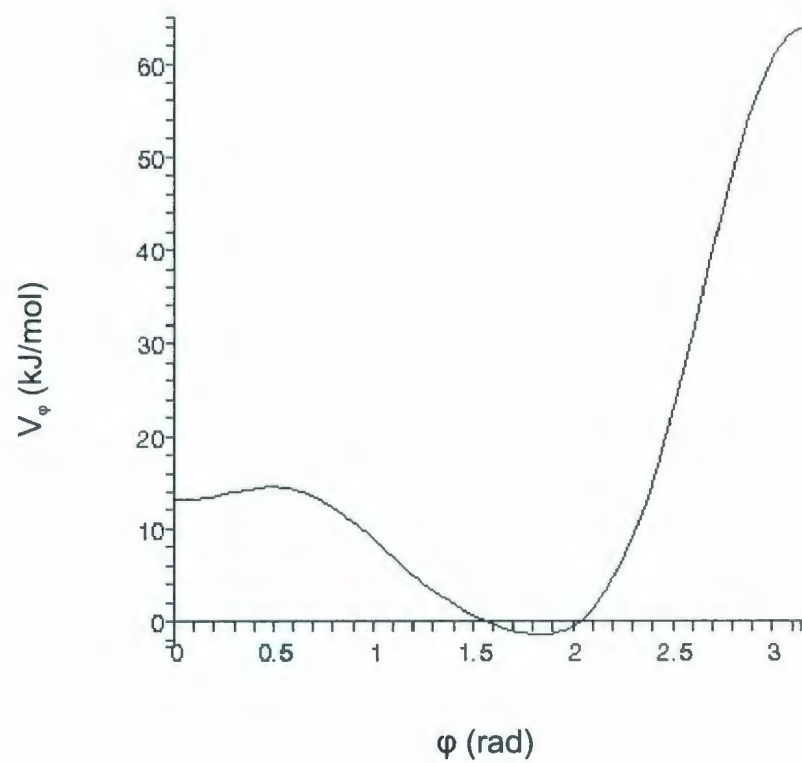
Figure 2.5: Ryckaert-Bellemans potential for H2-H2 dihedral angle.

2.2.3 Restraining potentials

Finally, restraining potentials may be used when a particular desired configuration of a molecule or group of molecules needs to be kept intact throughout the simulation. In the case of position restraints, a harmonic potential (with a given force constant) is applied to each atom so that they are relatively fixed in place. As mentioned earlier, position restraints may frequently be applied to a peptide or protein to keep it stationary while allowing the remainder of the system to relax to a more stable configuration. Restraints can be applied to other properties as well such as the distances between atoms, the angles between different bonds, etc.

Chapter 3

Method of MD Simulations

The method for simulating the dynamics of Mini-B in monolayers with DPPC and/or POPG molecules is discussed in the following. The monolayers were originally obtained from previously equilibrated bilayer structure files from the websites of Peter Tieleman (for DPPC) [48] and Mikko Kartunnen (for POPG) [49] by Doyle Rose [12], who then removed one leaflet of each bilayer to obtain the monolayers. The number of lipids in these monolayers were then doubled on each side in the X and Y directions and energy minimized by Jennifer Rendell and John Bartlett. Here the positive z-axis is assumed to point from the lipid tailgroups towards the headgroups. The DPPC monolayer for these simulations was originally close to 70 Å^2 per lipid, and the POPG monolayer was close to 60 Å^2 per lipid. In the following, the POPG molecules consist of roughly equal amounts of molecules of opposite chiralities (though the effects

of chirality is mostly ignored in the subsequent analysis of the simulations).

The force-field used for these simulations, as mentioned in Chapter 2, was based primarily on the OPLS all-atom force-field. It included a different set of parameters, adapted for the OPLS all-atom force-field, for the lipids that were derived from the Berger lipid parameter set [50, 47]. This was obtained from Luca Monticelli, who was then a member of Peter Tieleman's group, along with a molecular topology for DPPC also adapted for the OPLS all-atom force-field in 2006. The topology files for POPG adapted for the OPLS all-atom force-field were also obtained from Luca Monticelli in 2007. The topologies for POPG were modified by Jennifer Rendell for certain hydrogen atoms by changing the names of the atom type while keeping the charge and Lennard-Jones $\epsilon - \sigma$ parameters (see Subsection 2.2.1) the same in order that the compiler for the GROMACS simulation program would run properly.

A mixed monolayer of DPPC and POPG in a molecular ratio of 7:3 was prepared by looking at the coordinates of the molecules in the DPPC monolayer row by row and replacing three molecules of DPPC after each set of seven with POPG molecules to obtain the appropriate molecular ratio. Although the placement of the POPG molecules in this manner was not random, it was expected that the lipid molecules should be able to diffuse rapidly enough in the course of a 20 - 40 ns simulation that this initial periodic arrangement of the POPG molecules should not have any affect on the results. The mixed monolayer and the monolayer of pure POPG were used in the subsequent sim-

ulations. It was ensured that the monolayers were prepared at 60 Å^2 per lipid. The desired cross-sectional area per lipid was achieved by scaling the monolayer up or down in area as needed in small steps uniformly in the X and Y directions, with an energy minimization following each step. The idea for this procedure was borrowed from an article by Peter Tieleman and co-workers [51].

The DPPC:POPG monolayer that was used had 117 lipid molecules, while the POPG monolayers had 118 lipid molecules. 5555 solvent molecules were added to the DPPC:POPG monolayer, while 3047 solvent molecules were added to the POPG monolayer. Sodium ions were added to the POPG monolayers to balance the charge of the POPG headgroups, which is -1 for each POPG molecule. Energy minimizations were performed after the addition of the solvent and also after the addition of the sodium ions.

Energy minimization itself was performed in steps with the steepest descents algorithm by starting with a large tolerance value of 200000 kJ/mol·nm. This tolerance value is gradually reduced (generally by one-half at each step) until a final tolerance value of 100 kJ/mol · nm is reached (this is in accord with the general considerations described in Section 2.1). The steepest descents algorithm may be inadequate for reducing the tolerance value at some steps, particularly when the tolerance value approaches 100 kJ/mol · nm, and so the conjugate gradients algorithm is used instead. A script program written in Perl, used by the author, for executing a combination of steepest descents

and conjugate gradients algorithms in this manner was prepared by Jennifer Rendell.

VMD graphics of the initial setup involving a DPPC:POPG monolayer with solvent and sodium ions are provided below in Fig. 3.1, though the sodium ions are not made visible. A color-coding scheme, which is in addition to the color-coding scheme provided in Fig. 1.4 for the peptide Mini-B, is given at the top of the graphic. Although the L-POPG molecules are colored green here and the D-POPG molecules are colored red, both chiralities will be given the same color red in subsequent graphics since the effects of chirality are ignored.

DPPC = cyan, L-POPG = green, D-POPG = red,
solvent = grey. In subsequent graphics both D-POPG and
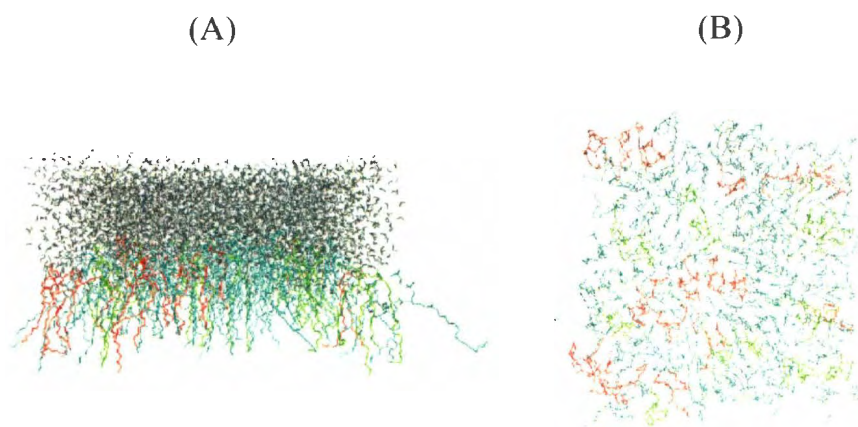L-POPG molecules are colored red.

(A) **(B)**

Figure 3.1: DPPC:POPG monolayer for (A) side view with solvent and (B) top view with solvent removed.

The monolayers which started at a cross-sectional area of 60 Å^2 per lipid were then run through a 10 ps "position-restrained" MD run, though in the absence of a peptide no molecule actually had any restraints placed on it. Instead, initial velocities were generated from a Maxwellian distribution of ideal gas molecules at 310 K during this stage of the procedure. Although the atoms of the system were not actually part of an ideal gas, they were treated as ideal gas constituents in assigning initial velocities. The monolayers were then simulated until the total energies of the systems reached a value that seemed to be more or less constant within the fluctuations that were present. This was one way of attempting to confirm that the system had come to thermal equilbrium, though it is by no means completely reliable because the energy may remain constant only for an undetermined amount of time before decreasing again. Further considerations in regard to whether or not the system reached an equilibrium or steady state configuration are discussed below. The monolayers were then scaled down in small steps to obtain a cross-sectional area of 55 Å^2 per lipid using a procedure similar to that as described above for the initial preparation of the monolayers. The monolayers at 55 Å^2 were then also simulated.

These monolayers were simulated on ACE-net and Westgrid clusters. The clusters are described in further detail below. The DPPC:POPG monolayers were simulated for 30 ns for the monolayer at 55 Å^2 per lipid and 20 ns for the monolayer at 60 Å^2 per lipid. The POPG monolayers were simulated for

approximately 65 ns for the monolayer at 55 Å^2 per lipid and 60 ns for the monolayer at 60 Å^2 per lipid.

The MD parameters that were used for the simulations are provided in Table 3.1. These parameters were also used for the simulations with Mini-B below. Although there are small differences between the parameters used for energy minimization, position restraints, and the final MD run the results should not be much affected due to the brevity of the first two steps. Small differences in cut-off values may not have much effect on the results anyway, but it is particularly expected that a small difference in cut-off values will not affect the final results very much for the energy minimization and position restraints steps because they are short in duration.

The temperature was kept fixed throughout the simulations. As mentioned previously in Chapter 2, no pressure coupling was applied in these simulations, but the volume of the system was kept constant instead. The total number of molecules was also fixed, so this corresponds to an NVT ensemble [52]. With regard to the value of τ used, it was reasonably short enough that temperature deviations would decay rapidly according to equation 2.3 relative to the timescale of the simulations.

For long-range electrostatics, the Particle Mesh Ewald algorithm was used. The parameters used during position restrained MD run were essentially identical to those used for the full-scale MD runs except for the restraining potential. In this case the restraining potential had a force constant of 1000

Table 3.1: Summary of some MD parameters used in the simulations

	Energy minimization parameters	Mdrun parameters
Step size	0.002 ps	0.002 ps
Neighbour list cutoff radious	1.25 nm	0.9 nm
Coulomb cut-off radius	1.0 nm	0.9 nm
Van der Waals cut-off radius	1.0 nm	1.4 nm
Temperature coupling type	N/A	Berendsen
Temperature	N/A	310 K
Temperature coupling time constant	N/A	0.1 ps

$kJ \cdot mol^{-1} \cdot nm^{-2}$.

Once the monolayers were prepared in different compositions, the next step was to place Mini-B inside the monolayers. The structure of Mini-B was obtained from the Protein Data Bank in .pdb format (PDB ID: 2DWF) [31], and then converted to a Gromacs structure file. VMD graphics of the structure by itself are provided in Fig. 1.4. In order to place the peptide inside the monolayer, a hole was made first. This is to facilitate the interaction between the lipid chains of the monolayer and the hydrophobic residues of

Mini-B, which would be expected to draw the peptide into the interior of the monolayer. Fig. 3.2 below illustrates how the hole appears, as well as how the peptide is placed relative to it. Note that when a hole is created and Mini-B is placed in the monolayer, the effective cross-sectional area per lipid (ie. the surface area of the monolayer in a single box divided by the total number of lipids) increases. However in making reference to the systems in subsequent discussions, the original cross-sectional area of 55 Å^2 per lipid or 60 Å^2 per lipid will be mentioned as opposed to its actual effective cross-sectional area per lipid.

In Fig. 3.2, Mini-B is reduced in size in order to minimize the effects of close contacts between the peptide and the lipids. This is the general procedure that was used for all simulations with Mini-B. The peptide had been shrunk uniformly in the X and Y directions so it was 10% smaller in each direction, ie. 19% smaller in cross-sectional area overall. Without doing this first, it is possible that close contacts between peptide molecules and lipid molecules could have lead to very strong interactions between them. These strong interactions could have ended up disrupting the structure of the peptide. Although using this procedure would initially shrink the distances between atoms of the peptide, it was expected that the peptide would relax back to a larger conformation during the course of the simulation.

Once the hole is created, the peptide is placed at different starting depths of 12 Å, 16 Å, and 20 Å with respect to the top of the monolayer. The
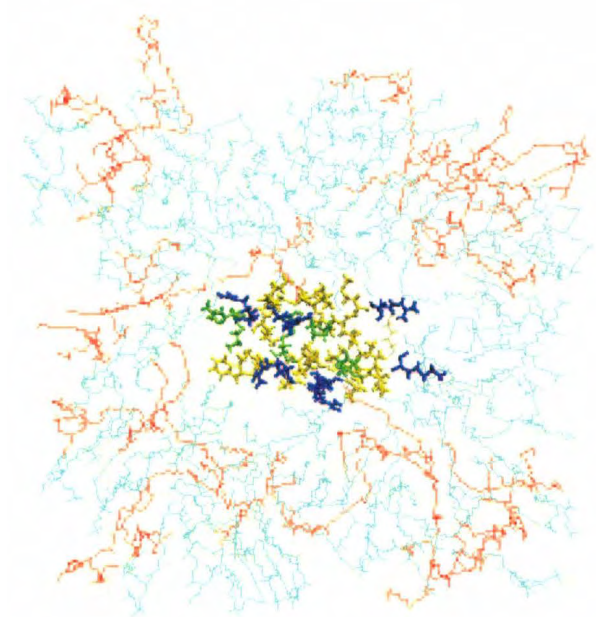
41

Figure 3.2: Mini-B is placed in hole in center of monolayer.

initial depth in this case was measured by determining the difference in vertical height of the top-most atom of the monolayer and the bottom-most atom of the peptide. An alternative measure for the initial depth is used in Section 4.1, but it will still be useful to refer to the initial depths according to this first measure.

