Molecular Dynamics Computer Simulations of Lipid Bilayers Containing Cholesterol

Zhancheng Zhang

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Chemistry.

Chapel Hill

2009

Approved by:

Dr. Max, Berkowitz

Dr. Andrew, Moran

Dr. Garyk, Papoian

Dr. Nancy, Thompson

Dr. Gary, Pielak

ABSTRACT

Zhancheng Zhang: Molecular Dynamics Computer Simulations of Lipid Bilayers Containing Cholesterol

(Under the direction of Professor Max L. Berkowitz)

This thesis presents my research on molecular dynamics (MD) simulations of lipid bilayers. Some of the simulations were done on lipid bilayers containing one lipid component, while others were on lipid bilayers containing cholesterol and different species of lipids. The impetus behind these simulations is to use MD simulation technique to help understand the mechanism of "rafts" formation in model membranes, in which cholesterol molecules play an important role. Main results include: comparisons of structural properties of bilayers with and without cholesterol; hydrogen bonding network in bilayers containing cholesterol; saturated and unsaturated lipid tails' distribution around cholesterol; interaction potential energy distributions of cholesterol molecules with neighboring lipids; free energy calculation of transfer cholesterol between different lipid bilayers. One of the chapters will be dedicated to the dynamics of water molecules in bilayers with different hydration levels.

ACKNOWLEDGEMENT

First of all, I would like to thank my advisor Dr. Berkowitz for his encouragement and support during all these years. I learnt a lot physical chemistry from him, but more importantly, the persistency and the positive attitude towards research. He helped me in every step in my growth as a Ph.D. student. He helped me to make better presentations; he taught me how to write papers in every detail; he corrected things for me of this thesis and he helped me found a job. He is a great mentor and a big encouragement for me to do good science in the future. I would also like to give my gratitude to a former post doctoral, Dr. Bhide and a former student Dr. Lu in our group from whom I learnt a lot about molecular dynamics simulations. They were great lab mates and are great friends. We are still in touch with each other these days and it is always a great pleasure talking with them about science and life. I consider myself to be lucky to have my current office mates Charles Davis and Changsun Eun. These years spent with them as a group will never be forgotten. Whether it was just a small group meeting or a big conference in another city, it is worth cherished. I wish them the best of luck in everything in the future.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

CHAPTER 1: BACKGROUND

1.1 Molecular Dynamics Simulations

1.1.1 Basics Ideas

It is generally impossible to analytically determine the structural and dynamical properties of a molecular system containing a large number of atoms. In principle, these properties can be calculated using quantum mechanics by numerically solving the time-dependent Schrödinger equation for all the electrons and nuclei in the system. However, the computer power needed to perform such calculations is not available today, despite the fast developments in computer hardware and software. Fortunately, for the study of many molecular systems, we do not have to rely on Quantum Mechanics. The movements of the atoms in a system, such as a protein molecule in a solution, can be well described by classical mechanics. The first molecular dynamics (MD) simulations were performed in the late 1950s[1] by Alder & Wainwright to study a system of hard core spheres. In a MD simulation, Newton's equations of motion for a system of N interacting atoms are solved:

$$m_i \frac{d^2 \mathbf{r}_i}{dt^2} = \mathbf{F}i, \qquad i = 1 \ldots N \tag{1.1}$$

Where m_i and $\mathbf{r}_i$ are the mass and position of atom i, t is time, and $\mathbf{F}_i$ is the force on this atom, which is the negative gradient of the potential energy V ($\mathbf{r}_1$, $\mathbf{r}_2$, …$\mathbf{r}_N$) along $\mathbf{r}_i$:

$$\mathbf{F}_i = -\nabla_n V(\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_N) \tag{1.2}$$

After the forces on the atoms are calculated, the accelerations of the atoms can be determined. The velocities of the atoms can be updated after a very short time (usually on the order of 1fs) according to the accelerations. Subsequently, the positions of the atoms are updated according to the new velocities. After all the positions of the atoms are updated, the potential energy, and thus the forces on the atoms will change and be recalculated. Then the above-mentioned steps will be repeated for the duration of the simulation. This process is called *integrating the equations of motion*. The position, velocity and force on each atom will be saved with a certain time interval. The saved data are called a *trajectory* of the system. Assuming the potential energy function is a good approximation of the interactions between the atoms, a MD simulation can provide a very detailed picture of the system: both structural and dynamical properties can be described on an atomic level. The first macromolecular MD simulation was performed in 1977 by McCammon[2] et al. on a system containing 500 atoms, the simulation time was 9.2 ps. Since this watershed simulation, the field has witnessed an explosive development. In 2006, the Theoretical and Computational Biophysics Group from University of Illinois at Urbana Champaign performed an MD simulation of the complete satellite tobacco mosaic virus[3] to probe the mechanism of virus assembly. Their system contained 1 million atoms and was simulated for 50 ns. This simulation would take a single 2006 desktop computer around 35 years to complete, thus necessitating the use of many processors concurrently in parallel. The same year, a simulation of the folding of the Villin Headpiece in full atomic detail was performed by a group from Stanford University under the leadership of Vijay Panda[4]. The Villin

Headpiece consisted of 20,000 atoms, and it was simulated for 500 μs. This simulation was run on 200,000 CPU's from around the world using a distributed computing architecture. These are just two of many great examples of what a powerful tool a MD simulation can be. They also demonstrate that a MD simulation can require a significant amount of computer resources and time. However, with better parallelization and further software development (like the development of the popular MD simulation package GROMACS[5,6], with which all the work mentioned in this thesis was done), the time and length scale in MD simulations keep increasing. Today, MD simulations have become an indispensable tool for biologists, chemists and physicists for studying fundamental phenomenon at an atomic level not available to experimental techniques.

1.1.2 Force Fields

One of the crucial requirements for a MD simulation to be useful is a good choice of the potential energy function $V(\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_N)$. This function should provide a good description of the underlying interactions between the atoms in the system while can easily be evaluated on a computer. It is always the case that many approximations need to be made to achieve these requirements. The final formula of this function, together with the parameters used in it, is called a *force field.* The potential energy is usually divided into different contributions from bonded interactions and nonbonded ones. The bonded interactions include interactions due to bond length variations, bond angle bending and bond rotations, or dihedral angle variations. The potential energy

due to bond length variation between two atoms *i* and *j* is usually described by a harmonic potential:

$$V_{bond}(r_{ij}) = \frac{k_{ij}^{b}}{2}(r_{ij} - r_{ij}^{0})^{2} \tag{1.3}$$

where r_{ij} is the bond length between atoms *i* and *j* at the moment the potential is calculated, k_{ij}^{b} is the force constant representing the stiffness of the bond being described. r_{ij}^{0} is the equilibrium bond length. A similar harmonic potential is usually adopted to describe the energy of an angle bending motion of three atoms *i*, *j* and *k,* with atom j at the vertex of the angle:

$$V_{angle}(\theta_{i\,jk}) = k_{ijk}^{\theta}(\vartheta_{ijk} - \theta_{ijk}^{0})^{2} \tag{1.4}$$

where θ_{ijk} is the value of the angle formed by atoms *i*, *j* and *k* at the moment the potential is calculated, k_{ijk}^{ϑ} is the force constant describing how difficult it is to distort the angle and θ_{ijk}^{0} is the equilibrium value of the angle. The dihedral angle energy term is the energy due to the rotation around the middle bond in a sequence of four atoms as shown in Fig. 1.1 on the left. The dihedral potential is usually modeled with a periodic function:

$$V_{dihedral}(\varphi_{i\,jkl}) = k_{jk}^{\varphi}[1 + \cos(n\varphi - \varphi_{0})] \tag{1.5}$$

There are two different ways to define the zero point of the angle; here it is defined according to the IUPAC/IUB convention, with zero corresponding to the *cis* configuration (atom *i* and *l* on the same side of bond *jk*). Here, k_{jk}^{φ} determines how hard it is to rotate around bond *jk*, and *n* determines the number of minima during a full rotation of the bond while φ_{0} their positions. The appearance of such a simple function is shown in Fig. 1.1 on the right.

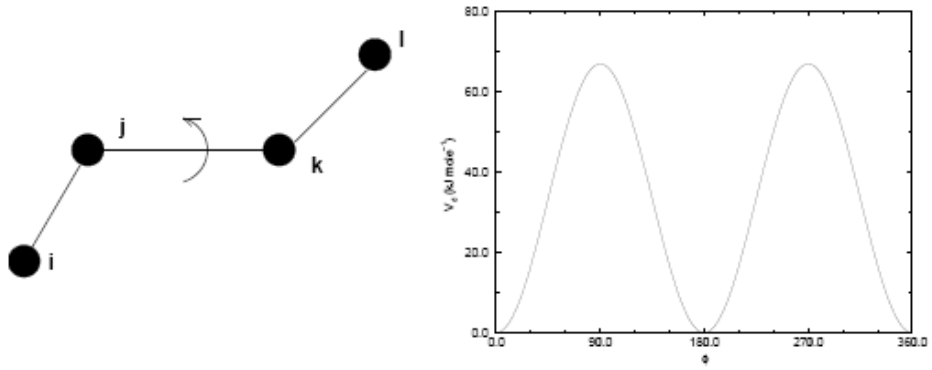

Fig 1.1 Definition of a dihedral angle (left) and the dihedral angle potential (right). (Source: GROMACS manual-3.2)

For alkanes, the Ryckaert-Bellemans dihedral potential[7] of the following form is usually used:

$$V_{rb}(\varphi_{ijkl}) = \sum_{n=0}^{5} C_n (\cos(\phi))^n \qquad (1.6)$$

where $\phi = \varphi - 180°$.