An additional parameter that was varied in these simulations was the tilt angle of the helical segments of Mini-B relative to the interface. This is typically measured with respect to the normal to the surface of the monolayer,

and is a parameter that is frequently studied with respect to other fragments of SP-B [53, 54]. In these simulations starting tilt angles of 90° and 45° with respect to the normal were used, as illustrated in Fig. 3.3. These values for the initial tilt angles are approximate, and a more rigorous measure for the tilt angle will be provided in Section 4.3. However, it still useful to refer to the initial tilt angles as being either 90° and 45°, and so these will be referred to as the "nominal" initial tilt angles.
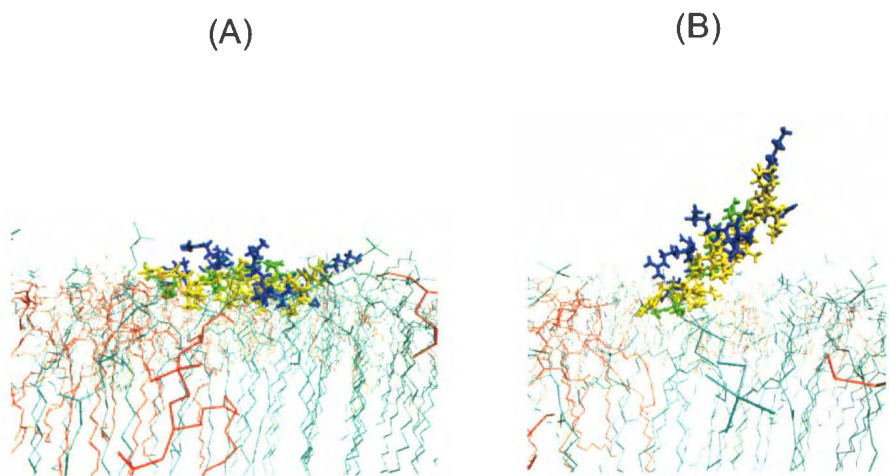
(A)   (B)

Figure 3.3: Illustration of starting tilt angles of (A) 90° and (B) 45°.

Once the systems were set up with Mini-B at different depths and tilt angles, energy minimization was performed. Roughly 5000-6000 water molecules were added to the systems, followed by another energy minimization. Sodium ions were then added to balance the charge of the POPG headgroups, followed by a third energy minimization. Position restrained MD runs, with position

restraints on the peptide, were then performed for 10 ps. Finally, these simulations were submitted in parallel to clusters on the ACE-net and Westgrid networks. No further restraints were placed on the peptide during the actual simulations.

The clusters that were used on ACE-net were Mahone2 [55] and Placentia2 [56] (these are now referred to simply as Mahone and Placentia). Mahone2 is a cluster of 139 nodes with 2 processors per node and 2 CPU cores per processor for a total of 568 cores. Most of the nodes have 16 GB of RAM memory per node, though 16 nodes have 64 GB of RAM per node. It uses Myrinet type interconnects, and its processors use the Linux AS4 operating system delivered by Red Hat Enterprise. Placentia2 is a cluster of 74 nodes with different numbers of CPU cores assigned to different sets of nodes. It also has a total of 568 cores. Some nodes use 16 GB of RAM memory per node, while others use 64 GB per node. The cluster uses a Gigabit Ethernet interconnect, and its processors also use the Linux AS4 operating system delivered by Red Hat Enterprise.

In most cases, 8 processing nodes were used on the Mahone2 and Placentia2 clusters. For a 40 ns MD run, a job on either of the ACE-net clusters usually took 6 to 10 days, which includes the waiting time between when the job was submitted and when the job actually started running. In some cases this waiting time could be neglible while in other cases it could be a few days or longer, depending on how many users had jobs on the cluster. Once the job

was submitted, it would usually run continuously until finished.

The cluster used on the Westgrid network was Matrix [57]. The Matrix cluster consists of 128 nodes and has 2 processors per node for a total of 256 processors [58, 59]. 2 GB of RAM memory are used by the processors on each node. The processors, which are of the AMD Opteron brand, also use the Linux operating system. Infiniband and Gigabit Ethernet interconnects are used in between the nodes.

In most cases the MD run jobs on Matrix were run on 8 processors. One limit to the Matrix cluster is that it had a walltime of 3 days on the amount of time a job could be run on it. In other words, a job would be terminated if it ran longer then 3 days, and this led to having to break the job up into individual steps that each would take less than 3 days. A script was available for submitting the steps one after the other in sequence without further need for user intervention. Despite the fact that the jobs had to be submitted in steps, a job of 40 ns still only took 6 to 10 days, similar to the time required for jobs on the ACE-net clusters. This may be because waiting times for jobs to start running could be very short. One way of overcoming the walltime limitation in the Matrix cluster is to use a different cluster such as the Glacier cluster, which has no walltime limit on jobs [60]. However, because of the absence of a walltime this cluster may also have many more users competing for processors and therefore waiting times may be longer.

The systems were mainly simulated between 20 - 60 ns in length. One

45

system was simulated for 80 ns in length, and this corresponded to the system for which the sodium ion concentration is varied as discussed below. Several of the systems were simulated for longer than 20 ns because they did not appear to be equilbrated (even some which had been run for 40 ns or 60 ns did not appear to be equilibrated).

One indicator that was used to determine whether or not the system had achieved an equilibrium configuration was the total energy of the system. This should approach a constant value, within thermal fluctuations, near equilibrium. A typical plot of the total energy is given in Fig. 3.4.

Initially the total energy decays with time due to its coupling to the constant temperature bath. In the absence of temperature coupling, energy is conserved according to Newtonian dynamics. In the presence of temperature coupling, the system relaxes to a more stable configuration from an initially more unstable configuration (unstable in the sense that it is higher in energy). The effect is similar to that of the dissipation of heat, though no heat is actually transferred to any substance within the system, and the origin of the effect is artificial. Due to the initial decay in total energy, the average total energy of the system was obtained normally by averaging over the last 10 to 20 ns of the trajectory.

Another means of confirming that the system had reached an equilibrium state was to look at the pressure components in the X and Y directions. These should have very similar values. However, the interpretation is questionable
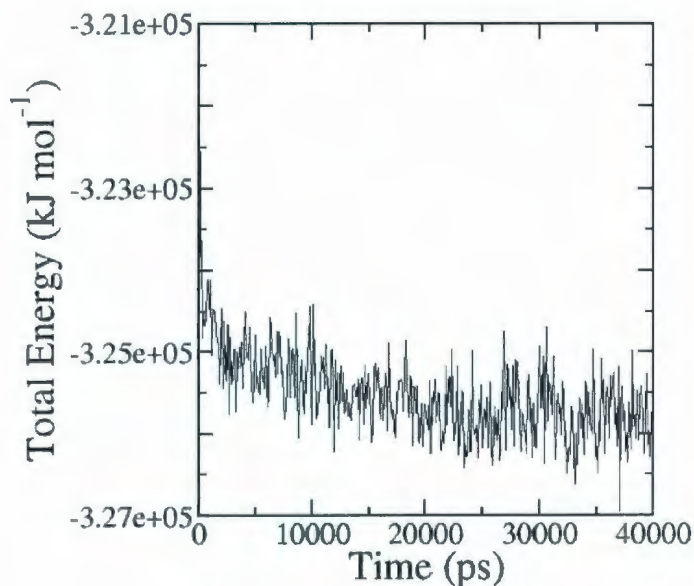
46

Figure 3.4: Total energy versus time.

as the pressure often tended to fluctuate wildly in these simulations. This is illustrated in Fig. 3.5 below. The pressure in the X direction in Fig. 3.5(A) has average -53.8 Bar with fluctuations of ±45.8 Bar. The pressure in the Y direction in Fig. 3.5(B) has average -56.0 Bar with fluctuations of ±49.0 Bar.

After the simulations, an analysis was carried out in which the position, secondary structure, and orientation of the peptide were determined for the above simulations. This will be discussed in further detail within the next chapter. In addition, the effects of sodium ions on the interaction between the
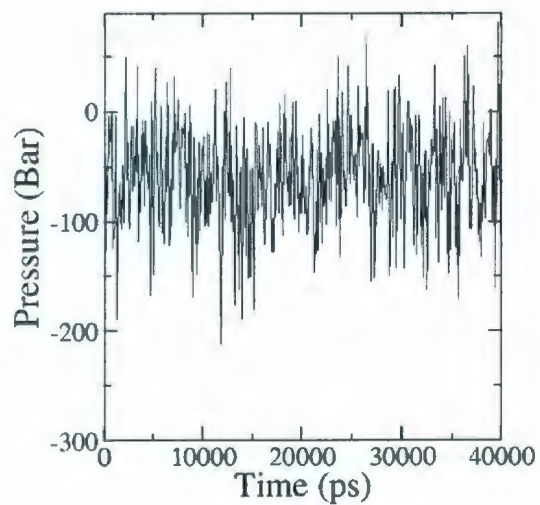
47
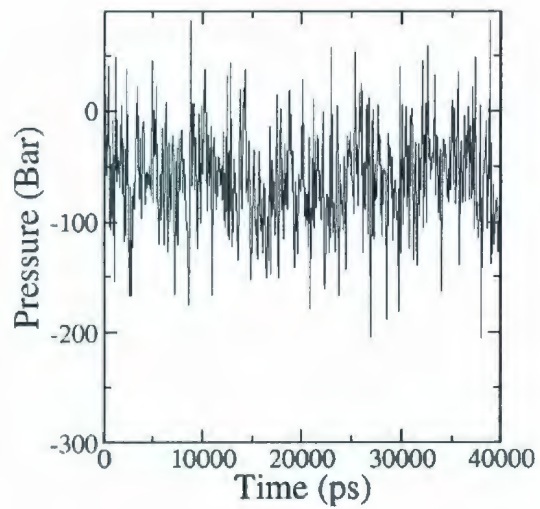
Figure 3.5: Typical plots of the pressure in the (A) X direction and (B) Y direction versus time.

cationic residues of Mini-B and the anionic POPG headgroups was studied for the system with a DPPC:POPG monolayer with original monolayer at 60 Å^2 per lipid and with the peptide initially at 12 Å depth and 45° tilt. The sodium ion concentration was varied between roughly 50% and roughly 150% of the original concentration that was used to keep the system charge neutral. The sodium ion concentration was modified from that of the original system after the first 40 ns of its trajectory, and then each individual system was simulated for a further 40 ns. The original system (ie. which had 100% concentration) was simulated up to 80 ns.

The effects of varying the sodium ion concentration were determined by using radial distribution functions to quantify the proximity of the positively charged Arginine and Lysine residues with respect to the POPG headgroups. The radial distribution function of molecules of type mol1 with respect to molecules of type mol2 is as follows [61]:

$$g(r) = \frac{1}{\rho N_{\text{mol1}}} \sum_{i=1}^{N_{\text{mol1}}} \sum_{j=1}^{N_{\text{mol2}}} \frac{\delta(r_{ij} - r)}{4\pi r^2}. \tag{3.1}$$

Here N_{mol1} is the number of molecules of type mol1, N_{mol2} is the number of molecules of type mol2, r is the radial distance, r_{ij} is the radial distance between the ith molecule of type mol1 and the jth molecule of type mol2 and ρ is a normalization constant. These radial distribution functions will also be discussed further in chapter 4.

Chapter 4

Results

The results from the simulations described in Chapter 3 were analyzed using a combination of ready made programs that were part of the GROMACS package as well as tools from the Octave software package. The analysis was primarily geared towards characterizing the position, secondary structure, and orientation Mini-B adopted inside the DPPC:POPG and POPG monolayers. In addition, the effects of using different sodium ion concentrations on the interaction between cationic Arginine and Lysine residues of Mini-B and anionic POPG headgroups were studied.

Results showed that in the primary set of simulations the properties of the fragment depended on the initial position and orientation of the fragment. Nonetheless, a number of similarities were also found in the simulations. Varying the initial conditions is important because it can enable a better grasp of

what behaviour of the fragment is possible. Ultimately, knowledge of properties such as the position, secondary structure, and orientation of Mini-B is important because it can be used to infer properties of full length SP-B and hence the mechanism of its behaviour.

In the description of properties which were obtained from different simulations, a notation involving a letter followed by two numbers will be used to refer to each of the individual simulations, with exception of the ones which were at different sodium ion concentrations. Simulations which involved the DPPC:POPG monolayer are labelled starting with D, while simulations in the pure POPG monolayers are labelled starting with P. The second character is either 5 or 6 depending on whether the cross-sectional area of the monolayer prior to adding Mini-B was 55 Å^2 per lipid or 60 Å^2 per lipid respectively. The third character is either 1, 2, or 3, depending on whether the initial depth is 12 Å, 16 Å, or 20 Å respectively. Finally, the notation is ended by "-45" or "-90" depending on whether the initial tilt angle is 45° or 90°.

Hence, for example, the system involving a DPPC:POPG monolayer with starting cross-sectional area of 55 Å^2 per lipid, in which Mini-B is at an initial depth of 20 Å and tilt angle of 45° would be labelled D53-45.

A couple of general remarks about the simulations can be made before going into further detail with regard to the results. Analysis of the total energy of the system indicated that some systems were not completely equilibrated. This was mostly the case for the systems with POPG monolayers,

although there was one case, D51-45 (run for 40 ns), where a system with a DPPC:POPG monolayer did not seem to be close enough to a steady state. Other systems that did not seem quite close enough to a steady state were: P51-45 (60ns), P53-45 (40ns), P52-90 (40ns), P62-45 (60ns), P63-45 (60ns), and P63-90 (40ns). The effects of this on the results, however, may not have been very significant. A likely reason that the systems with POPG monolayers required longer equilibration times is that the presence of negatively charged POPG headgroups created a repulsive effect that put limits on the motion of the lipids within the monolayer.

The hole around the peptide, which was initially used to place the peptide in the monolayer, could close rapidly around the peptide. This could occur well before the end of the simulations. This is illustrated below in Fig. 4.1, showing the system at the start and end of the simulation. Here it seems that the peptide was able to incorporate itself within the monolayers without any major gaps between the peptide and the lipids.

4.1   Position of Mini-B within Monolayers

The following discusses the assessment of the final position or depth of Mini-B in the monolayers. This was determined as a function of initial depth and tilt angle. Results are shown for different compositions and cross-sectional areas per lipid of the monolayers.
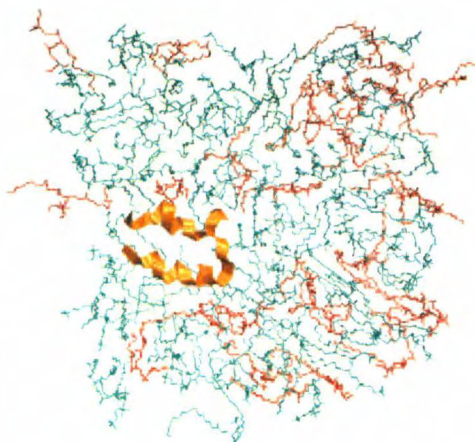
(A)
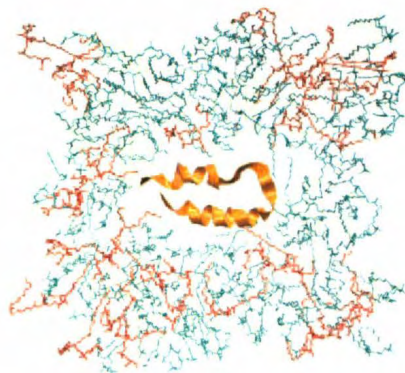
(B)

Figure 4.1: Example graphics illustrating hole closing around peptide for (A) start of simulation and (B) end of simulation.

As mentioned in the previous chapter, the initial depth was first measured for the preparation of the system as the difference in position of the topmost atom of the monolayer and the bottom-most atom of the peptide (Measure 1). This measure functioned well for the preparation of the system, but a different measure was needed to more properly characterize the depth of the peptide. An alternative measure for the depth involved the z-component of the distance vector between the center of mass of the peptide and the center of mass of the phosphorus atoms at the top of the monolayer (Measure 2). This is illustrated in Fig. 4.2 below. Note that it is assumed the positive z-axis points from the monolayer to the water. Along with the determination of the final depth of the peptide, this scheme led to different values for the initial depth of the peptide as well. Examples of initial and final depths using this scheme for some systems starting with labels D6 are given in Table 4.1 below. From here on the last two indices in the notation that is used are removed when referring to all systems of a given composition and cross-sectional area per lipid, eg., the systems starting with labels D6 will simply be referred to as the D6 systems.

In this case the vector joining the center-of-masses of the two molecular groups has its endpoints such that a positive depth implies the center of mass of Mini-B is above that of the phosphorus atoms, while a negative depth implies the center-of-mass of Mini-B is below that of the phosphorus atoms. In the case of Fig. 4.2, Mini-B has a negative depth according to Measure 2. It was not
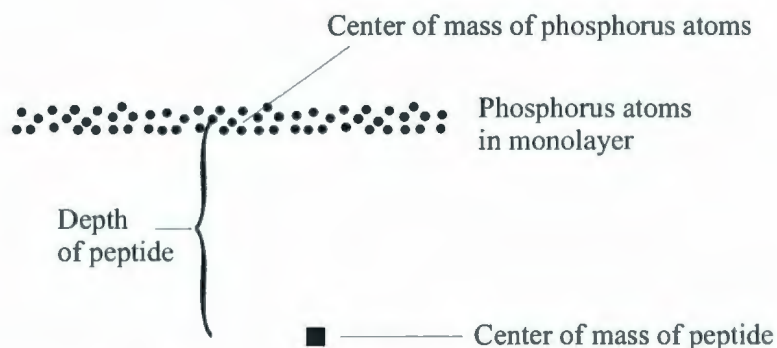
54

Figure 4.2: Illustration of depth of peptide according to Measure 2.

Table 4.1: Depths of Mini-B using two different measures for D6 systems with initial tilt angles of 45 degrees

Initial Depth, Measure 1 (Å)	Initial Depth, Measure 2 (Å)	Final Depth, Measure 2 (Å)
12	+7.9	-2.4 +/- 0.9
16	+3.9	-6.1 +/- 1.1
20	-0.14	-5.5 +/- 0.8

necessary to specify a positive or negative value for the initial depths according to Measure 1 since the bottom of the peptide in these cases is always below the topmost atom of the monolayer. However the convention for Measure 2 will continue to be used since the center of mass of the peptide is sometimes above the center of mass of the phosphorus atoms and at other times below. The final depth was averaged over the last 10 or 20 ns, depending on what seemed appropriate from the energy data. The error associated with the final

55

depths is due to taking into account the RMS deviations of the temperature fluctuations of the depth of the peptide.

It can already be seen from this table that Mini-B tends to attain a greater depth within the monolayers by the end of the simulations. This is what is generally observed in the simulations, that Mini-B either sinks lower within the monolayer or maintains the same depth (this latter case occurs mainly when the peptide is placed initially at a low enough depth). The data which supports these observations are depicted in Figs. 4.3 and 4.4 below. Notice that the error bars in the final depths take into account the temperature fluctuations. Note also that the error bars of systems with similar initial depth according to Measure 1 can overlap, as can be seen for example in Fig. 4.3(A) for the systems with initial depth (Measure 1) of 20 Å. The initial depths according to Measure 2 are also plotted along with the final depths according to Measure 2 in order to compare the two.

Although there is some scatter in some of the plots, it can be seen that there are several cases, particularly at lower starting depths, where there is agreement in the final depth of the peptide for a given composition. By agreement it is meant that the range of final depths implied by the error bars for one particular system overlaps with that of another system. This may be significant because it would be expected experimentally that the peptide have a final depth that is not overly sensitive to initial condititions when the system comes to equilibrium. This in turn suggests that it may be necessary to insert
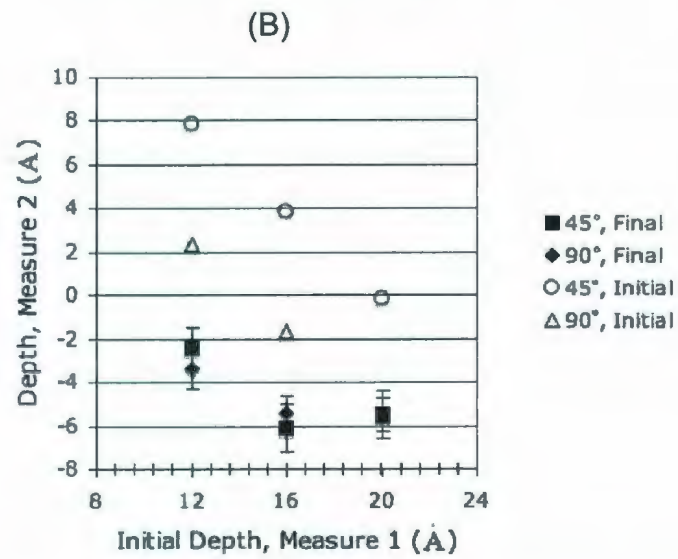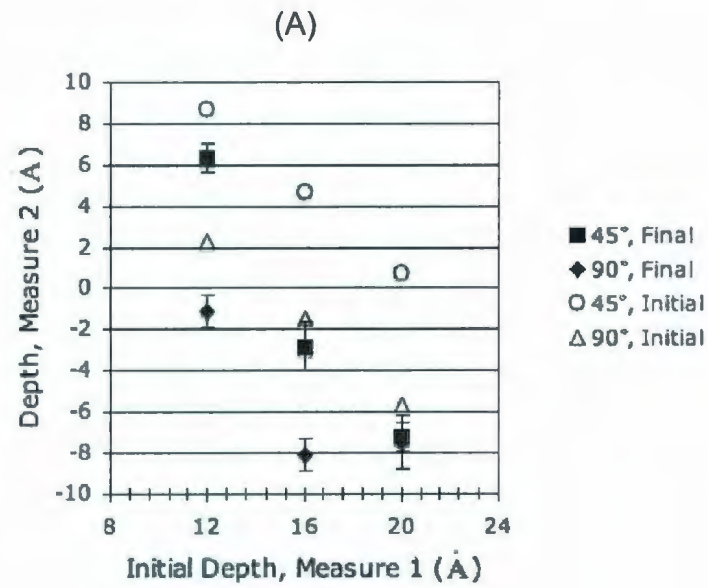
Figure 4.3: Initial and final depths (Measure 2) versus initial depth (Measure 1) for (A) D5 systems and (B) D6 systems.
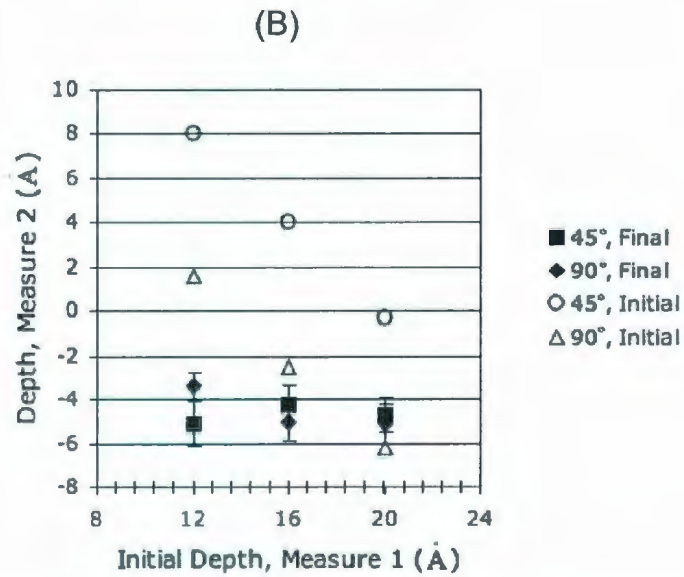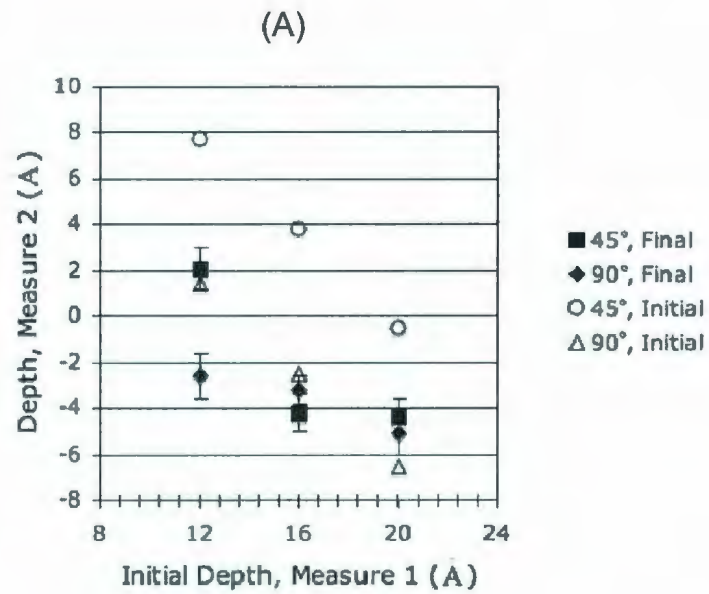
Figure 4.4: Initial and final depths (Measure 2) versus initial depth (Measure 1) for (A) P5 systems and(B) P6 systems.

the peptide at a lower depth of close to 20 Å so that the peptide can attain a depth that is not too sensitive to initial conditions, which may be more consistent with experiment.

An average final depth can be obtained from each of the plots, ie. for a given composition and cross-sectional area per lipid of the monolayer. For the D5 systems the average depth is -3.4 +/- 0.9 Å, and for the D6 systems the average depth is -4.7 +/- 0.9 Å. For the P5 systems the average depth is -2.9 +/- 0.8 Å, and for the P6 systems it is -4.6 +/- 0.9 Å. From this it would seem that the cross-sectional area per lipid has a much greater affect on the depth of the peptide than composition, suggesting that the peptide is to some degree pushed upwards towards the water when the monolayer is compressed by a small amount.

On the other hand, it seems that in most cases the system with initial depth of 12 Å and tilt angle of 45° has a range of final depths that is significantly different from those of the other systems within the same plot. This may have had an effect on the average final depth values obtained above. It is of interest, therefore, to determine alternative values for the average final depth from each plot when the system with this set of values for initial depth and tilt angle is excluded.

Two of these systems with initial depth of 12 Å and tilt angle of 45° did not come to an equilibrium configuration as mentioned in Section 4.1, the D51-45 and P51-45 systems, and hence there is reason for excluding these

systems. The P61-45 system did not have as subtantial a deviation in final depth from the other systems in the same plot, and hence excluding it would not make as much difference in determining the average final depth. Finally, although the D61-45 system did come to a more or less contant energy, it will be excluded to keep consistency in the method of determining an average final depth. It is more difficult to justify excluding the D61-45 system, but it is still possible that this system did not achieve a proper equilibrium configurations and is merely in a local energy minimum. It is still of interest, therefore, to determine alternative values for the average final depth when these systems are excluded.

Hence for the D5 systems omitting D51-45 the average depth is -5.3 +/- 1.0 Å, and for the D6 systems omitting D61-45 the average depth is -5.2 +/- 0.9 Å. For the P5 systems omitting P51-45 the average depth is -3.9 +/- 0.8 Å, and for the P6 systems omitting P61-45 the average depth is -4.5 +/- 0.8 Å. Using this alternative method, it can be seen that for the DPPC:POPG systems the average depth is essentially the same for the different cross-sectional areas per lipid. While the average depths for the POPG systems seem to overlap, the average depth for the P5 systems seems to be shifted somewhat higher in comparison to that of the P6 systems. Furthermore, the ranges of values for the average depth of Mini-B in the POPG monolayers seem to be somewhat shifted higher compared to those of the DPPC:POPG monolayers. These results for the average depth using this alternative method may suggest that

60

Mini-B is attracted more upwards in the POPG monolayers compared to the DPPC:POPG ones. A simple cause for the peptide having a higher range of depths in POPG compared to mixed monolayers may be the Coulombic interaction between the anionic POPG headgroups and the positively charged residues of Mini-B.

Also shown are example graphics of the peptide within the monolayer after having attained its final positions. Fig. 4.5(A) shows a side-view graphic of the peptide in a DPPC:POPG monolayer (D53-90) while Fig. 4.5(B) shows a graphic of the peptide in a POPG monolayer (P62-90). (In both cases the N-TERM helix faces front). These graphics are cases where the peptide lies more or less completely below the top of the monolayer. This situation occurs in the majority of the cases, though there are some simulations where there is a significant part of the peptide that comes above the top of the monolayer. A rough idea of this can be obtained by looking at the mass density profiles, ie. mass density versus distance in the z direction, of the peptide (in this case with the hydrogens removed) with respect to those of the phosphorus atoms. Some examples for the D5 systems are shown below in Figs. 4.6 and 4.7. In Fig. 4.6(A) and Fig. 4.7(A), part of the peptide lies above the phosphorus atoms. In Fig. 4.6(B) and Fig. 4.7(B), the peptide lies entirely below the phosphorus atoms. These mass density profiles are averaged over the last 10 or 20 ns of the trajectories.
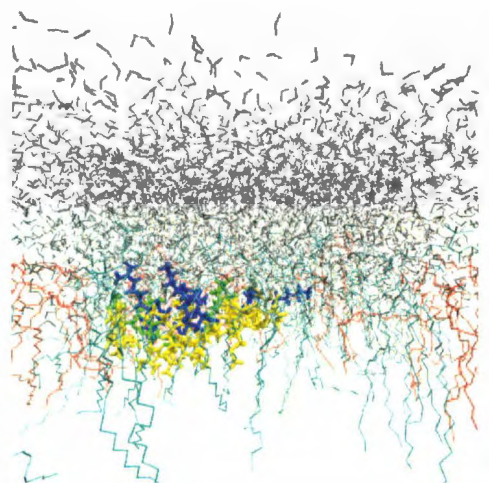
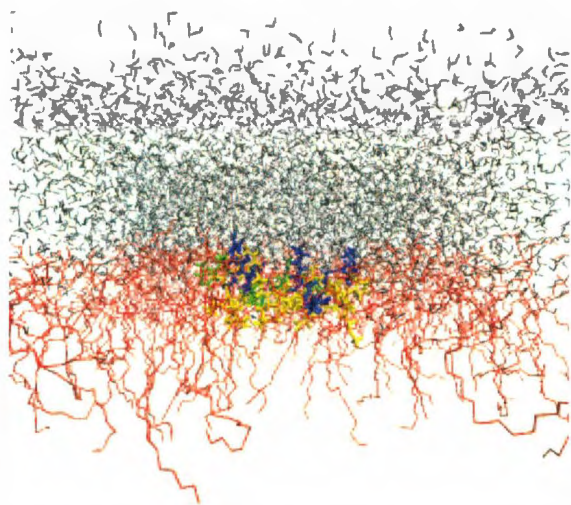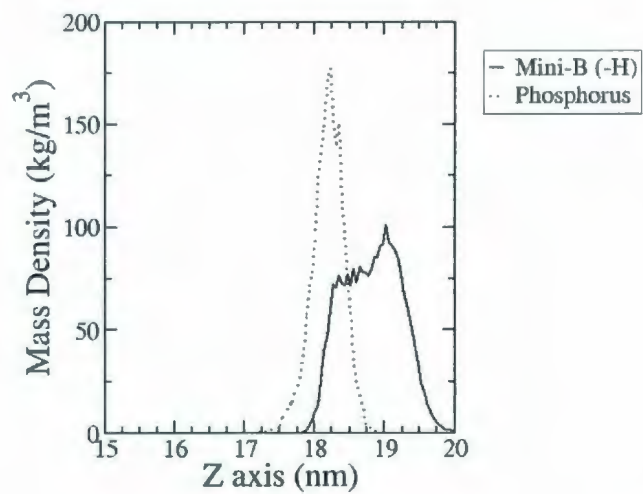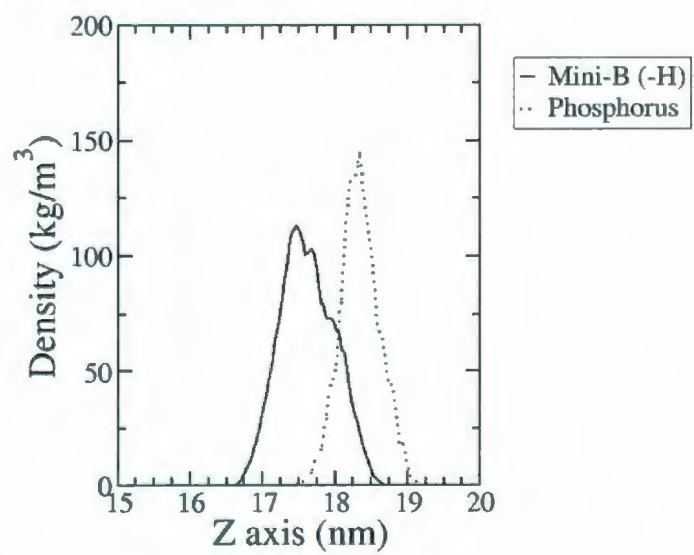Figure 4.5: Graphics of Mini-B in monolayers for (A) D53-90 and (B) P62-90.