Fig. 1.2 shows the appearance of this function.

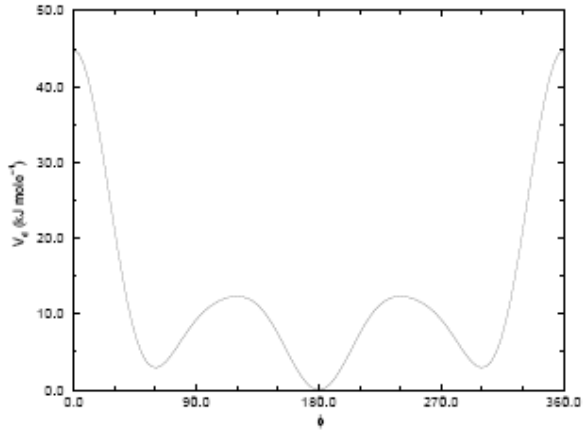

Figure 1.2: Rychkaert-Bellemans dihedral potential. (Source: GROMACS manual-3.2)

This function takes into account that the two *gauche* states at $\varphi = \pm 60°$ are less favorable than the *trans* state at $\varphi = 180°$. Also the transition from a *gauche* state to the *trans* state without crossing the *cis* one should be much easier than a transition that crosses the *cis* state. These two features can be seen from Fig. 1.2. One should keep in mind that when the simple periodic dihedral potential (1.5) is used, the nonbonded interaction between atom *i* and *l* should be included, while if the Rychkaert-Bellemans potential is used to describe a dihedral angle, the nonbonded interaction between atom *i* and *l* should be excluded. These dihedral energy functions are used to describe the normal dihedral interaction or the proper dihedral interaction. Another form of dihedral potentials, improper dihedrals, is used to keep planar groups planar (e.g. aromatic rings) or to prevent molecules from flipping over to their mirror images. This potential only applies to certain groups in a molecule. Usually a harmonic potential function is employed for this purpose:

$$V_{id}(\xi_{ijkl}) = k_{\xi}(\xi_{ijkl} - \xi_0)^2 \tag{1.7}$$

where k_{ξ} is the force constant and ξ_0 is the equilibrium dihedral angle value.

The bonded interactions discussed above are local interactions. The number of these interactions will be proportional to the number of atoms in the system. As a result, they are not very expensive to calculate in a MD simulation. However, the CPU time required for the calculation of nonbonded interactions is a totally different story.

Nonbonded interactions include the van der Waals interaction between atoms as well as the electrostatic interaction between charged particles. The Lennard-Jones

potential function is often used to represent the van der Waals interaction:

$$V_{LJ}(r_{ij}) = \frac{C_{ij}^{(12)}}{r_{ij}^{12}} - \frac{C_{ij}^{(6)}}{r_{ij}^{6}} \tag{1.8}$$

where r_{ij} is the distance between the pair of atoms and $C_{ij}^{(12)}, C_{ij}^{(6)}$ are interaction parameters which depend on the type of atoms between which the interaction being calculated. The first term represents the repulsive forces in short atomic distances due to the overlap of electronic wave functions. Sometimes r_{ij}^{9} instead of r_{ij}^{12} can be used in the denominator. The second term represents the attractive London dispersion forces, named after the German-American physicist Fritz London. These attractive forces are always present and arise from the interactions between temporary multipoles in molecules without permanent multipole moments. Fig 1.3 shows a Lennard-Jones interaction function:

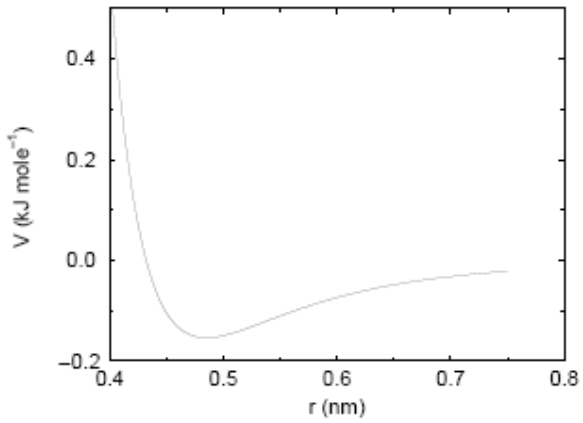

Fig 1.3 The Lennard-Jones interaction. (Source: GROMACS manual-3.2)

This weak interaction approaches zero at relatively short distance, thus, to decrease computational load, these interactions are truncated at a cutoff distance at around

1nm.

The electrostatic interactions between charged particles are given by:

$$V_{coulomb}(r_{ij}) = \frac{q_i q_j}{4\pi\varepsilon_0 \varepsilon_r r_{ij}^2} \tag{1.9}$$

where q_i and q_j are the charges, ε_0 is the permittivity of vacuum and ε_r is the relative permittivity. These interactions decay much slower than the weak Lennard-Jones potential, which makes it one of the biggest challenges in MD simulations. A longer cutoff distance (2 nm, e.g.) can be used to truncate these interactions, but a simple truncation can still introduce significant errors. The reaction field method is another way to calculate the coulomb interactions for a homogeneous system. It assumes a constant electric environment beyond the cutoff distance. The interaction becomes:

$$V_{crf}(r_{ij}) = \frac{1}{4\pi\varepsilon_0}\frac{q_i q_j}{r_{ij}}[1 + \frac{\varepsilon_{rf} - 1}{2\varepsilon_{rf} + 1}\frac{r_{ij}^3}{r_c^3}] - \frac{1}{4\pi\varepsilon_0}\frac{q_i q_j}{r_c}\frac{3\varepsilon_{rf}}{2\varepsilon_{rf} + 1} \tag{1.10}$$

Here, r_c is the cutoff distance and ε_{rf} is the dielectric constant of the environment beyond the cutoff distance. The constant term on the right makes the potential zero at the cutoff. Still yet another method to calculate the long range electrostatics in an MD simulation is the Ewald summation, which was first introduced to calculate the long range interactions of the periodic images in crystals[8]. The electrostatic energy of *N* particles and their periodic images is given by:

$$V = \frac{1}{8\pi\varepsilon_0}\sum_{n_x}\sum_{n_y}\sum_{n_z}^{*}\sum_{i}^{N}\sum_{j}^{N}\frac{q_i q_j}{r_{ij,\mathbf{n}}} \tag{1.11}$$

where **n** = (n_x, n_y, n_z) is the box index vector. The star means when **n** = (0, 0, 0), the interaction between particles $i = j$ should be excluded. The summation converges

conditionally and slowly. In Ewald summation, the sum in (1.11) is split into two quickly-converging terms and a constant term, as shown below:

$$V = V_{dir} + V_{rec} + V_0 \tag{1.12}$$

$$V_{dir} = \frac{1}{8\pi\varepsilon_0} \sum_{i,j}^{N} \sum_{n_x} \sum_{n_y} \sum_{n_z}^{*} q_i q_j \frac{erfc(\beta r_{ij,\mathbf{n}})}{r_{ij,\mathbf{n}}} \tag{1.13}$$

$$V_{rec} = \frac{2\varepsilon_0}{V} \sum_{i,j}^{N} q_i q_j \sum_{m_x} \sum_{m_y} \sum_{m_z}^{*} \frac{\exp(-(\pi \mathbf{m}/\beta)^2 + 2\pi i \mathbf{m} \cdot (\mathbf{r}_i - \mathbf{r}_j))}{\mathbf{m}^2} \tag{1.14}$$

$$V_0 = -4\sqrt{\pi}\varepsilon_0 \beta \sum_{i}^{N} q_i^2 \tag{1.15}$$

where β is a parameter that determines the relative weights of the direct and reciprocal sums and $\mathbf{m}=(m_x, m_y, m_z)$ is the reciprocal-space vector. V in equation (1.14) is the volume of the crystal, or the volume of the simulation box. A short cutoff can be used to truncate both the real space sum (on the order of 1nm) and the reciprocal space (10 wave vectors in each direction). The computational time of the reciprocal sum increases as N^2, which makes the Ewald summation not suitable for simulations of large systems. The Particle-Mesh Ewald proposed by Tom Darden[9,10] scales as $N\log(N)$ and is much faster than the original Ewald summation for large systems. Today, PME is the most popular scheme for calculating electrostatic interactions in MD simulations. As mentioned previously, the nonbonded interactions are not confined to atoms connected to each other and the total number of these interactions can be proportional to the square of the number of atoms in the system. Thus, the calculation of the nonbonded interaction in an MD simulation is the most time consuming part. There are other functional forms available to describe the energy terms mentioned

above, such as using the Morse potential[11] for the bonding stretching, or a more flexible and realistic exponential term for the repulsive part in the nonbonded interaction between atoms, as in the Buckingham potential. However, the exponential terms in these functions make them more expensive to calculate. Different force fields may take different functional forms. One should realize that a potential form better representing the underlying physics is not necessarily a better potential function for a simulation. The simplicity of a function for evaluation by a computer is always an essential factor in choosing a particular function form. As the force field is always an approximation of the interactions between atoms and it is parameterized in such a way that the equilibrium properties of the system are reproduced as well as possible. Thus, a simple function that can be easily evaluated but bears more approximations may prove to be much more useful than a more complicated one that describes the underlying physics in more detail. It is also worth mentioning that the time step used in the integration of the equations of motion can be part of a force field and a force field has a temperature range within which it gives reasonable descriptions of the interactions in the system simulated.