Figure 4.6: Mass density profile for (A) the D51-45 system and (B) the D53-45 system.
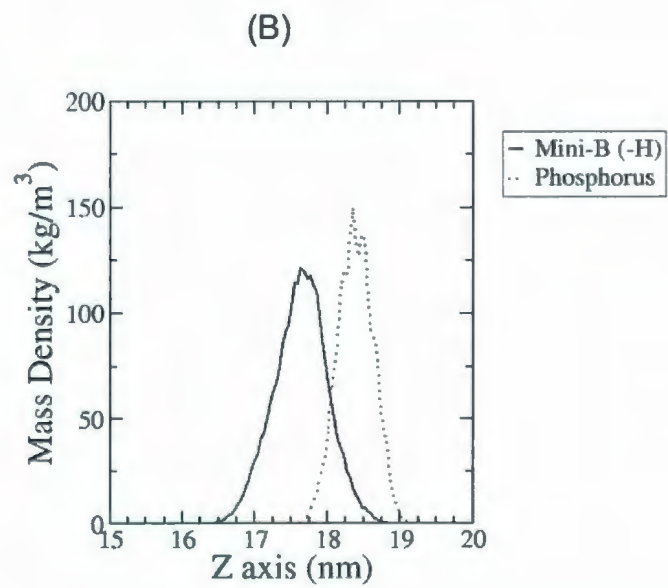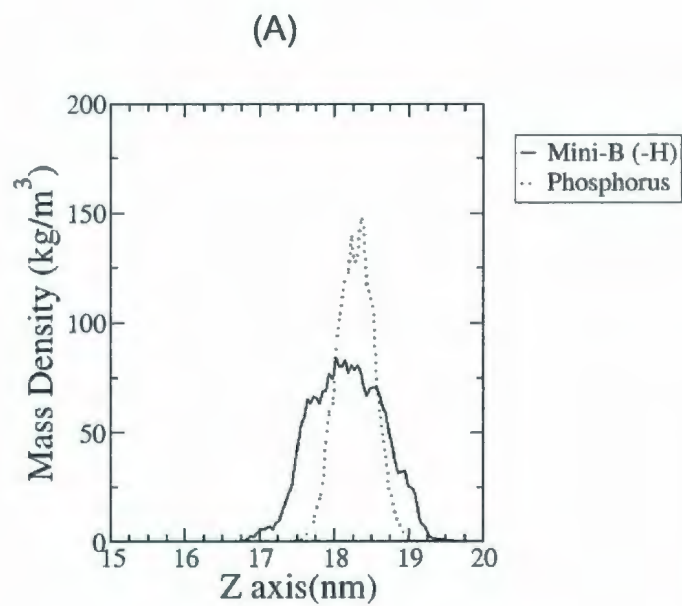
Figure 4.7: Mass density profile for (A) the D51-90 system and (B) the D53-90 system.

4.2 Secondary structure of Mini-B

The following discusses an analysis of the secondary structure of Mini-B. In particular, effort was made to determine whether or not Mini-B kept its alpha helices in the monolayers, as no restraints had been placed on the peptide during the simulations. At first this was done by looking at secondary structure plots of the residues of Mini-B provided by the DSSP program [62] (which comes with the GROMACS package). Some example plots are shown in Fig. 4.8, where residues in blue form part of an alpha helical structure while residues in yellow correspond to ordinary turns.

The results from using this program indicated, however, that in most cases Mini-B lost at least one alpha helix by the end of the simulations. For example, Fig. 4.8(A) shows the system D62-45 in which the N-TERM helix is retained although shortened, while the C-TERM helix dissolves into a 3-10 helix structure. From these plots it seemed the criteria by which the DSSP program determines whether or not a particular residue of the peptide exists as part of an alpha helix might be too restrictive.

The DSSP program primarily analyzes the arrangement of hydrogen bonds of the peptide structure to determine if a residue exists as a turn, and then looks at the arrangement of turns to determine if the residue exists as part of an alpha helix. The criteria used by the DSSP program to determine whether or not a residue exists as part of an alpha helix may be too restrictive in the
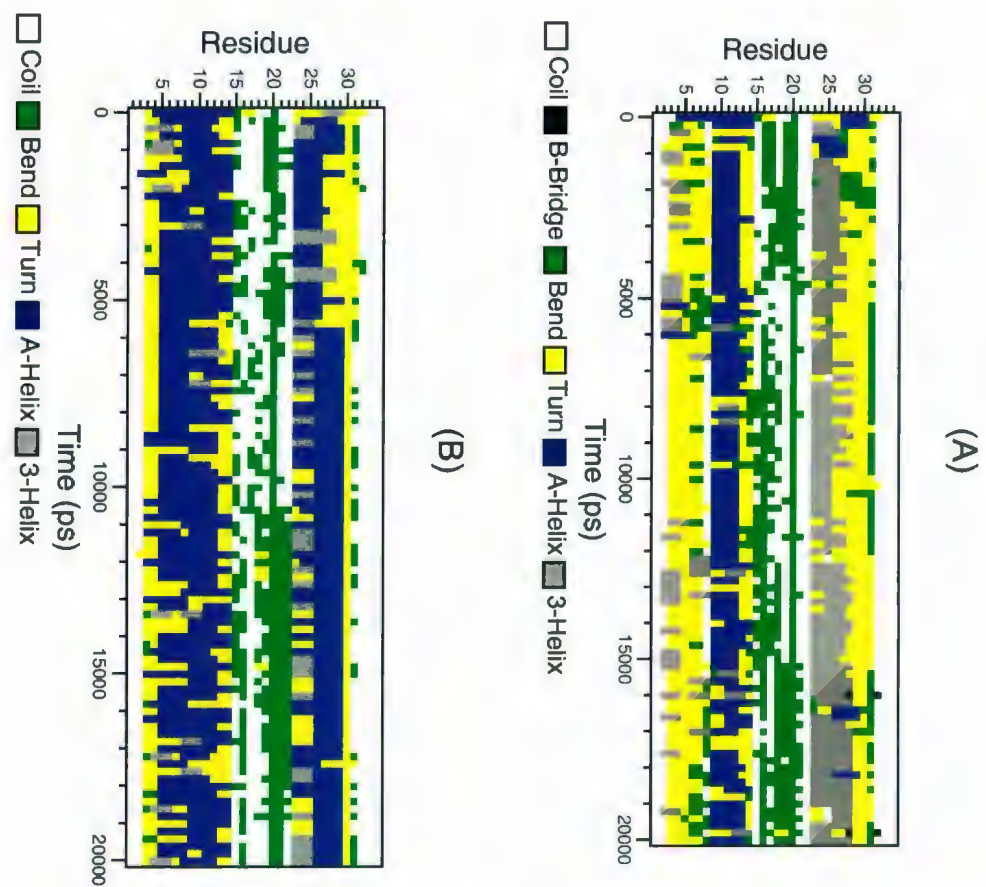
Figure 4.8: Secondary structure plots for (A) D62-45 and (B) D62-90

sense that it requires the alpha helix to be too close to an ideal alpha helx. The alpha helix may undergo changes and distortions from an ideal alpha helical configuration, but it may still be close enough to an alpha helical structure to still be considered alpha helical. Hence an alternative means for quantifying these changes and distortions was used, as discussed below.

One might also assume that the simulations simply do not hold up well with experiment, in which case one would have to conclude that most of the systems did not really attain an equilibrium configuration. However, there was no real reason to suppose that the topologies and MD run parameters chosen for the systems would not yield results that were consistent with experiment as far as something as critical as the secondary structure of the peptide is concerned. Hence it was considered prudent to check to see how the secondary structure of the peptide changed from another standpoint, and so another measure for the change in alpha helical structure was used.

The alternative means for quantifying changes in alpha helical structure was through the root mean square (RMS) deviations of the distances between pairs of backbone atoms in each alpha helix. If x_0^{ij} is the distance between the backbone atoms i and j at the start of the simulation, and x^{ij} is the distance between these atoms at the end of the simulation, then the RMS deviations of the corresponding alpha helix is defined as

$$\left(\frac{1}{N} \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} (x^{ij} - x_0^{ij})^2 \right)^{1/2} \tag{4.1}$$

where N is the number of backbone atoms in the helix.

Plots of the RMS deviations, typically averaged over the last 10 to 20 ns of a simulation (again depending on what seemed reasonable from the energy data) are provided in Figs. 4.9 to 4.12. The average RMS deviations for different systems ranged in value from 0.4 Å to 2.1 Å.

The RMS deviations seemed to show consistency with the secondary structure plots. For example, the secondary structure plot in Fig. 4.8(B) for D62-90 had RMS deviations of on average 0.5 Å for the N-TERM and 1.0 Å for the C-TERM. Also, the system D53-45 had RMS deviations on average of 0.7 Å for the N-TERM and 0.7 Å for the C-TERM and it was observed to retain both alpha helices in the secondary structure plot, though the secondary structure plot is not shown here. There were some cases where an alpha helix underwent RMS deviations on average of up to 1.2 Å and still retained its secondary structure. On the other hand, the system D62-45 in Fig. 4.8(A) had RMS deviations of 1.2 Å for both the N-TERM and C-TERM helix even though the C-TERM helix was lost according to the secondary structure plot. These considerations, as well as observations of other secondary structure plots, suggested that a cut-off of 1.4 Å on the RMS deviations could be used for determining whether on not an alpha helix was retained. If the lower limit of

Figure 4.9: RMS deviations versus initial depth for (A) N-TERM alpha helix and (B) C-TERM alpha helix of the D5 systems.

69

Figure 4.10: RMS deviations versus initial depth for (A) N-TERM alpha helix and (B) C-TERM alpha helix of the D6 systems.

(A)

(B)

Figure 4.11: RMS deviations versus initial depth for (A) N-TERM alpha helix and (B) C-TERM alpha helix of the P5 systems.
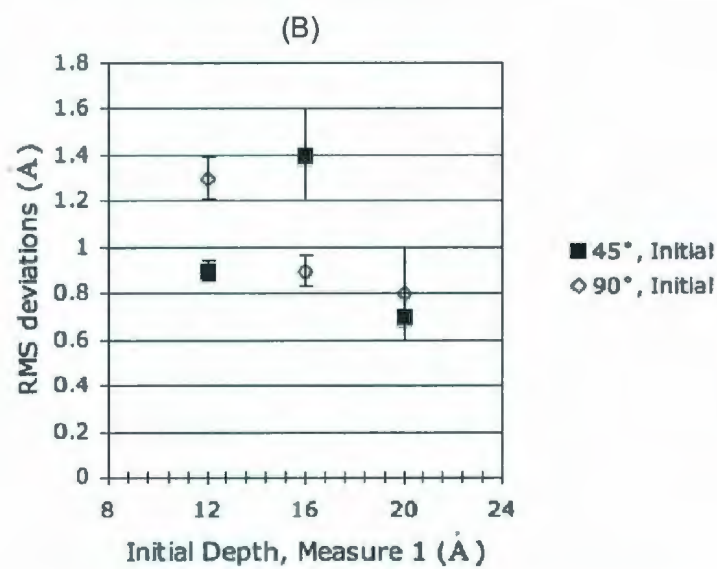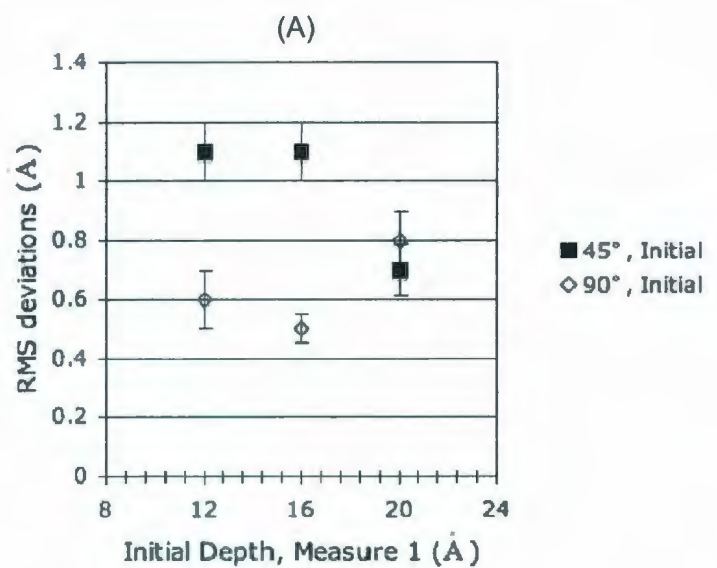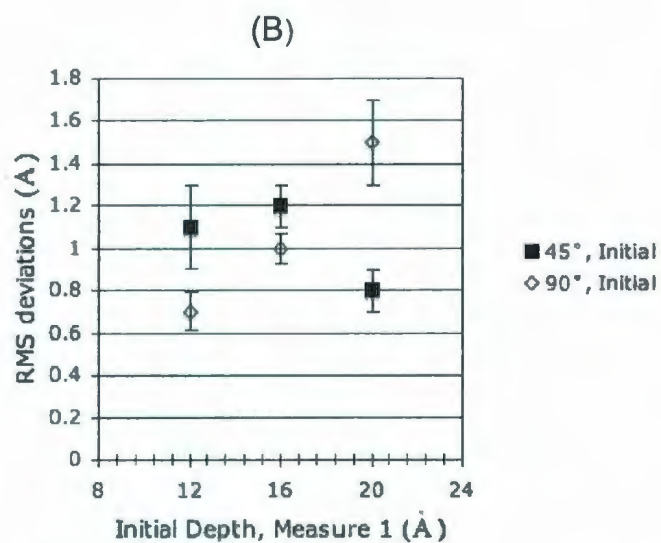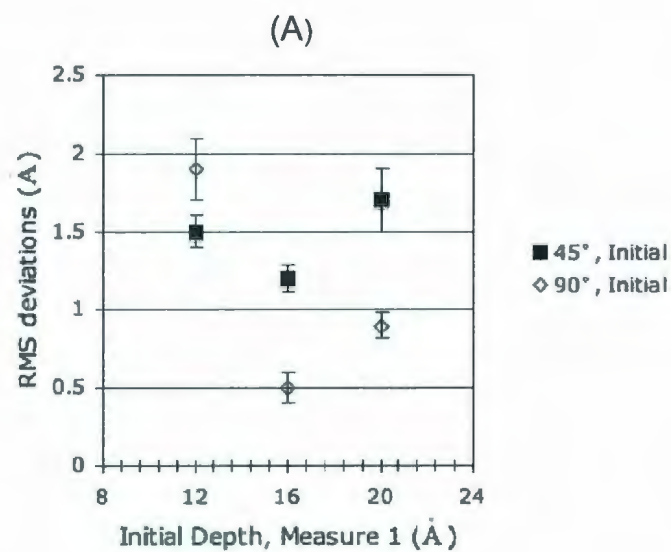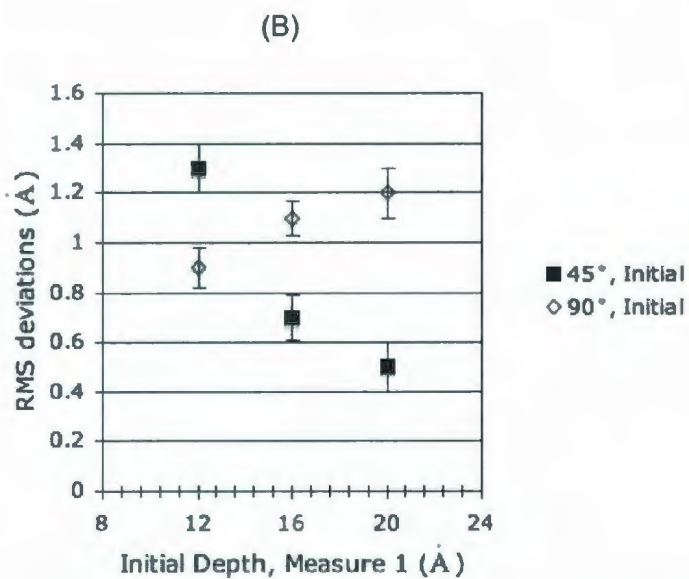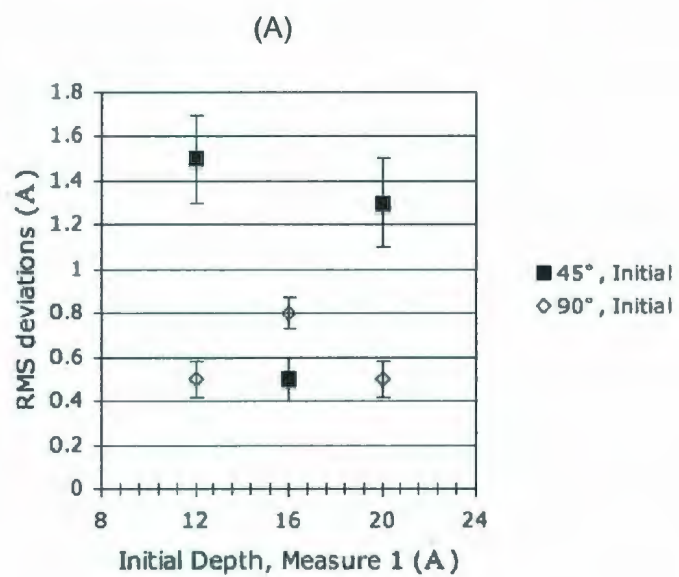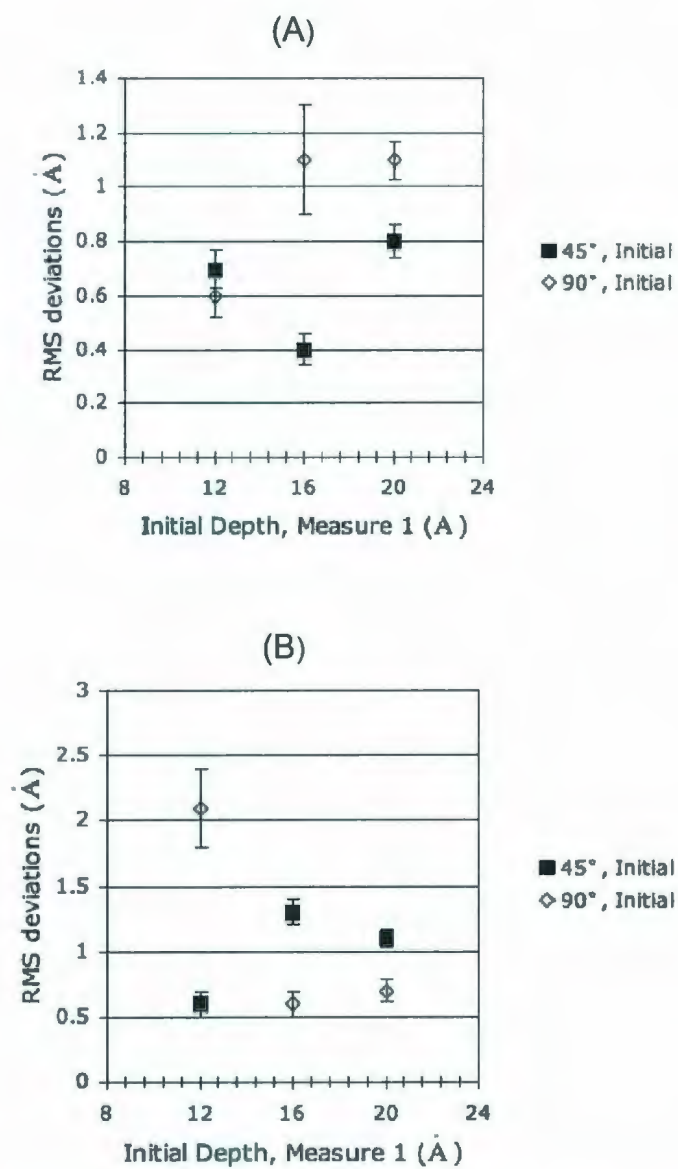
71

(A)

(B)

Figure 4.12: RMS deviations versus initial depth for (A) N-TERM alpha helix and (B) C-TERM alpha helix of the P6 systems.

the RMS deviations of an alpha helix was above this cut-off, the alpha helix was considered to have altered its structure by too much to still be considered alpha helical.

With this method of determining whether or not a system retained its alpha helices, it was found that overall there were three systems that did not retain both alpha helices: D63-45, D61-90, and P61-90. One alpha helix was lost from each of these systems. It is difficult to determine any correlation with regard to which systems tended to retain or not to retain both alpha helices since only three systems did not retain both alpha helices. These results were also not conclusive about what final length the alpha helices adopted in the monolayers. Nonetheless, it is still useful to know whether or not the structure of Mini-B that was determined in SDS micelles retained its alpha helices in the model lung surfactant monolayers of these simulations.

In general, except for the few cases mentioned, the plots of the RMS deviations seem to indicate that the structure of Mini-B determined in SDS micelles retained both its alpha helices in both types of monolayers, although the final lengths of the alpha helices remain undetermined. This is mainly due to the fact that the secondary structure plots may be too restrictive in determining whether or not a residue is part of an alpha helical structure, and therefore could not be used for this latter purpose. It is likely that the RMS deviations themselves cannot be used to determine the lengths of the alpha helices.

4.3 Final orientation of Mini-B

The final orientation or tilt angle of Mini-B was assessed for systems with DPPC:POPG monolayers. This was determined for each alpha helix, even though the length of each alpha helix may have changed in the simulations. It was also determined even for the systems which had lost an alpha helix (ie. D63-45 and D61-90 from the previous section). Although these systems may have lost an alpha helix, it is still of interest to see how the final structures are oriented.

The tilt angle was determined from the angle between the vector joining the center of masses of the backbone atoms of the first and last residues of a particular helix and the positive z-axis. This vector was oriented so that its tail end coincided with the residue at the terminal end of the helix and its head coincided with the residue at the other end of the helix, as illustrated in Fig. 4.13. If axes are drawn so that the head end of the vector coincides with the origin, then the angle between the positive z-axis and the vector corresponds to the tilt angle ϕ. The tilt angle was averaged over the last 10 or 20 ns.

The initial tilt angles of the helices were determined according to this measure from the systems which started at 55 Å^2 per lipid with Mini-B initially at 12 Å depth (ie. the D51-45 and D51-90 systems). Since the systems were prepared at initial tilt angles in the same way at other depths, it was assumed that it would be sufficient to determine the initial tilt angle only from the 12 Å
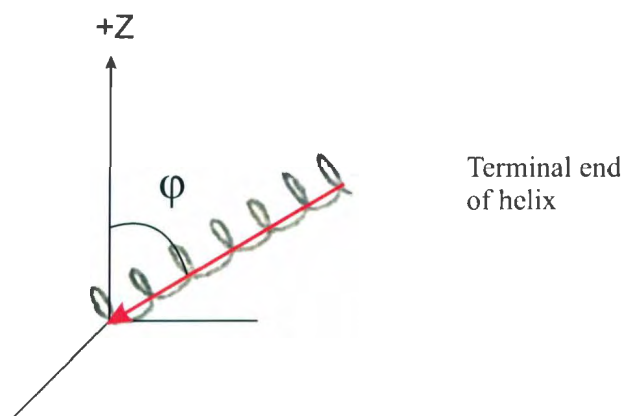
Figure 4.13: Illustration of tilt angle for a helix. This could correspond either to a C-terminal helix or an N-terminal helix.

depth case. These initial tilt angles were also found to be similar for systems that started at 60Å^2 per lipid with Mini-B initially at 12 Å depth. The initial tilt angles for the D51-45 system was 50.5° for the N-TERM helix and 50.9° for the C-TERM helix. The initial tilt angles for the D51-90 system was 94.4° for the N-TERM helix and 96.5° for the C-TERM helix. These values are similar to the nominal initial tilt angles of 45° and 90° discussed in chapter 3, which had previously been guessed mainly from visual inspection. It will be convenient to continue referring to the initial tilt angles by their nominal values. In other words, the initial tilt angle of the helices will continue to be referred to as either 45° or 90° as appropriate, consistent with the labelling in the notation. Note that in the case of the systems for which the initial tilt angles had been found to be close to 90°, the initial orientation of the

peptide is unchanged from what it originally had when it was obtained from the Protein Data Bank.

Plots of the final tilt angles for the D5 and D6 systems are provided in Figs. 4.14 and 4.15 for both the N-TERM and C-TERM alpha helices. These plots show that the final tilt angle of either helix lies in the range of approximately 70° to 110° if the peptide is initially placed deep enough in the monolayer, ie. typically at 20 Å (this is true for both the D5 and D6 systems). In other words, the final orientation of either helix of the peptide is fairly close to perpendicular to the normal to the interface, regardless of initial orientation, provided it is initially placed at 20 Å depth. This is of interest since the tilt angle of the original structure as determined experimentally in SDS micelles was close to 90° as mentioned above.

With the exception of the N-TERM helix in the D6 systems, the final tilt angles of the helices tend to remain relatively close to their initial tilt angles when the peptide is started off at 12 Å depth. In general these results indicate that the peptide must be placed low enough in the monolayer in order to ensure that the helices will maintain an orientation of closer to 90° with respect to the positive z-axis.

A graphic showing a backbone ribbon structure of Mini-B by itself towards the end of the simulation for the D53-90 system is provided in Fig. 4.16. This graphic also shows the positive z-axis as a blue arrow to compare it with the orientation of the helices. The C-TERM helix appears below the N-TERM

76
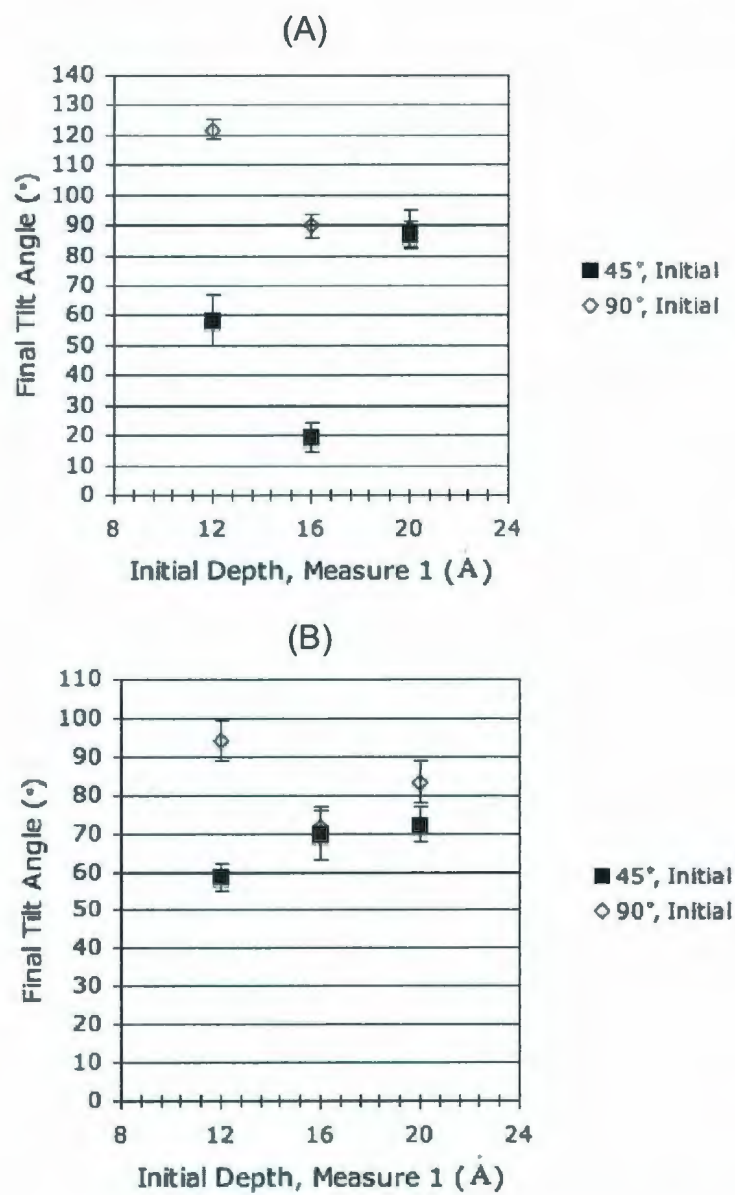
Figure 4.14: Final tilt angle versus initial depth (Measure 1) for (A) N-TERM helix and (B) C-TERM helix for D5 systems.