1.1.3 Update Configuration

After the forces on all the atoms have been calculated, the positions and velocities of the atoms are updated according to the equations of motion (1.1) and (1.2). These equations are solved numerically and there are many algorithms available. Among these algorithms, the one developed by Verlet[12] has proved to be the best for

molecular dynamics. The idea behind this algorithm is the addition and subtraction of the Taylor expansion of an atom's position at time $t - \Delta t$ and $t + \Delta t$:

$$\mathbf{r}_i(t - \Delta t) = \mathbf{r}_i(t) - \Delta t \frac{d}{dt}\mathbf{r}_i(t) + \frac{(\Delta t)^2}{2!}\frac{d^2}{dt^2}\mathbf{r}_i(t) - \frac{(\Delta t)^3}{3!}\frac{d^3}{dt^3}\mathbf{r}_i(t) + \mathrm{O}(\Delta t^4) \tag{1.16}$$

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \frac{d}{dt}\mathbf{r}_i(t) + \frac{(\Delta t)^2}{2!}\frac{d^2}{dt^2}\mathbf{r}_i(t) + \frac{(\Delta t)^3}{3!}\frac{d^3}{dt^3}\mathbf{r}_i(t) + \mathrm{O}(\Delta t^4) \tag{1.17}$$

The addition of these two expansions, together with equation (1.1) gives:

$$\mathbf{r}_i(t + \Delta t) \approx -\mathbf{r}_i(t - \Delta t) + 2\mathbf{r}_i(t) + \frac{\Delta t^2}{m_i}\mathbf{F}_i \tag{1.18}$$

With a truncation error on the order of $(\Delta t)^4$. The subtraction of equations (1.17) and (1.16) gives:

$$\mathbf{v}_i(t) = \frac{d}{dt}\mathbf{r}_i(t) \approx \frac{1}{2\Delta t}[\mathbf{r}_i(t + \Delta t) - \mathbf{r}_i(t - \Delta t)] \tag{1.19}$$

With a truncation error on the order of $(\Delta t)^3$. However, a practical problem with the calculation of velocities in this way is that the velocities are evaluated as the difference of two quantities of the same magnitude, thus the evaluation is very sensitive to numerical precision and round-off errors. The Leap-Frog scheme[13] circumvented this problem by updating the positions and velocities according to the following equations:

$$\mathbf{r}_i(t + \Delta t) \approx \mathbf{r}_i(t) + \Delta t \mathbf{v}_i(t + \Delta t/2) \tag{1.20}$$

$$\mathbf{v}_i(t + \Delta t/2) \approx \mathbf{v}_i(t - \Delta t/2) + \frac{\Delta t}{m_i}\mathbf{F}_i \tag{1.21}$$

It can be easily shown that equations (1.20) and (1.21) are equivalent to equations (1.16) and (1.17). One drawback of this scheme is that the positions and velocities are not calculated at the same time (offset by half a time step, thus the naming of this

algorithm). The velocities at whole time steps, however, can always be calculated as the average of the velocities at plus and minus half a time step of that whole step. There are other more advanced integrators available, like the Gear algorithm and other predictor-corrector schemes[14-16]. These algorithms can give more accurate positions and velocities if a small time step is used than the simple integrators can. When a large time step is used, these advanced schemes will give much bigger energy drift than the simple schemes do. Since one wants to use a time step that is as big as possible in a MD simulation to speed up the sampling, there is no point in using a very small time step. The reason behind this is the *shadow trajectory*[17] argument: since the equations of motion for all atoms are solved numerically, some numerical errors will always exit. And these errors grow exponentially as the simulation proceeds. Thus, in a simulation, it is impossible to reproduce the "exact trajectory" a particular atom will follow in a real system. This does not matter though. The equilibrium properties of the system are not sensitive to the details of an individual atom's trajectory. One only needs to guarantee that a long enough part of the trajectory extracted from a simulation would statistically provide a fair description of a particle with the same initial conditions. This trajectory is called a *shadow trajectory.* In this sense, the Leap-Frog scheme does an excellent job. It has relatively large energy fluctuations in short time but gives very low energy drifts in the long run. Together with the efficiency it has in integrating the equations, the Leap-Frog scheme is by far the most popular integration scheme in MD simulations.

1.1.4 Limitations

As any technology used in science, MD simulation has its own limitations and to know the limitations of a method is as necessary as to know its utilities.

First, equation (1.1) and (1.2) tell that an MD simulation is completely classical. This is fine for many heavy atoms at normal temperature but for the motion of a light hydrogen atom, it is well known that many aspects of this motion are quantum mechanical in nature, like tunneling through an energy barrier in the course of transfer of a hydrogen bond. Also, electrons are supposed to be in their ground state and not described explicitly in a MD simulation. Instead, the *Born-Oppenheimer* approximation is used: the electrons will adjust their positions instantly as the positions of the atoms change. As a result, electron transfer cannot be explicitly modeled, like chemical reactions where bonds can form and break.

Second, the force fields are approximate. From the form of the energy functions to the parameters used in them, many approximations are made. Most force fields are pair-additive, and non pair-additive interactions, like interactions through atomic polarizability are represented by effective pair-additive potentials. This can incur problems where parameterizations are not consistent amongst different systems under separate conditions. As a result, one should always be cautious when applying a force field to a system with different conditions. Results should be checked carefully and, sometimes, re-parameterization will be needed. Another approximation with the force fields is that long range interactions are truncated at a distance of 1-2 nm. The effect of using a cutoff may be minor for the Lennard-Jones interactions but it is certainly

not for the electrostatic interactions. Though there are other methods to handle this portion of energy calculations, like the PME[9,10], they also have their own pitfalls (see below).

Third, the boundary conditions are unnatural. Due to the limited computational capability, a simulated system contains much less particles compared to a macroscopic system (a system containing 100,000 particles is huge in a simulation). As a result, a large portion of the particles will be close to an unnatural boundary of vacuum and this can cause too many artifacts on the properties of the system. To reduce this edge effect, *periodic boundary conditions (PBC)* are used in simulations. The idea of *PBC* is that the simulation box is surrounded by translated copies of itself in all directions and atoms that cross one side of the box enter from the other side of it immediately. The Lennard-Jones interactions between atoms are only calculated between the closest images of the atoms. A *PBC* is desired if one simulates a crystal structure. However, liquids or solutions are not crystal and *PBC* itself causes errors. The application of PME to calculate the electrostatic interaction in a periodic system may introduce extra order to it.

As the development of better parallelization, better algorithm and faster hardware proceeds, the length and time scales accessible to MD simulations will continue to increase. The boundary effect may be the first thing we can erase from our concerns. Also, there are methods (e.g. QM/MM[18]) developed to handle the electrons' degrees of freedom in some parts of a simulated system, and chemical reactions can be simulated. The parameterization of the force field remains a big challenge, especially

for new systems.

1.2 Free Energy Calculations in MD Simulations

1.2.1 Statistical Background

In a canonical ensemble with N particles, the Helmholtz free energy is given by:

$$A = -\frac{1}{\beta}\ln Q_{NVT} \tag{1.22}$$

where $\beta = 1/k_B T$. k_B is the Boltzmann constant and T is the temperature of the system. Q_{NVT} is the partition function of the system:

$$Q_{NVT} = \frac{1}{h^{3N} N!}\int\int \exp[-\beta H(\mathbf{x}, \mathbf{p}_x)] d\mathbf{x} d\mathbf{p}_x \tag{1.23}$$

where $H(\mathbf{x},\mathbf{p}_x)$ is the Hamiltonian of the system. $\mathbf{x}$ and $\mathbf{p}_x$ are $3N$-dimentional vectors containing the atomic coordinates and momenta of the particles in the system. h is the Planck constant. The $N!$ term is due to the fact that the N particles are indistinguishable. The probability to find the system in a particular microscopic configuration ($\mathbf{x}$, $\mathbf{p}_x$) is defined as:

$$P(\mathbf{x}, \mathbf{p}_x) = \frac{\exp[-\beta H(\mathbf{x}, \mathbf{p}_x)]}{\int\int \exp[-\beta H(\mathbf{x}', \mathbf{p}'_x)] d\mathbf{x}' d\mathbf{p}'_x} \tag{1.24}$$

From the definition of the free energy and partition function for a system, it is obvious that it is virtually impossible to calculate the free energy of a system from a finite MD simulation: a finite MD simulation won't be able to sampling all the configurations in limited time. However, in practice, people are generally interested in a free energy difference between two systems, like a reference system and a mutated one, or the Potential of Mean Force (PMF) as a function of a reaction coordinate when a system

is transformed from state A to state B. These quantities are possible to obtain from a simulation. Two approaches: the Free Energy Perturbation (FEP) and Umbrella Sampling methods[19-27], are usually used to solve these problems.