77
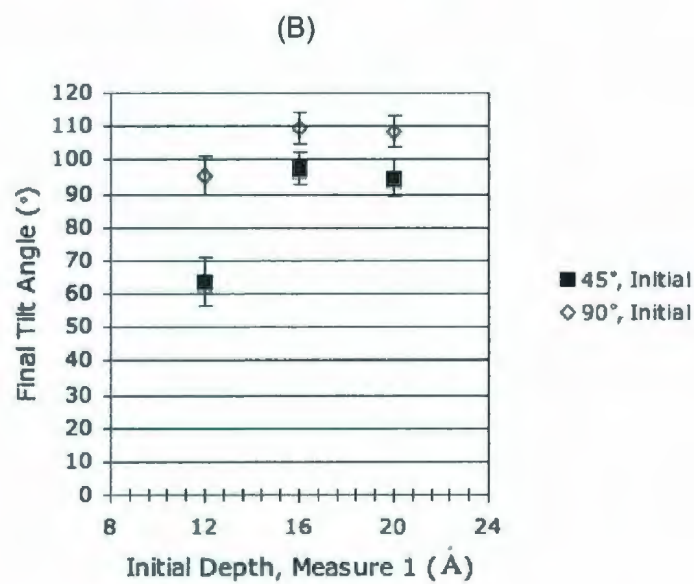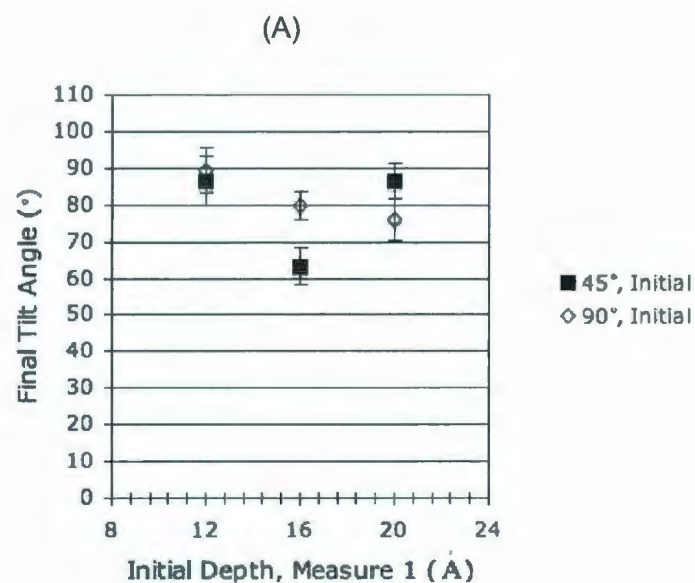
Figure 4.15: Final tilt angle versus initial depth (Measure 1) for (A) N-TERM helix and (B) C-TERM helix for D6 systems.

78

helix because the peptide is rotated as a whole about the axis passing between the helices by a small amount. Nonetheless, each helix is still close to 90° with respect to the positive z-axis.
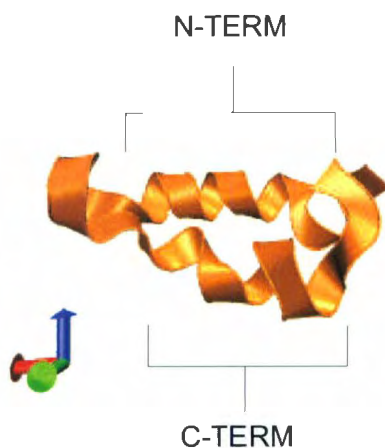
Figure 4.16: System D53-90 with N-TERM and C-TERM helices at roughly 90° to the normal to the interface. Blue arrow gives the direction of the positive z-axis.

4.4 Electrostatic interactions and the variation of sodium ions

After these initial simulations which involved varying the initial depth and orientation and analyzing the configuration of Mini-B in the monolayers, the Na+ concentration was varied for the D61-45 system (in other words which initially started at 12 Å depth and 45° tilt in the DPPC:POPG monolayer that was originally at 60 Å^2 per lipid). This system was chosen because results for the total energy had indicated that it was lowest in energy for the D6 systems.

Plots of the total energy are shown for the systems with DPPC:POPG monolayers in Fig. 4.17 below. In addition to showing that the total energy for the D6 systems is lowest for the system with starting depth of 12 Å and starting tilt angle of 45°, it also shows that the systems with starting tilt angles of 45° have lower energies compared to the systems with starting tilt angles of 90°. Similar results are seen for the D5 systems, although the total energy overall is higher at the smaller cross-sectional area per lipid, possibly due to closer packing of lipids. The origin of the difference in energy between systems with starting tilt angles of 45° and starting tilt angles of 90° is unclear and would likely require further analysis, but it does suggest that the initial tilt angle is a useful parameter to vary in attempting to determine the lowest energy configuration.

The discussion continues with the effects of varying sodium ion concentra-
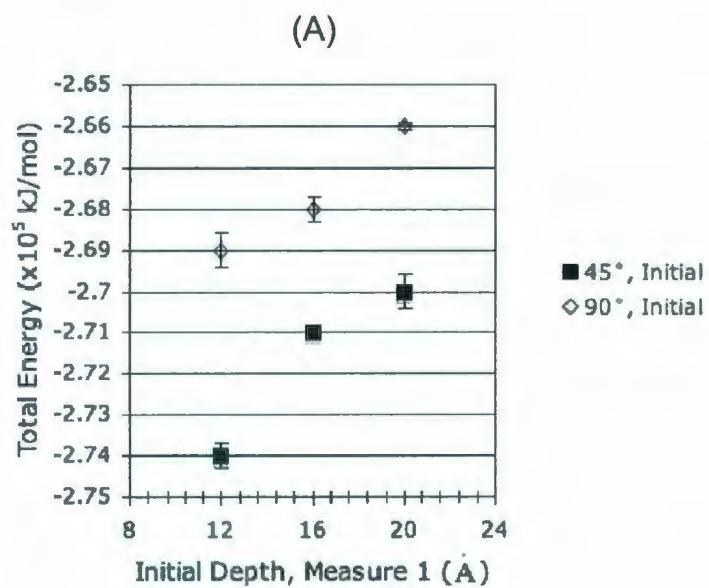
Figure 4.17: Total energy versus initial depth (Measure 1) for (A) D5 systems and (B) D6 systems.
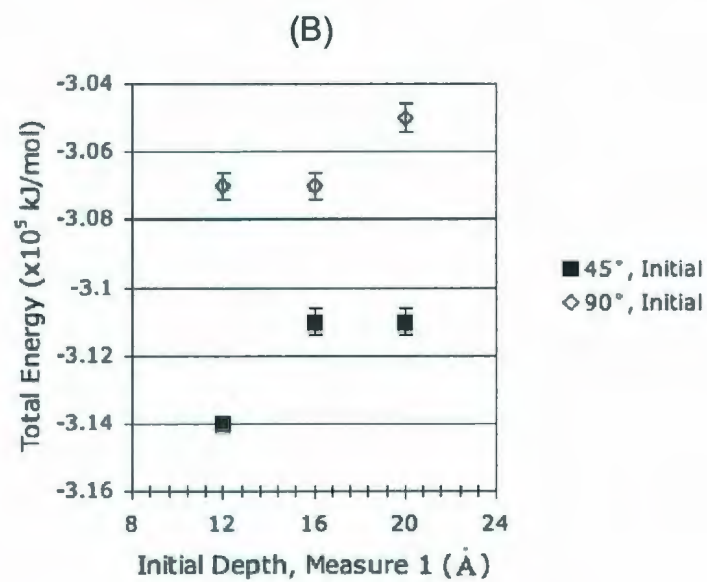
81

tions. In the previous simulations discussed in this chapter, the net charge for each of the systems was set to be zero. The sodium ion concentration was varied between roughly 50% to roughly 150% of the starting concentration. Each of these systems were simulated for 40 ns, following the initial 40 ns of the simulation for the system with 100% sodium ion concentration. Note that if the sodium ion concentration is greater than 100%, the number of sodium ions is greater than the amount needed to keep the system charge neutral. If the sodium ion concentration is less than 100%, the number of sodium ions is less than the amount required to keep the system charge neutral.

Afterwards the proximity of the cationic Arginine and Lysine residues to the POPG headgroups was assessed using radial distributions plots according to Eqn. 3.1. An example radial distribution plot for the system with 50% sodium ion concentration is provided in Fig. 4.18. A radial distribution function shows how one molecular group is distributed in space with respect to another molecular group. Generally it should have the form of a peak followed by a rapid decay. These radial distribution functions are used to describe the way in which the Arginine residues are distributed in space with respect to the POPG phosphorus atoms as well as how the Lysine residues are distributed with respect to the POPG phosphorus atoms.

In particular, the radial distances corresponding to the location of the peaks in these plots can be used as an indicator for where the cationic residues lie with respect to the POPG headgroups. The locations of the peaks were
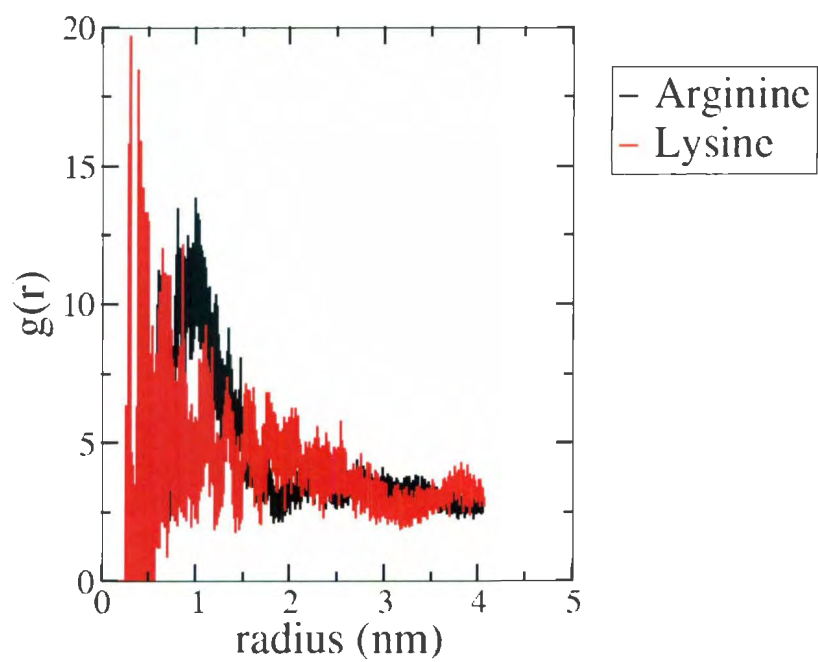
Figure 4.18: Radial distribution function of cationic residues with respect to POPG phosphorus atoms for system with 50% sodium ion concentration (averaged over last 10 ns).

determined from the radial distances at which the radial distribution functions had a maximum value. There were times where thin spikes in the graph at smaller radial distance could confuse the determination of these "peak radial distances". However it was assumed that the correct maximums would be found in the broader peaks of the types shown in Fig. 4.18. Hence the radial distribution plots were examined in each case and it was made sure by visual inspection that the correct maximum was obtained. In the following it is shown how these peak radial distances are affected by different sodium ion concentrations, and the relationship of the corresponding trends with the effect of electrostatic screening is discussed.

Plots of the peak radial distances as a function of percent sodium ion concentration are provided in Fig. 4.19 for the Arginine residue and Fig. 4.20 for the Lysine residue. Additionally trendlines are shown describing how these peak radial distances may change as the sodium ion concentration is increased. In determining these trends, some points had to be excluded from the data sets that were used in the plots because they seemed to fall off the main trend. These points are denoted by asterisks in the plots. For the plot of the Arginine residue, the excluded point corresponded to the system at 88% sodium ion concentration. For the plot of the Lysine residue, there were two excluded points which corresponded to the systems at 125% sodium ion concentration and 150% sodium ion concentration.

In Fig. 4.19 the excluded point was chosen because it seemed to be most

different from the other points in the plots, and the trend shown is the apparent trend. The R-squared value for the trendline is 0.2985. In Fig. 4.20 the points to exclude from the trendline were chosen based on the assumption that a maximum of only two points can be excluded. With this assumption, the R-squared values for all possible trendlines were determined. The results for the R-squared values are shown in Table 4.2 for different possible sets of points. The top of the table shows the R-squared value when no points are excluded. The next 9 cases show the R-squared values when only a single point is excluded. The remaining cases show the R-squared values when two points are excluded. From this table, it is seen that the best R-squared value is obtained when the points excluded are the same as in Fig. 4.20. The corresponding R-squared value is 0.7563, which is significantly above the next best value which is 0.3116. As to why the excluded points have such low values, it can only be guessed at this time that it is because of numerical error. The plots were checked for these cases and the main peaks were clearly shifted to lower radial distances.

These results for the trends indicate that the peak radial distances for the Arginine residue does not change drastically with respect to sodium ion concentration. However, the change in the peak radial distance of the Lysine residue from a sodium ion concentration of 50% to a sodium ion concentration of 137% is roughly 1 nm (note that the monolayer itself is approximately 8 nm wide in each direction).
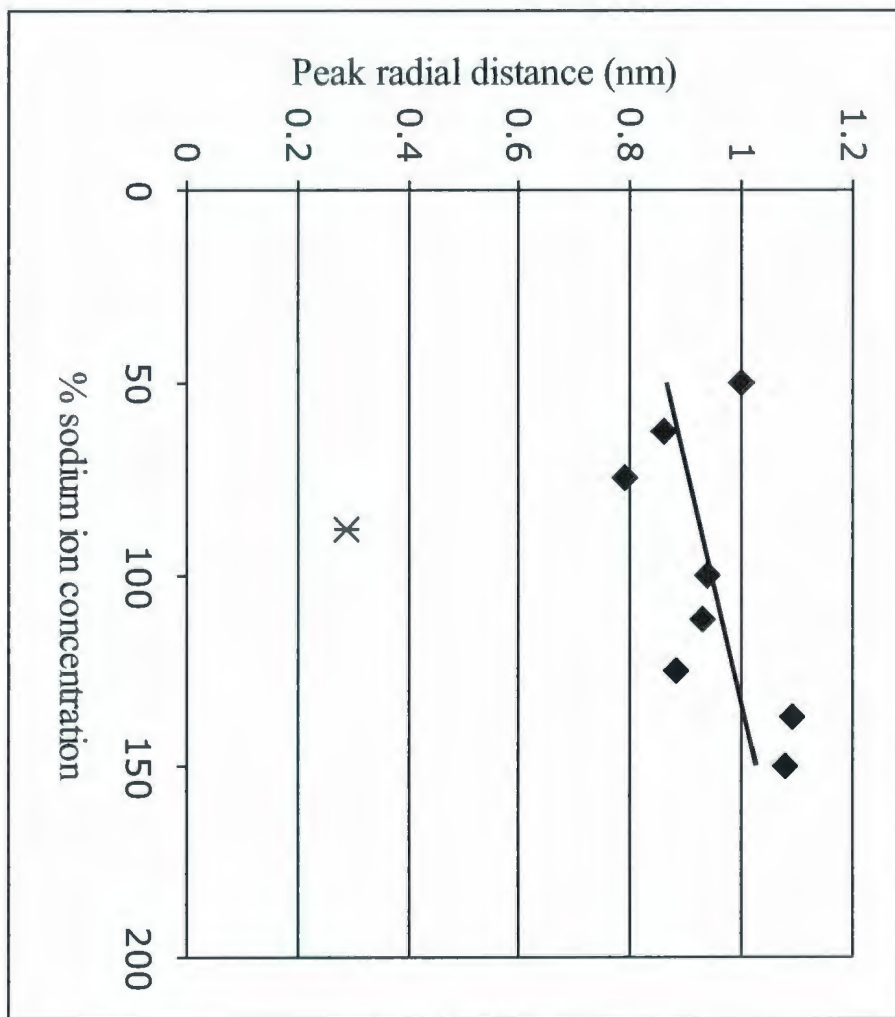
Figure 4.19: Plot of peak radial distance versus % sodium ion concentration for Arginine residue.
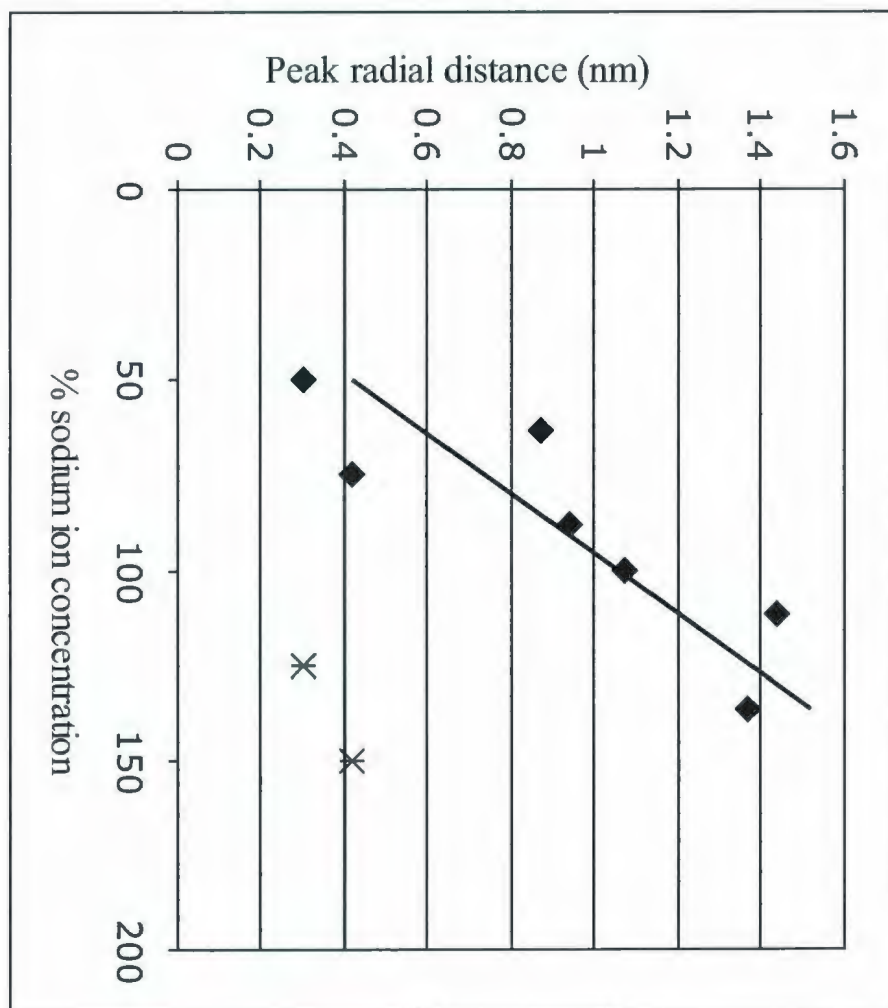
Figure 4.20: Plot of peak radial distance versus % sodium ion concentration for Lysine residue.