1.2.2 Free Energy Perturbation

Suppose one is interested in estimating the free energy difference between a reference system, **a**, and a target system, **b**. Assume that system **a** is described by Hamiltonian $H_a(\mathbf{x}, \mathbf{p}_x)$ and system **b** by $H_b(\mathbf{x}, \mathbf{p}_x)$, such that:

$$H_b(\mathbf{x}, \mathbf{p}_x) = H_a(\mathbf{x}, \mathbf{p}_x) + \Delta H(\mathbf{x}, \mathbf{p}_x) \tag{1.25}$$

Here, $\Delta H(\mathbf{x},\mathbf{p}_x)$ is the difference between the two Hamiltonians, or the perturbation between the final and initial states. According to equation (1.22), the free energy difference between system **b** and system **a** can be written as:

$$\Delta A_{a\to b} = -\frac{1}{\beta}\ln\frac{Q^b_{NVT}}{Q^a_{NVT}} \tag{1.26}$$

Substituting equation (1.23) and (1.25) into equation (1.26), we get:

$$\begin{aligned}\Delta A_{a\to b} &= -\frac{1}{\beta}\ln\frac{\int\int \exp[-\beta H_b(\mathbf{x}, \mathbf{p}_x)]d\mathbf{x}d\mathbf{p}_x}{\int\int \exp[-\beta H_a(\mathbf{x}, \mathbf{p}_x)]d\mathbf{x}d\mathbf{p}_x} \\ &= -\frac{1}{\beta}\ln\frac{\int\int \exp[-\beta \Delta H(\mathbf{x}, \mathbf{p}_x)]\exp[-\beta H_a(\mathbf{x}, \mathbf{p}_x)]d\mathbf{x}d\mathbf{p}_x}{\int\int \exp[-\beta H_a(\mathbf{x}, \mathbf{p}_x)]d\mathbf{x}d\mathbf{p}_x}\end{aligned} \tag{1.27}$$

Using the definition of $P(\mathbf{x}, \mathbf{p}_x)$ in equation (1.24), we get:

$$\Delta A_{a\to b} = -\frac{1}{\beta}\ln\int\int \exp[-\beta \Delta H(\mathbf{x}, \mathbf{p}_x)]P(\mathbf{x}, \mathbf{p}_x)d\mathbf{x}d\mathbf{p}_x \tag{1.28}$$

Or:
$$\Delta A_{a\to b} = -\frac{1}{\beta}\ln < \exp[-\beta \Delta H(\mathbf{x}, \mathbf{p}_x)] >_a \tag{1.29}$$

where $<\ldots>_a$ means an ensemble average over configurations of the reference system.

In practice, equation (1.29) only holds when the two systems **a** and **b** are close enough, which means the configuration ensembles of **a** and **b** overlap appropriately, for numerical reasons. If this condition is not met, which is true in most practical cases, the transformation from **a** to **b** is usually divided into many small steps. In each small step, the overlap condition is met, and the free energy change from system **a** to system **b** is the sum of the free energy change in each step. For each step, an intermediate state is introduced. For a particular intermediate state **k**, the Hamiltonian reads:

$$H_k(\mathbf{x},\mathbf{p}_x) = (1-\lambda_k)H_a(\mathbf{x},\mathbf{p}_x) + \lambda_k H_b(\mathbf{x},\mathbf{p}_x) \tag{1.30}$$

λ_k is called the coupling parameter. For system **a,** $\lambda_k = 0$, and for system **b,** $\lambda_k = 1$. For a series of N systems, including **a** and **b** (H_1=H_a, $H_N = H_b$)**,** the free energy change for the transformation from **a** to **b** can be expressed as a sum of N-1 free energy differences:

$$\Delta A_{a\to b} = -\frac{1}{\beta}\sum_{k=1}^{N-1} \ln < \exp\{-\beta[H_{k+1}(\mathbf{x},\mathbf{p}_x) - H_k(\mathbf{x},\mathbf{p}_x)]\} >_k \tag{1.31}$$

Equation (1.31) is the basic FEP equation.

1.2.3 Umbrella Sampling

The umbrella sampling method is commonly used to calculate the free energy change when a system is transformed from state **a** to state **b**; for example, folded and unfold states of a protein. Suppose the reaction coordinate for that transition is q and the potential energy for the system U(**R**, q) is a function of the reaction coordinate and all other complementing coordinates designated by the vector **R**. The free energy as a function of q can be written:

$$F(q_0) = -\frac{1}{\beta}\ln\int \delta(q-q_0)\exp[-\beta U(\mathbf{R},q)]dqd\mathbf{R} \qquad (1.32)$$

Here, $\beta = 1/k_B T$, k_B is the Boltzmann constant and T is the temperature. The δ function is the Dirac's delta function. We omitted the kinetic part of the Hamiltonian for the simplicity of our discussion. Included or not, the kinetic energy part will not make a difference in the final conclusion anyway (see below).

We are usually interested in a free energy difference such as $F(q_0)$ - $F(q_1)$. For the calculation of free energy difference, we can add a constant to each free energy without changing the value of their difference. (It is also true when a constant is subtracted from each free energy; this is why we could omit the kinetic part, since the kinetic energy contributes equally to each free energy). Here, we add the total free energy of the system to the free energy at a particular reaction coordinate q_0, and $F(q_0)$ reads:

$$F(q_0) = -\frac{1}{\beta}\ln\left[\frac{\int \delta(q-q_0)\exp[-\beta U(\mathbf{R},q)]d\mathbf{R}dq}{\int \exp[-\beta U(\mathbf{R},q)]d\mathbf{R}dq]}\right] = -\frac{1}{\beta}\ln < \delta(q-q_0)_U > \qquad (1.33)$$

Equation (1.33) expresses the free energy as an average of a function, which can be calculated in a MD simulation. However, if some of the q configurations are high improbable, it will be very difficult to obtain reliable statistics for these q in a limited simulation. Also, a simulation can be stuck in some local minimum during a simulation. Figure 1.4 below shows such a case. The thick line in the plot is the potential energy and the thin line represents the positions sampled by the system during a simulation. In simulations, the configurations close to the energy barrier are not likely to be sampled due to their high energy. Thus, it is possible for the system to

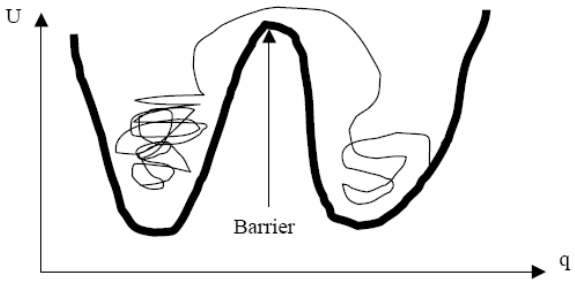

Fig 1.4 The projection of potential along the reaction coordinate *q*.

get stuck in a local energy minimum, such as the minimum on the left side of the barrier in Figure 1.4, so that the configurations on the right side of the barrier are not sampled sufficiently.