Table 4.2: R-squared values of trendline for cases where zero, one, or two points are excluded from plot of peak radial distances for Lysine residue (Fig. 4.20)

Point excluded (% sodium ion concentration)	Second point excluded (% sodium ion concentration)	R-squared value
nothing	nothing	0.0503
50	nothing	4×10^{-6}
63	nothing	0.0757
75	nothing	0.0231
88	nothing	0.0598
100	nothing	0.0532
112	nothing	0.0336
125	nothing	0.148
137	nothing	0.0012
150	nothing	0.2479
50	63	7×10^{-5}
50	75	0.0406
50	88	0.0003
50	100	0.0002
50	112	0.0027
50	125	0.0226
50	137	0.0598
50	150	0.0954

continued from previous page		
Point excluded (% sodium ion concentration)	Second point excluded (% sodium ion concentration)	R-squared value
63	75	0.0346
63	88	0.0967
63	100	0.0888
63	112	0.0739
63	125	0.1801
63	137	0.0102
63	150	0.3116
75	88	0.0292
75	100	0.0265
75	112	0.0134
75	125	0.1023
75	137	0.0015
75	150	0.2033
88	100	0.0661
88	112	0.0473
88	125	0.1569
88	137	0.0031
88	150	0.2583
100	112	0.0361
100	125	0.1486
100	137	0.0003
100	150	0.243
112	125	0.1227
112	137	0.0216
112	150	0.2015
125	137	0.0445
125	150	0.7563
137	150	0.0996

The data can be interpreted as suggesting that the sodium ions can screen the electrostatic interaction between the cationic residues of Mini-B and the anionic headgroups of POPG for higher sodium ion concentrations. This is particularly true for the Lysine residue. It is not clear why this trend does not appear to be stronger for the Arginine residue. It may be possible that Arginine residues that are close by to POPG phosphorus atoms are already attracted too close to the phosphate headgroups for sodium ions to come significantly in between. This has not been confirmed, however.

Fig. 4.21 shows the system at 63% sodium ion concentration. In Fig. 4.21 (B) a top view is shown and it can be seen that the sodium ions cluster around the POPG phosphorus atoms, although there are none close to the POPG phosphorus atoms which are in the proximity of Mini-B. It is likely that the positively charged residues of Mini-B repel the sodium ions that come near the vicinity of these phosphorus atoms. However, at a higher sodium ion concentration of 137%, it can be seen in Fig. 4.22(B) that sodium ions tend to come near the POPG phosphorus atoms which are close by to Mini-B. It seems that at the larger sodium ion concentration, some sodium ions are able to come in the vicinity of the phosphorus atoms and screen the interaction between positively charged residues and the negatively charge phosphate headgroups. In Fig. 4.22(A) it is seen that the Lysine side-chain seems to be bent further away from the phosphorus atom than in Fig. 4.21(A). In these graphics, the solvent is not shown and the POPG phosphorus atoms are colored gray while

90

the sodium ions are colored orange. In addition, the Lysine side-chains have been colored silver to distinguish them from the Arginine side-chains (all other colors remain the same).

This provides some confirmation that electrostatic screening is occuring in between the Lysine side-chains and the POPG phosphate headgroups. There may be a weak electrostatic screening effect occuring in the case of the Arginine residue given the slightly upward trend in Fig. 4.19, but this does not appear to be substantial. In a sense the existence of electrostatic screening, particularly in the case of the Lysine residue, is a confirmation of the strong electrostatic interactions than can occur between the Lysine residue and the POPG phosphate headgroups (if there were no strong electrostatic interactions, electrostatic screening effects would not be significant). While the interactions between the Arginine residues and the POPG phosphate headgroups may also be strong, it seems that this cannot be confirmed by varying the sodium ion concentration alone.

(A)

Lysine side-chain

POPG phosphorus atom

(B)

Lysine side-chain

POPG phosphorus atom
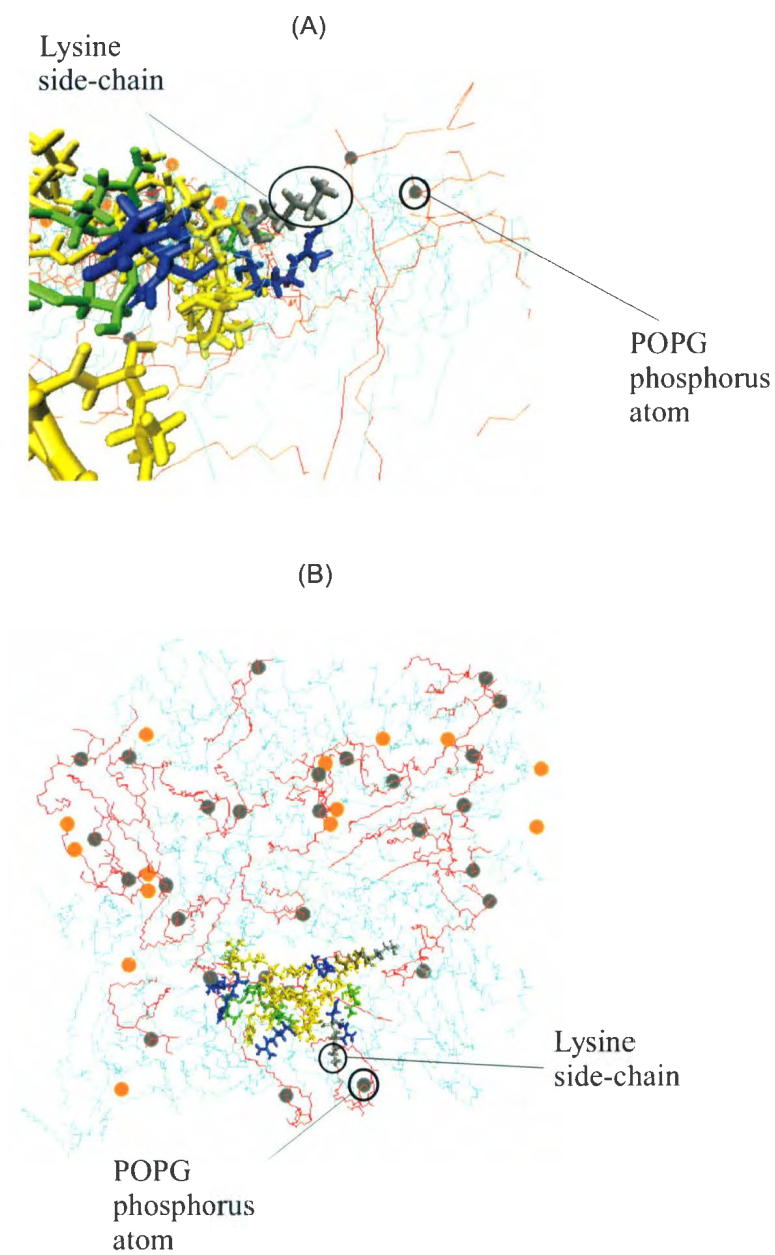
Figure 4.21: Graphics for system with 63% sodium ion concentration for (A) side-view and (B) top-view.
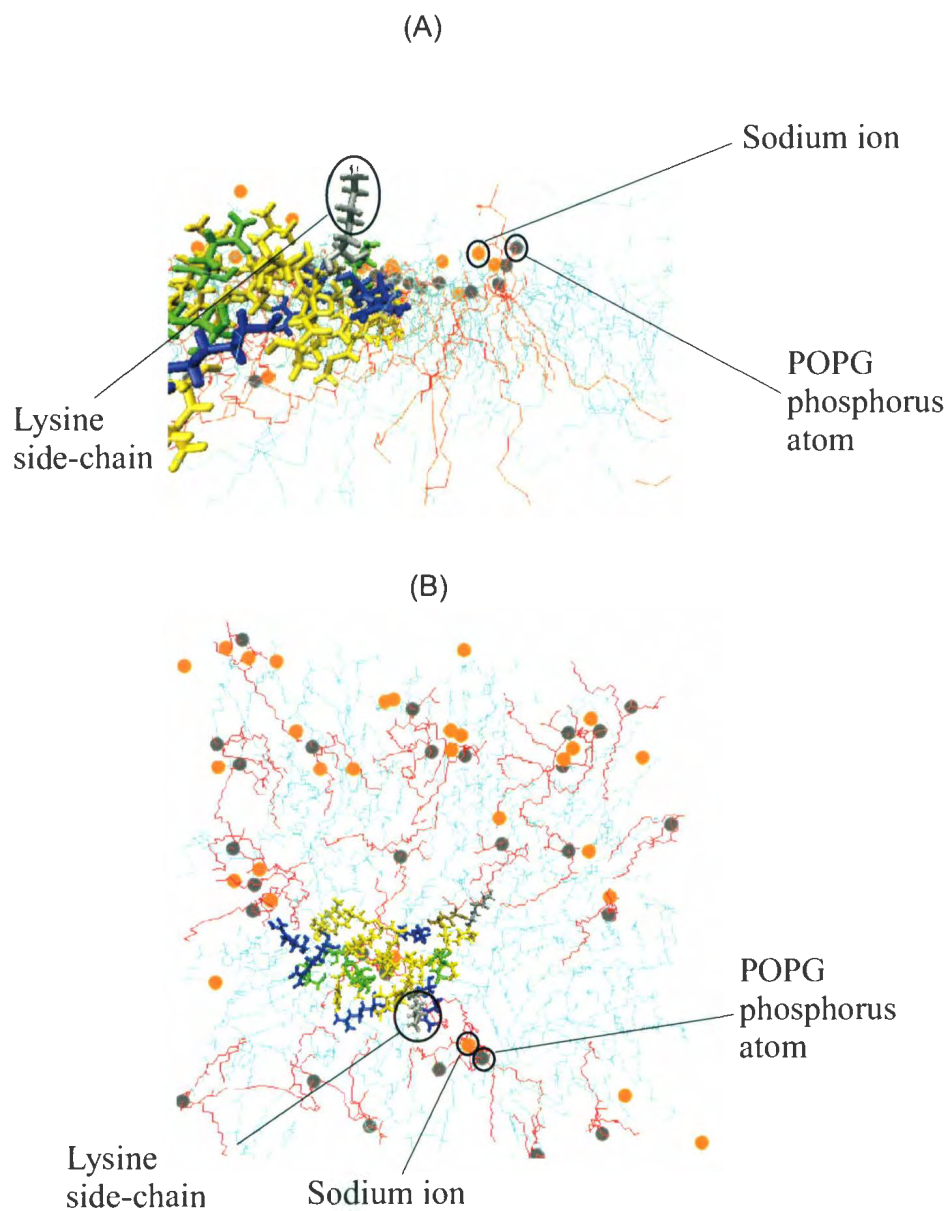
Figure 4.22: Graphics for system with 137% sodium ion concentration for (A) side-view and (B) top-view.

Chapter 5

Conclusions

The molecular dynamics of the peptide Mini-B was studied in model lung surfactant monolayers. It was found that the peptide had a final depth that tended to have similar values, independent of initial depth and orientation, as long as the peptide was placed low enough in the monolayers. In many cases "low enough" corresponded to a depth of 16 Å, though when the peptide was placed at an initial depth of 20 Å the peptide had a final depth that was independent of initial tilt angles for all cases of different compositions and cross-sectional areas per lipid.

An average final depth of the peptide was obtained for each case of a particular composition and cross-sectional area per lipid of the monolayer. For the DPPC:POPG systems the average depth is -3.4 +/- 0.9 Å for the systems with monolayers that started at 55 $Å^2$ per lipid and -4.7 +/- 0.9 Å for the systems

with monolayers that started at 60 Å^2 per lipid. For the POPG systems the average depth is -2.9 +/- 0.8 Å for the systems with monolayers that started at 55 Å^2 per lipid and -4.6 +/- 0.9 Å for the systems with monolayers that started at 60 Å^2 per lipid. From this it would seem that the cross-sectional area per lipid has a much greater affect on the depth of the peptide than composition, and suggests that the peptide is to some degree pushed upwards in the monolayer when the monolayer is compressed by a small amount.

An average final depth was also obtained when the system with initial depth of 12 Å and initial tilt angle of 45° was excluded from each of a particular composition and cross-sectional area per lipid. For this the final depth of Mini-B in the DPPC:POPG monolayers was found to be on average -5.3 +/- 1.0 Å for the systems with monolayers that started at 55 Å^2 per lipid and -5.2 +/- 0.9 Å for the systems with monolayers that started at 60 Å^2 per lipid. In addition, the final depth of Mini-B in the POPG monolayers was found to be on average -3.9 +/- 0.8 Å for the systems with monolayers that started at 55 Å^2 per lipid and -4.5 +/- 0.8 Å for the systems with monolayers that started at 60 Å^2 per lipid. The ranges of values seem to indicate that Mini-B was pulled higher towards the top of the POPG monolayers, quite possibly due to the attraction between the anionic headgroups of POPG and the cationic residues of the peptide.