The umbrella sampling method modifies the original potential energy in such a way that configurations with desired reaction coordinates can be sampled significantly. The original potential is distorted by another potential term:

$$U(\mathbf{R},q) \rightarrow U(\mathbf{R},q) + V(q) \tag{1.34}$$

where $V(q)$ is independent of $\mathbf{R}$ and is chosen so that it is centered around a certain value of $q^{(k)}$. The potential V(q) is sometimes called a biasing potential. A commonly used biasing potential is a simple harmonic one:

$$V(q) = V^{(k)}(q) = k(q - q^{(k)})^2 \tag{1.35}$$

$q^{(k)}$ is a constant. It is clear that after the biasing potential is added, it is favorable for the simulation to sample configurations with reaction coordinate around $q^{(k)}$. By using different constant in the biasing potential, we can ideally obtain statistics of the configurations with any reaction coordinate value we are interested in. One obvious problem is that the modified potential is not the true potential and we must find a way

to convert sampling data with the biased potential to statistical data without the bias, if there is such a way at all. Fortunately, this can be done. Let's start with the integrals:

$$\frac{\int \delta(q-q_0)\exp[-\beta U(\mathbf{R},q)]d\mathbf{R}dq}{\int \exp[-\beta U(\mathbf{R},q)]d\mathbf{R}dq} = \frac{\int \delta(q-q_0)\exp[\beta V(q)]\exp[-\beta(U(\mathbf{R},q)+V(q))]d\mathbf{R}dq}{\int \exp[-\beta(U(\mathbf{R},q)+V(q))]d\mathbf{R}dq}\frac{\int \exp[-\beta(U(\mathbf{R},q)+V(q))]d\mathbf{R}dq}{\int \exp[-\beta U(\mathbf{R},q)]d\mathbf{R}dq} \tag{1.36}$$

Both sides are equal because we just added and subtracted $V(q)$ from the exponent and multiplied and divided by the same integral. Let's look at the first one of the two ratios:

$$\frac{\int \delta(q-q_0)\exp[\beta V(q)]\exp[-\beta(U(\mathbf{R},q)+V(q))]d\mathbf{R}dq}{\int \exp[-\beta(U(\mathbf{R},q)+V(q))]d\mathbf{R}dq} = < \delta(q-q_0)\exp[\beta V(q)] >_{U+V}$$
$$= \exp[V(q_0)] < \delta(q-q_0) >_{U+V} \tag{1.37}$$

Notice that the average of the delta function is an average over the configurations sampled with the total potential $U+V$. The second ratio can also be written as an average:

$$\frac{\int \exp[-\beta(U(\mathbf{R},q)+V(q))]d\mathbf{R}dq}{\int \exp[-\beta U(\mathbf{R},q)]d\mathbf{R}dq} = \frac{\int \exp[-\beta(U(\mathbf{R},q)+V(q))]d\mathbf{R}dq}{\int \exp[\beta V(q)]\exp[-\beta(U(\mathbf{R},q)+V(q))]d\mathbf{R}dq} = \frac{1}{< \exp[\beta V(q)] >_{U+V}} \tag{1.38}$$

Put (1.37) and (1.38) together, we have:

$$< \delta(q-q_0) >_U = \frac{\exp[V(q_0)] < \delta(q-q_0) >_{U+V}}{< \exp[\beta V(q)] >_{U+V}} \tag{1.39}$$

The numerator above can be calculated simply enough from a simulation. The denominator, however, cannot. The reason is that the configurations in a simulation are sampled with a weight of exp[-β(U+V)]=exp[-βU]exp[-βV]. So the weight is exponentially small in V while the function we try to average is exponential big in V. Thus, in the simulation, either the weight or the function will be close to zero and the sampling is problematic for the calculation of this average. We need to find a way to avoid the calculation of the denominator.

Suppose we use two biasing potential $V^{(k)}$ and $V^{(k+1)}$ in two separate simulations, both with its minimum close to but not the same as q_0 . Since both simulations potentially center around q_0, the value $<\delta(q-q_0)>_U$ can be calculated using sampling data from either of the simulations, and we have:

$$< \delta(q-q_0) >_U = \frac{\exp[\beta V^{(k)}(q_0)] < \delta(q-q_0) >_{U+V^{(k)}}}{< \exp[\beta V^{(k)}(q)] >_{U+V^{(k)}}} \tag{1.40}$$

And

$$< \delta(q-q_0) >_U = \frac{\exp[\beta V^{(k+1)}(q_0)] < \delta(q-q_0) >_{U+V^{(k+1)}}}{< \exp[\beta V^{(k+1)}(q)] >_{U+V^{(k+1)}}} \tag{1.41}$$

So we have:

$$\frac{\exp[\beta V^{(k)}(q_0)] < \delta(q-q_0) >_{U+V^{(k)}}}{< \exp[\beta V^{(k)}(q)] >_{U+V_{(k)}}} = \frac{\exp[\beta V^{(k+1)}(q_0)] < \delta(q-q_0) >_{U+V^{(k+1)}}}{< \exp[\beta V^{(k+1)}(q)] >_{U+V_{(k+1)}}} \tag{1.42}$$

Take the log of both sides, we get:

$$\begin{aligned} &\beta V^{(k)}(q_0) + \ln < \delta(q-q_0) >_{U+V^{(k)}} - \beta V^{(k+1)}(q_0) - \ln < \delta(q-q_0) >_{U+V^{(k+1)}} \\ &= \ln < \exp[\beta V^{(k)}(q)] >_{U+V^{(k)}} - \ln < \exp[\beta V^{(k+1)}(q)] >_{U+V^{(k+1)}} \end{aligned} \tag{1.43}$$

The right hand side of (1.43) is independent of q_0. For the above equation to hold for any q_0, the right hand side of the equation must be equal to a constant. Equation (1.43)

is also true for q_1, which is close to q_0, so we have:

$$\begin{aligned}&\beta V^{(k)}(q_1) + \ln < \delta(q - q_1) >_{U+V^{(k)}} - \beta V^{(k+1)}(q_1) - \ln < \delta(q - q_1) >_{U+V^{(k+1)}} \\ &= \ln < \exp[\beta V^{(k)}(q)] >_{U+V^{(k)}} - \ln < \exp[\beta V^{(k+1)}(q)] >_{U+V^{(k+1)}}\end{aligned} \tag{1.44}$$

The free energy of the system at reaction coordinates q_0 and q_1 is:

$$\begin{aligned}&F(q_0) = -\frac{1}{\beta}\ln < \delta(q - q_0) >_U \\ &= -V^{(k)}(q_0) - \frac{1}{\beta}\ln < \delta(q - q_0) >_{U+V^{(k)}} + \frac{1}{\beta}\ln < \exp[\beta V^{(k)}(q)] >_{U+V^{(k)}}\end{aligned} \tag{1.45}$$

And

$$\begin{aligned}&F(q_1) = -\frac{1}{\beta}\ln < \delta(q - q_1) >_U \\ &= -V^{(k)}(q_1) - \frac{1}{\beta}\ln < \delta(q - q_1) >_{U+V^{(k)}} + \frac{1}{\beta}\ln < \exp[\beta V^{(k)}(q)] >_{U+V^{(k)}}\end{aligned} \tag{1.46}$$

These two equations are true if we use a biasing potential $V^{(k+1)}$ is close to $V^{(k)}$.

If we are primarily interested in the free energy difference of the system at q_0 and q_1, we don't need to calculate all three terms on the right hand side of equation (1.45) or (1.46), but instead, just calculate the first two terms. The third terms would cancel each other since they are the same. This is also true when biasing potential $V^{(k+1)}$ is applied in the simulation. If we plot the quantity

$-\frac{1}{\beta}\ln < \delta(q - q_i) >^{U+V(k)} - V^{(k)}(q_i)$ for different q-s and also for different k-s, it will look something like:

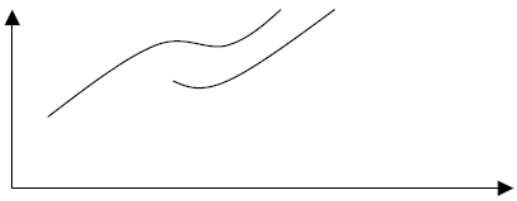

Fig 1.5 Free energy as a function of reaction coordinate q

The two curves correspond to two different biasing potentials. At the overlap region of these two curves, the differences between them at different *q*-*s* are a constant. Further more, these two curves are related to the real free energy by a constant. Though for different biasing potentials, the constant is different. By sliding the biasing potential along different positions of the reaction coordinate, we may get the free energy profile over a range of reaction coordinates. This process may look pictorially something like:

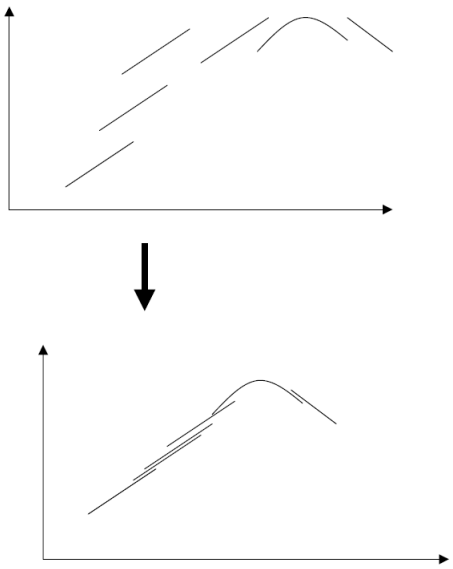

Fig 1.6 The process of constructing the free energy profile along the reaction coordinate.

1.2.4 The Weighted Histogram Analysis Method (WHAM).

From an umbrella sampling simulation with biasing potential $V_{bias}(q)$, the biased probability to find the system at a particular configuration (characterized by reaction coordinate *q*) is P'(*q*) and the unbiased free energy could be evaluated as following:

$$F(q) = -k_B T \ln P'(q) - V_{bias}(q) + Const \tag{1.47}$$

The "Const" term is not known and it depends on the biasing potential. In practice, a number of simulations are run with the minimum of the biasing potential at a different q for each simulation. From each of these simulation, a biased probability P'(q) for a particular reaction coordinate value q could be estimated using a histogram method, in which the range of sampled reaction coordinate is discretized into a number of bins with equal width. Each P'(q) from an individual simulation will carry its own error. There is a need to combine the results from all simulations in an appropriate way to give a best estimation for the unbiased probabilities P(q).