The secondary structure of the peptide was also assessed and it was found to have retained both alpha helices in the majority of cases. Here it was

95

assumed that the RMS deviations of any alpha helix could not exceed 1.4 Å. With this assumption it was found that three systems did not retain both alpha helices, having lost one alpha helix during the length of the simulation. This is of interest, as it suggests that the alpha helices of Mini-B are stable for unrestrained simulations within the monolayers.

The orientation of Mini-B was also determined in the DPPC:POPG monolayers, and it was found that if the peptide is placed initially low enough in the monolayers (typically at 20 Å depth), its helices tended to attain a tilt angle that is relatively close to 90°. However if the peptide was not placed deep enough in the monolayers, ie. at 12 Å depth, it tended to have except for one case a final tilt angle closer to 45° if it started off at 45° and closer to 90° if it started off at 90°. This suggests that if the peptide is initially placed deep enough in the monolayer, it will attain an orientation that is not very much different from the orientation of the peptide structure that was determined in SDS micelles by experimental NMR, assuming simulation times on the order of 20-80 ns.

Further work was carried out in determining the effects of varying sodium ion concentrations for one particular system, the D61-45 system. Although this system had the peptide placed initially fairly shallow in the monolayer, it was chosen mainly based on energy data. The sodium ion concentration was varied between 50% and 150% of what was required to keep the system charge neutral. Data from radial distribution plots suggested that the location

of the Arginine residues with respect to the POPG phosphorous atoms (and hence the POPG phosphate headgroups) does not depend much on sodium ion concentration. However, the peak radial distance for the Lysine residue changes by roughly 1 nm between sodium ion concentrations of 50% and 137% (the system at 150% sodium ion concentration had been excluded in this case).

The plots of the peak radial distances suggested that there was electrostatic screening of the interaction between the Lysine residue and the POPG headgroups by the sodium ions present. This was particularly true as the sodium ion concentration increased. It seemed that at low sodium ion concentrations the positively charged residues of the peptide may set up a repulsive barrier that tends to keep the sodium ions from coming too close nearby the POPG headgroups. However, as the sodium ion concentration is increased, there may be a greater tendency for the sodium ions to penetrate through this barrier and come close enough to the POPG headgroups to screen the interaction between the POPG headgroups and the cationic residues of the peptide. Why this did not seem to occur for the Arginine residue is not clear, but this may require further study.

In future work, it would be of interest to look in addition at what effects the peptide has on the morphology of monolayers when they are compressed to very small cross-sectional areas per lipid, close to the point where the monolayers exhibit a plateau in their surface pressure-area isotherms. A study of the orientation of the lipid tail groups could also reveal what effect the peptide

97

has on the phases of the monolayer.

A further study of the effects of sodium ions might involve initially placing the peptide at a lower depth in comparison to the D61-45 case. One feature of the D61-45 case is that part of the peptide protuded above the top of the monolayer. It is of interest to determine whether similar effects of screening can be observed when the peptide is entirely below the top of the monolayer.

Bibliography

[1] Notter, R. H., *Lung Surfactants: Basic Science and Clinical Applications*, Marcel Dekker, Inc., New York, NY, United States, 2000.

[2] Walti, H., and Monset-Couchard, M., "A Risk-Benefit Assessment of Natural and Synthetic Exogenous Surfactants in the Management of Neonatal Respiratory Distress Syndrome", *Drug Safety*, **18**, 5, pp. 321-337, 1998.

[3] Lyra, P. P. R., and Diniz, E. M. A., "The Importance of Surfactant on the Development of Neonatal Pulmonary Diseases", *Clinics*, **62**, 2, pp. 181-190, 2007.

[4] Zuo, Y. Y., et al, "Current Perspectives in Pulmonary Surfactant – Inhibition, Enhancement, and Evaluation", *Biochimica et Biophysica Acta – Biomembranes*, **1778**, 10, pp. 1947-1977, 2008.

[5] Lang, C., et al, "New Insights into the Thermal Dynamics of the Surfactant System from Warm and Cold Animals", in K. Nag (Ed.),

Lung Surfactant Function and Disorder, pp. 17-57, Taylor & Francis Group, Boca Raton, FL, United States, 2005.

[6] Goodwin, J., *Colloids and Interfaces with Surfactants and Polymers: an Introduction*, John Wiley & Sons, Ltd., West Sussex, England, 2004.

[7] Serrano, A., and Perez-Gil, J., "Protein-Lipid Interactions and Surface Activity in the Pulmonary Surfactant System", *Chemistry and Physics of Lipids*, **141**, pp. 105-118, 2006.

[8] Yu, S., and Possmayer, F., "Lipid Compositional Analysis of Pulmonary Surfactant Monolayers and Monolayer-Associated Reservoirs", *Journal of Lipid Research*, **44**, 3, pp. 621-629, 2003

[9] Baoukina, S., et al, "The Molecular Mechanism of Monolayer-Bilayer Transformations of Lung Surfactant from Molecular Dynamics Simulations", *Biophysical Journal*, **93**, pp. 3775-3782, 2007.

[10] Haagsman, H. P., and van Golde, L. M. G., "Synthesis and Assembly of Lung Surfactant", *Annual Review of Physiology*, **53**, pp. 441-464, 1991.

[11] Alberts, B., et al, *Molecular Biology of the Cell*, 4th ed., Garland Science, New York, NY, USA, 2002.

[12] Rose, D., *MSc. Thesis*, 2007.

[13] Nkadi, P. O., et al, "An Overview of Pulmonary Surfactant in the Neonate: Genetics, Metabolism, and the Role of Surfactant in Health and Disease", *Molecular Genetics and Metabolism*, **97**, 2, pp. 95-101, 2009

[14] Lee, K. Y. C., et al, "Synchotron X-Ray Study of Lung Surfactant-Specific Protein SP-B in Lipid Monolayers", *Biophysical Journal*, **81**, 1, pp. 572-585, 2001.

[15] Lhert, F., et al, "Effects of Hydrophobic Surfactant Proteins on Collapse of Pulmonary Surfactant Monolayers", *Biophysical Journal*, **93**, 12, pp. 4237-4243, 2007.

[16] de Gennes, P., and Brochard-Wyart, F., *Capillarity and Wetting Phenomena: Drops, Bubbles, Pearls, Waves*, Springer-Verlag, Inc., New York, NY, United States, 2004.

[17] Rosen, M. J., *Surfactants and Interfacial Phenomena*, 3rd ed., John Wiley & Sons, Inc., Hoboken, New Jersey, United States, 2004.

[18] Petty, M., "Langmuir-Blodgett Films: an Introduction", Cambridge University Press, Cambridge, NY, United States, 1996.

[19] Cruz, A., et al, "Microstructure and Dynamic Surface Properties of Surfactant Protein SP-B/Dipalmitoylphosphatidylcholine Interfa-

cial Films Spread from Lipid-Protein Bilayers", *European Biophysics Journal with Biophysics Letters*, **29**, 3, pp. 204-213, 2000.

[20] Walters, R. W., et al, "Distinct Steps in the Adsorption of Pulmonary Surfactant to an Air-Liquid Interface", *Biophysical Journal*, **78**, 1, Part 1, pp. 257-266, 2000.

[21] Rugonyi, S., and Hall, S. B., "The Basis of Low Surface Tension in the Lungs", in K. Nag (Ed.), *Lung Surfactant Function and Disorder*, pp. 173-189, Taylor & Francis Group, Boca Raton, FL, United States, 2005.

[22] Lee, K. Y. C., "Collapse Mechanisms of Langmuir Monolayers", *Annual Review of Physical Chemistry*, **59**, pp. 771-791, 2008.

[23] Rugonyi, S., et al, "The Biophysical Function of Pulmonary Surfactant", *Respiratory Physiology and Neurobiology*, **163**, 1-3, pp. 244-255, 2008.

[24] Palaniyar, N., et al, "Receptors and Ligands for Collectins Surfactant Proteins A and D", in K. Nag (Ed.), *Lung Surfactant Function and Disorder*, pp. 87-123, Taylor & Francis Group, Boca Raton, FL, United States, 2005.

102

[25] Klenz, U., et al, "Influence of Lipid Saturation Grade and Headgroup Charge: A Refined Lung Surfactant Adsorption Model", *Biophysical Journal*, **95**, 2, pp. 699-709, 2008.

[26] Sarker, M., *MSc Thesis*, 2006.

[27] Kuroki, Y., and Voelker, D., "Pulmonary Surfactant Proteins", *The Journal of Biological Chemistry*, **269**, 42, pp. 25943-25946, 1994.

[28] Hawgood, S., et al, "Structure and Properties of Surfactant Protein B", *Biochimica et Biophysica Acta*, **1408**, 2-3, pp. 150-160, 1998.

[29] Branden, C., and Tooze, J., *Introduction to Protein Structure*, 2nd ed., Garland Publishing, Inc., New York, NY, United States, 1999.

[30] Booth, V., et al, "NMR Structures of the C-terminal Segment of Surfactant Protein B in Detergent Micelles and Hexafluoro-2-propanol", *Biochemistry*, **43**, 48, pp. 15187-15194, 2004.

[31] Sarker, M., et al, "Structure of Mini-B, a Functional Fragment of Surfactant Protein B, in Organic Solvent and Detergent Micelles", *Biochemistry*, **46**, 39, pp. 11047-11056, 2007.

[32] Waring, A. J., et al, "The Role of Charged Amphipathic Helices in the Structure and Function of Surfactant Protein B", *Journal of Peptide Research*, **66**, 6, pp. 364-374, 2005.

[33] Humphrey, W., et al, "VMD - Visual Molecular Dynamics", *Journal of Molecular Graphics*, **14**, 1, pp. 33-38, 1996.

[34] van der Spoel, D., et al, *Gromacs User Manual version 3.3*, www.gromacs.org, 2005.

[35] Lindahl, E., et al, "Gromacs 3.0: a Package for Molecular Simulation and Trajectory Analysis", *Journal of Molecular Modeling*, **7**, 8, pp. 306-317, 2001.

[36] van der Spoel, D., et al, "Gromacs: Fast, Flexible, and Free", *Journal of Computational Chemistry*, **26**, 16, pp. 1701-1718, 2005

[37] Berendsen, H. J. C., et al, "GROMACS: a Message-Passing Parallel Molecular-Dynamics Implementation", *Computer Physics Communications*, **91**, 1-3, pp. 43-56, 1995.

[38] Berendsen, H. J. C., et al, "Molecular Dynamics with Coupling to an External Bath", *The Journal of Chemical Physics*, **81**, 8, pp. 3684-3690, 1984.

[39] Hess, B., et al, "LINCS: A Linear Constraint Solver for Molecular Simulations", *Journal of Computational Chemistry*, **118**, 12, pp. 1463-1472, 1997.

[40] Scott, W. R. P., et al, "The GROMOS Biomolecular Simulation Program Package", *Journal of Physical Chemistry A*, **103**, 19, pp. 3596 - 3607, 1999.

[41] Jorgensen, W. L., et al, "Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids", *Journal of the American Chemical Society*, **118**, 45, pp. 11225-11236, 1996.

[42] Jorgensen, W. L., and Nguyen, T. B., "Monte-Carlo Simulations of the Hydration of Substituted Benzenes with OPLS Potential Functions", *Journal of Computational Chemistry*, **14**, 2, pp. 195-205, 1993.

[43] Sagui, C., and Darden, T. A., "Molecular Dynamics Simulations of Biomolecules: Long Range Electrostatic Effects", *Annual Review of Biophysics and Biomolecular Structure*, **28**, pp. 155-179, 1999.

[44] Darden, T., et al, "Particle Mesh Ewald: an N·log(N) Method for Ewald Sums in Large Systems", *Journal of Chemical Physics*, **98**, 12, pp. 10089-10092, 1993.

[45] Essmann, U., et al, "A Smooth Particle Mesh Ewald Method", *Journal of Chemical Physics*, **103**, 19, pp. 8577 - 8593, 1995.

[46] Ryckaert, J. P., and Bellemans, A., "Molecular Dynamics of Liquid Alkanes", *Faraday Discussions of the Chemical Society*, **66**, 66, pp. 95-106, 1978.

[47] Berger, O., et al, "Molecular Dynamics Simulations of a Fluid Bilayer of Dipalmitoylphosphatidylcholine at Full Hydration, Constant Pressure, and Constant Temperature", *Biophysical Journal*, **72**, 5, pp. 2002-2013, 1997.

[48] Tieleman, P., "Structures and Topologies", from http:// moose.bio.ucalgary.ca/index.php?page=Structures_and_Topologies.

[49] Zhao, W., et al, "Atomic-Scale Structure and Electrostatics of Anionic Palmitoyloleoylphosphatidylglycerol Lipid Bilayers with Na+ Counterions", *Biophysical Journal*, **92**, 4, pp. 1114-1124, 2007.

[50] Tieleman, D. P., et al, "Membrane Protein Simulations with a United-Atom Lipid and All-Atom Protein Model: Lipid-Protein Interactions, Side Chain Transfer Energies and Model Proteins", *Journal of Physics: Condensed Matter*, **18**, 28, S1221-S1234, 2006.

[51] Kandt, C., et al, "Setting Up and Running Molecular Dynamics Simulations of Membrane Proteins", *Methods*, **41**, 4, pp. 475-488, 2007.

[52] Fraternali, F., "Restrained and Unrestrained Molecular Dynamics Simulations in the NVT Ensemble of Alamethicin", *Biopolymers*, **30**, 11-12, pp. 1083-1099, 1990.

[53] Kaznessis, Y. N., et al, "Specific Mode of Interaction Between Components of Model Pulmonary Surfactants Using Computer Simulations", *Journal of Molecular Biology*, **322**, 3, pp. 569-582, 2002.

[54] Lee, H., et al, "Molecular Dynamics Simulations of the Anchoring and Tilting of the Lung Surfactant Peptide SP-B$_{1-25}$ in Palmitic Acid Monolayers", *Biophysical Journal*, **89**, 6, pp. 3807-3821, 2005.

[55] Atlantic Computational Excellence Network, *Mahone*, Retrieved June 26, 2009, from http://wiki.ace-net.ca/index.php/Mahone, 2009.

[56] Atlantic Computational Excellence Network, *Placentia*, Retrieved June 26, 2009, from http://wiki.ace-net.ca/index.php/Placentia, 2009.

[57] Western Canada Research Grid, *GROMACS on Matrix*, Retrieved June 26, 2009, from http://www.westgrid.ca/support/software/gromacs/matrix, 2008.

[58] Western Canada Research Grid, *Matrix Quick-Start Guide*, Retrieved June 26, 2009, from http://www.westgrid.ca/support/quickstart/matrix, 2008.

107

[59] Western Canada Research Grid, *Programming on the Westgrid Matrix System*, Retrieved June 26, 2009, from http://www.westgrid.ca/support/programming/matrix, 2008.

[60] Western Canada Research Grid, *Glacier Quick-Start Guide*, Retrieved September 22, 2009, from http://www.westgrid.ca/support/quickstart/glacier, 2008.

[61] Tseng, H. J. C., et al, "Master Curves and Radial distribution Functions for Shear Dilatancy of Liquid n-Hexadecane via Nonequilibrium Molecular Dynamics Simulations", *Journal of Chemical Physics*, **130**, 16, No. 164515, 2009.

[62] Kabsch, W., and Sander, C., "Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features", *Biopolymers*, **22**, 12, pp. 2577-2637, 1983.

`