The Weighted Histogram Analysis Method (WHAM)[28] is an extension of Ferrerberg and Swendsen's Multiple Histogram Technique[29]. The basic idea of this method is to weight the estimations for the density of states from each simulation in such a way so that these estimations with smaller errors get more weight. However the error associated with each estimation is not known and it is related to the true density of states itself. The WHAM equations shown in (1.48) and (1.49) provide a method to best estimate the unbiased probability distribution of states:

$$P(q) = \frac{\sum_{i=1}^{N_{sims}} n_i(q)}{\sum_{i=1}^{N_{sims}} N_i \exp[(F_i - V_{bias}(q))/k_B T]} \tag{1.48}$$

$$F_i = -k_B T \ln\{\sum_{N_{bins}} P(q) \exp[-V_{bias}(q)/k_B T]\} \tag{1.49}$$

where N_{sims} is the number of simulations, $n_i(q)$ is the number of counts in histogram bin associated with q, F_i is the free energy of the system in simulation i. and $P(q)$ is the best estimation of unbiased probability distribution. F_i and P(q) are not known and

the above two equations are solved by iteration until self consistency is reached.

1.3 Lipid Rafts in Cell and Model Membranes

The concept of lipid rafts was formulated about twenty years ago from studies on lipid polarity and lipid sorting in epithelial cell[30,31] . The main postulate was the existence of lipid domains, consisting of dynamic assemblies of cholesterol, sphingolipids and certain proteins, in the exoplasmic leaflet of the cell membrane. This concept was popularized in 1997[32] and since, the field has undergone wide acceptance and expansion. Lipid rafts have been believed to be involved in many cellular processes such as signaling[33-35], protein and lipid sorting[36], cellular entry by toxins and viruses and viral budding[37,38] etc. A consensus definition of lipid rafts was given in the recent Keystone Symposium[39] as "Membrane rafts are small (10-200 nm), heterogeneous, highly dynamic, sterol- and sphingolipid-enriched domains that compartmentalize cellular processes. Small rafts can sometimes be stabilized to form larger platforms through protein-protein and protein-lipid interactions".

The first lipid domains were detected in human and hamster fibroblasts as detergent-resistant membranes (DRMs) that were enriched of glycosphingolipids, cholesterol and glycoproteins[40,41]. The term "lipid raft" is defined operationally by the way in which they are extracted from cells using a non-ionic detergent Triton X-100 at 4℃. These rafts float to a low density during a gradient centrifugation due to their high lipid content[42,43]. Proteins found to be associated with these sphingolipid-cholesterol rafts include glycosylphosphatidylinositol (GPI)-anchored

proteins and trans-membrane proteins, like the influenza virus proteins neuraminidase and haemagglutinin[32]. The operational definition of lipid rafts had led to some controversies on their existence in living cells as the small size of these rafts preclude the possibility of direct observation by light microscopy. However, many recent studies with improved methodologies, such as fluorescence resonance energy transfer measurements[44], biochemical crosslinking of GPI-anchored proteins[45], antibody crosslinking of raft proteins into patches segregating from non-raft proteins[46], photonic force microscopy measurements[47] and electron microscopy visualization[48] have provided more and more evidence on the existence of lipid rafts. Nevertheless, the size of different domains, the mechanisms behind the formation of domains and the dynamics and the structural properties of these domains are far from clear. Are proteins necessary for the formation of a domain, or they are just recruited after the domains have formed. Are there corresponding domains in the inner leaflet of a cell membrane, which are primarily enriched in phosphatidylethanolamines (PE) and unsaturated phosphatidylcholines (PC)? If yes, are they coupled with the domains in the exoplasmic leaflet? How?

Since biological membranes are very complex entities consisting of different kinds of phospholipids, glycolipids, sterols and proteins, it is very difficult to study such a system and understand the role each component plays in the functioning of such an assembly. In practice, simple model membranes containing one species of lipid or well-controlled mixtures of two or three components are usually synthesized for study. Experimental observations have demonstrated that physical properties of

these model membranes are similar to the physical properties of natural membranes. In both natural and synthetic membranes, the lipid rafts domains are enriched in cholesterol and saturated lipids, like sphingolipids. Sphingolipids differ from many biological lipids in containing mostly saturated acyl tails. This allows for them to pack tightly together, which is one of the reasons why sphingolipids have a much higher melting temperature (T_m)[32] than biological phospholipids (like phosphatidylcholine, or PC), which usually carry kinked, unsaturated acyl chains. Researchers have shown that the ability for sphingolipids to pack their acyl tails tightly is a key feature in raft forming[38,49,50]. In fact, at physiological temperature, a model membrane made of pure saturated sphingolipids would be in a gel phase (usually abbreviated as s_o), while one made of unsaturated phospholipids would be in a disordered state (l_c or l_d). In the gel phase, tails of lipid molecules are highly ordered and lipid molecules have very low mobility. In the liquid disordered phase, the lipid tails are disordered and have a high degree of lateral mobility.

In lipid rafts, the lipid tails are also highly ordered similar to a gel phase, however, due to the high concentration of cholesterol molecules, molecules in a lipid rafts have a lateral mobility close to those in a l_d phase, and such a phase is usually termed a liquid ordered phase (l_o). Lipid rafts are in a liquid ordered phase due to the special role cholesterol molecules play in forming such a phase. It is well known that cholesterol has important effects on the phase behavior of certain lipids. Different binary mixtures of cholesterol with lipids as well as ternary mixtures containing cholesterol have been studied, which brought up many interesting observations as

well as questions. In experiments, McConnell and collaborators[51,52] observed that for monolayers containing a binary mixture of cholesterol/ di(10:0) PC or cholesterol/di(12:0)PC or cholesterol/di(14:0)PC at low surface pressures, two coexisting liquid phases were present. At higher surface pressures, the two phases merged into a single phase, displaying an upper miscibility critical point. When the mixture contained cholesterol and, for example, di(15:0) PC, two upper critical points appeared. The phase diagrams for monolayers of cholesterol/di(14:0) PC and cholesterol/di(15:0) PC mixtures are shown in Fig. 1.7:

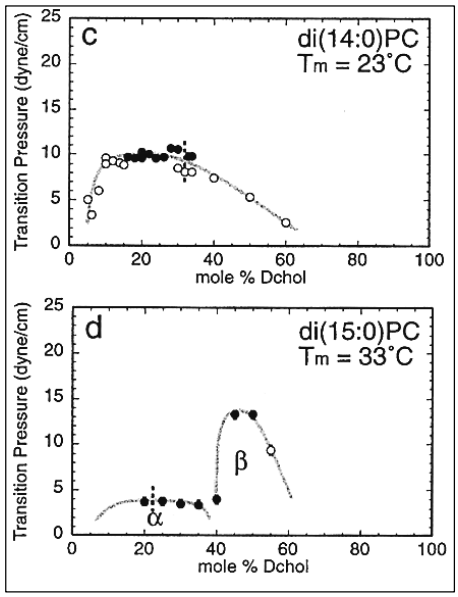

Fig 1.7 The phase diagrams for monolayers of chol/di(14:0)PC and chol/di(15:0)PC mixtures

As can be seen from this figure, there are two regions which are two phase regions in the cholesterol/di(15:0) PC mixture. According to McConnell and his collaborators, the existence of two critical points in the phase diagram indicates the formation of a "condensed complex" between cholesterol (C) and phospholipid (P). This process can be described by the reaction

$$nqC + npP \Leftrightarrow C_{nq}P_{np}, \qquad (1.50)$$

where q and p are stoichiometric integers and n is the measure of the size of the complex. n also shows the degree of cooperativity in the complex formation.

Using the notion of the condensed complex, we can assign one region of the phase diagram, region α, as due to immiscibility of pure phospholipid and complex and the other region, β, as due to immiscibility of cholesterol and complex. What is the molecular structure of the complex? It was emphasized in the work of McConnell's group that no molecular picture of a cholesterol/phospholipid complex exists at the present time. There is also no clear understanding what are the values of q, p and n, although some indications are that q=1 or 2 and p=2 or 3 and n~2-10.

Micron-scale liquid domains and liquid-liquid phase separation in artificial membranes were observed in giant unilamellar vesicles (GUVs)[53] and in black lipid membranes[54] when these systems contained three lipid components. The two liquid phases in these systems were the liquid ordered (l_o) phase that constitutes the domains (rafts) and the liquid disordered (l_d) phase in the rest of the membrane. One of the ternary mixtures initially studied was a mixture containing a relatively saturated lipid such as sphingomyeline (SM), cholesterol and unsaturated lipid such as dioleoylphosphatidylcholine (DOPC) at ratio 1:1:1. The lipid raft domains were observed to be enriched by SM and cholesterol, while the liquid disordered phase outside the rafts was enriched in DOPC[55]. Later, the phase diagram for this mixture was mapped out and showed the existence for the region of liquid-liquid phase coexistence (see Fig. 1.8a for the diagram; the inserts are micrographs of vesicles. The diagram is from the review by Veatch and Keller[56]). The choice of SM molecule as a saturated lipid in a mixture perhaps was dictated by the observation that natural membranes contain a large proportion of SM. One can choose another saturated lipid,

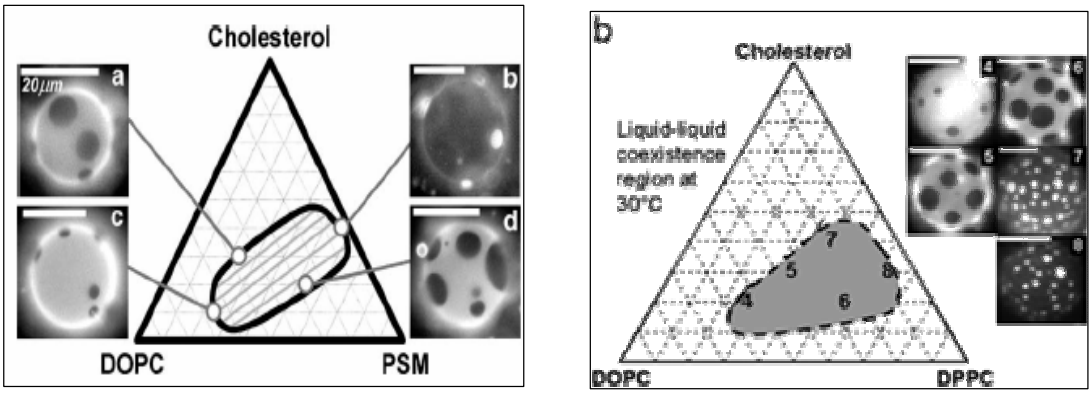

Fig 1.8 Phase diagrams of ternary mixtures containing cholesterol and phospholipids.

such as the well studied dipamitoylphosphatidylcholine (DPPC) instead and still observe the liquid-liquid coexistence region in the phase diagram for a ternary mixture of DOPC/DPPC/CHOL (see Fig 1.8b). Liquid-liquid phase separation can also be observed in the phase diagram of the DPPC/DLPC/CHOL mixture. What kind of interactions between cholesterol and phospholipids are responsible of the liquid-liquid coexistence regions in the phase diagrams shown above? Are they very specific? What is the role of the phospholipid headgroup/cholesterol interaction or the phospholipid tail/cholesterol interaction? To answer these questions, one needs to look into molecular details of the interactions between these molecules and it is very hard to measure these data from experiments by using microscopy or calorimetric methods. Since molecular dynamics simulations can provide atomic level resolution pictures of the system under study, performing computer simulations of lipid mixtures containing cholesterol will help to shed some light on these questions.

CHPATER 2: A SIMULATION OF ASYMMETRIC LIPID BILAYER

2.1 Introduction

Biological membranes are very complex entities containing many distinct lipid species[57]. The distribution of these lipids in cell membranes is not homogeneous. Thus, for example, lipids such as sphingomyelin (SM) and phosphatidylcholine (PC) can be often found in the outer leaflet of plasma membranes, while two other typical lipids-phosphatidylserine (PS) and phosphatidylethanolamine (PE)-are found in the inner leaflet of the membrane[58,59]. Another important lipid found in membranes is cholesterol (CHOL), but it is not known if it is distributed in the same amount in the two leaflets of the membrane[58]. In view of the inhomogeneous character of natural membranes, it was proposed that these membranes contain domains where lipid exist in a liquid ordered phase (l_o) surrounded by assemblies of lipids in the liquid-disordered phase (l_d)[32]. The l_o domains are rich in cholesterol and saturated lipids; these domains are also often called lipid rafts[32,60]. The issues related to our understanding of the composition, structure, dimensions, and properties of lipid rafts in natural biomembranes are very far from being clarified[61].

To study properties of lipid rafts, model membranes are intensely investigated. Numerous studies have been performed on bilayers containing binary mixtures of lipids with cholesterol as one of the components[56,62]. Specifically, in view of the

importance of the SM molecule for rafts, recent studies were done to understand the nature of interactions between SM and CHOL[63-65]. Also, studies were done on model bilayers containing ternary mixtures of cholesterol, SM, and unsaturated PC and phase diagrams for these mixtures were mapped out[56,66]. It was observed that in giant unilamellar vesicles containing the three components mentioned above, a phase separation occurs: the bilayers contain l_o domains (rafts) where CHOL and SM can be found in an enriched amount. The composition of both leaflets in model systems was the same and it was observed that l_o domains were created simultaneously in both leaflets and the domains were in the same location in the inner and outer leaflets. What is the situation with the domains in natural membranes containing asymmetric (different in lipid composition) leaflets? As of today, no clear understanding of this issue exists[58].

Computer simulations of bilayers containing PC lipids or mixtures of PC with cholesterol provided molecular level information on structural and dynamical properties of such bilayers[67-72]. Recently, results were reported on simulation studies done on SM bilayers[73-76] as well as on bilayers containing a binary mixture of SM and CHOL[77]. Some preliminary, but thought-provoking simulation work was done on ternary mixtures containing dioleoylphosphatidylcholine, SM, and CHOL by Pandit et al[78,79].

Nearly all simulations on bilayers containing lipid mixtures that were reported in the literature were performed on symmetric bilayers containing the same composition in both leaflets. In this chapter, a molecular dynamics (MD) simulation performed on

an asymmetric bilayer containing a mixture of CHOL and (18:0) SM in one leaflet and stearoyl-oleoyl-phosphatidylserine (SOPS) and CHOL in the other leaflet is reported. For comparison purposes, two simulations on symmetric bilayers: first simulation was performed on a bilayer containing a binary mixture of SOPS and CHOL, with second containing a mixture of (18:0) SM and CHOL were also performed. We chose to simulate a mixture of PS with CHOL because PS molecules interact more favorably with CHOL compared to phophatidylethanolamine molecules[60]. SOPS molecule is chosen to represent PS molecules, which are found in the cytoplasmic leaflet of natural membranes[80]. The raft-forming concentration of cholesterol in the SM+CHOL bilayer was chosen for the study. Since the interaction between PS and CHOL is favorable and the PS molecules are condensed in the bilayer due to interlipid hydrogen bonding interactions and condensing effect due to conterions[81,82], we expect that the phase of the SOPS and CHOL mixture will also be the l_o phase. Therefore, we expect that in our simulation we represent a patch of and l_o domain in an asymmetric bilayer.

2.2 Computational Details

We have performed MD simulations on two symmetric bilayers containing mixtures of phospholipid molecules with cholesterol and on simulation on an asymmetric bilayer containing different mixtures in each leaflet. The first bilayer contained 84 SOPS molecules, 44 cholesterol (CHOL) molecules, and 84 Na^+ counterions. The second system had 84 (18:0) sphingomyelin (SM) molecules and 44

CHOL molecules. Both bilayers were hydrated with 3840 water molecules. The molecular structures of SOPS, (18:0) SM, and CHOL were generated using SYBYL, Ver 7.0 (Tripos, St. Louis, MO) and are shown in Fig. 2. 1. The initial structure of a bilayer leaflet was obtained by generating an 8 X 8 array of 64 lipids (SOPS or (18:0) SM) in the *xy* plane by random rotation of each lipid around the *z* direction. Then 22 of these phospholipids were randomly chosen and replaced by CHOL molecules. The second leaflet was obtained by reflection and translation of this first layer. A water slab was added on both sides to solvate the headgroups. In the case of SOPS bilayer, this was followed by a random replacement of 42 water molecules on each side of the bilayer by Na^+ ions. For both bilayers, the bilayer normal is directed along the z direction.

Our simulations were performed using the GROMACS package[5,6]. The LINCS algorithm was used to constrain all bonds in the system[83] allowing an integration time step of 2 fs. Periodic boundary conditions were applied in all three dimensions and long-range eletrostatics was handled using the SPME algorithm[10] with real-space cutoff of 10 Å, fourth-order interpolation, and a tolerance of 10^{-5}. A 12 Å cutoff was utilized for van der Waals' interactions. The temperature in the simulations was maintained at 310 K using the Nosé-Hoover scheme[84] with a thermostat relaxation time of 0.5 ps. The system was simulated in an NPT ensemble using the Parrinello-Rahman semi-siotropic pressure coupling scheme[85] with a barostat time constant of 2.0 ps at a pressure of 1 atm.

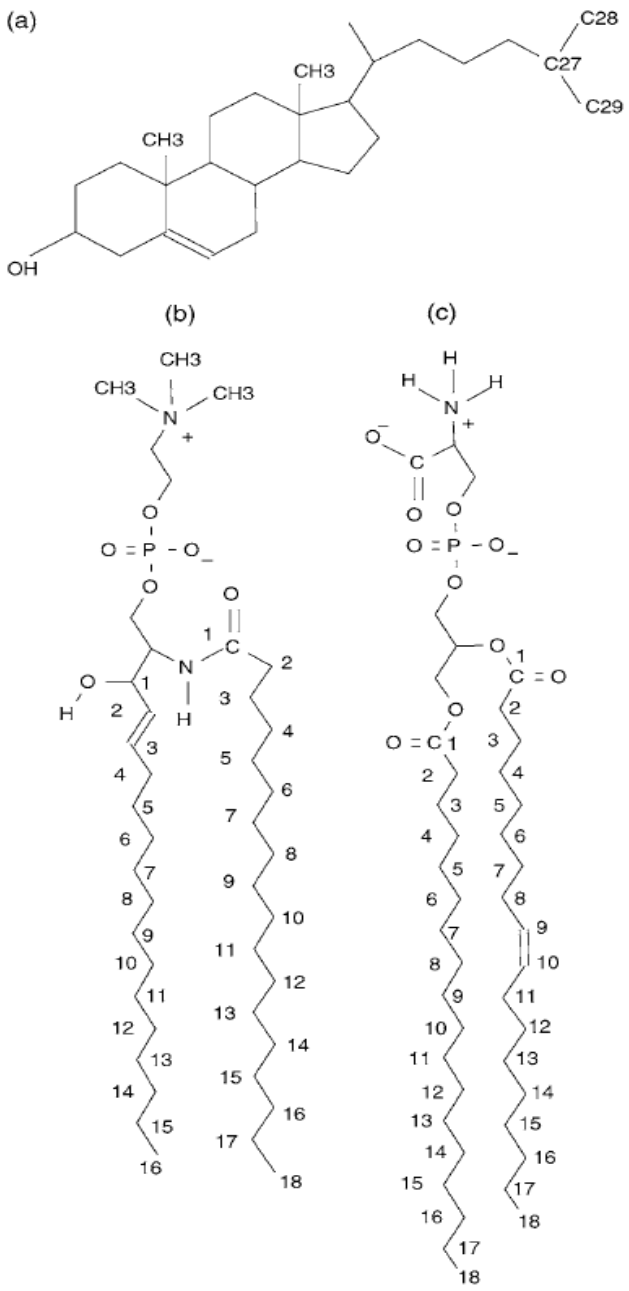

Fig 2.1 Structure of the lipid molecules (a) cholesterol (CHOL). (b) (18:0) sphingomyelin (SM), and (c) SOPS.

The SPC/E model of water[86] was used in the simulations. Force field for the SOPS was based on the parameters of Berger et al.[87] and GROMOS87[88] parameters. The

carboxylate group charges were taken from the aspartic acid side chain. The partial charges for amine group were the same as used in palmytoyl-oleoyl-phosphatidylcholine[89]. The force field for the (18:0) was the same as used by Niemelä[90] et al. The force field parameters of CHOL as used in the study by Pandit et al[72]. were also used in this study.

The SOPS and (18:0) SM bilayers were simulated for 60 ns and 64 ns respectively. The positions and velocities of the system were saved every 1 ps. To construct the asymmetric bilayer we combined configurations of two lipid leaflets, each from the simulation with the symmetric SM+CHOL and SOPS+CHOL bilayers. A snapshot of the asymmetric bilayer is shown in Fig 2.2.

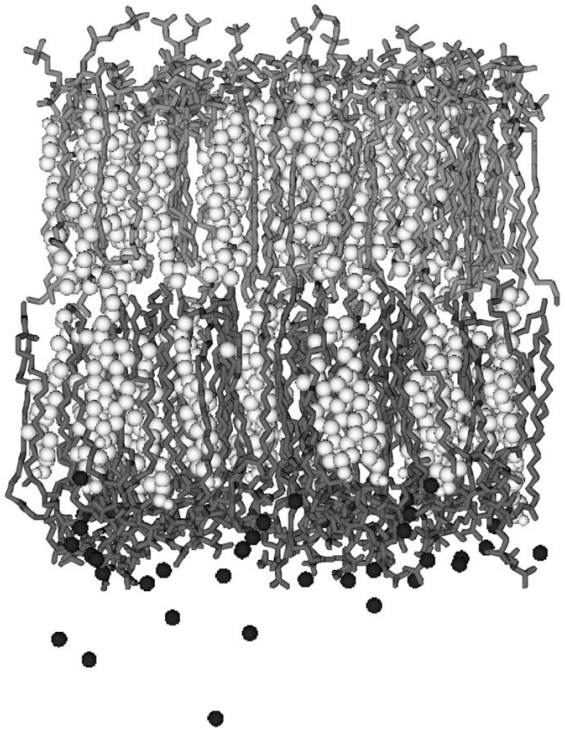

Fig 2.2 The asymmetric bilayer: SOPS (solid-stick representation), CHOL (open spheres), Na^+ (solid spheres) in the lower leaflet, SM (shaded-stick representation), and CHOL (open spheres) in the upper leaflet. The solvent water molecules are not

shown.

The configurations were obtained from the end of 50 ns and 60 ns MD runs, respectively. Both leaflets contained 42 phospholipid and 22 CHOL molecules. In addition, the SOPS+CHOL leaflet contained 42 Na^+ counterions. A water slab of 1920 water molecules was generated using GROMACS utility genbox and added to each leaflet of the bilayer. The asymmetric bilayer structure was initially energy-minimized and then an MD simulation was carried out for 50 ns in an NPT ensemble at 1 atm pressure and at 310 K with the three-dimensional periodic boundary conditions. For the asymmetric bilayer, the force field parameters and the simulation parameters were the same as described above for the symmetric bilayers. The trajectory data were saved every 1ps.

The analyses of saved data were performed using the utilities available in GROMACS as well as programs written by us. The analyses have been carried out over the last 45 ns and 38 ns of the trajectory for symmetric SOPS and (18:0) SM bilayers, respectively. For the asymmetric bilayer, the analyses were carried out over the last 45 ns of the trajectory. Henceforth, we refer to (18:0) sphingomyelin as SM, unless mentioned otherwise.

2.3 Results and Discussion

2.3.1 Structural properties

Fig. 2.3 (a) shows the values for the areas per lipid from the two simulations

performed on the symmetric bilayers.

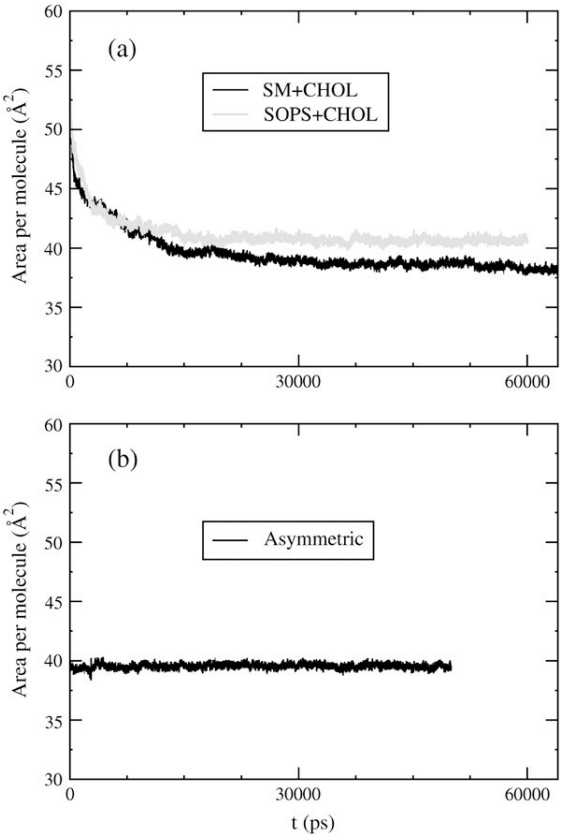

Fig 2.3 The area per molecule for (a) symmetric SOPS+CHOL and SM+CHOL bilayers and (b) asymmetric bilayer.

The values for the average area per lipid are 40.7±0.3 and 38.6±0.3 Å^2 for the bilayers containing SOPS and SM phospholipids, respectively. These numbers seem to be somewhat small, when compared to areas per lipid of ~50-70 Å^2 measured in monolayers or bilayers containing only one lipid component[91]. Since we have cholesterol in our bilayers, the area condenses. That the numbers given above are reasonable can be understood from the following argument. Let us assume that cholesterol is shielded from water by phospholipids, as it is suggested in the umbrella

model[92]. In this case one gets that the area per phospholipid is $a_{PL} = A/N_{PL}$, where a_{PL} is the area per headgroup of the phospholipid (SOPS or SM), A is the xy area of the simulation box, and N_{PL} is the number of phospholipids. From the umbrella model one gets for phospholipids area values of 62 $Å^2$ for SOPS and 58.8 $Å^2$ for SM. The value fro SOPS is close to the measured value of $\approx$ 64 $Å^2$ obtained for dioleoyl-phosphatidylserine (DOPS)[93]. It is expected that the area per SOPS should be substantially smaller than the one for DOPS, because SOPS only has one unsaturated acyl tail. The area per SM as estimated above is also larger than the area per SM of $\sim$ 51 $Å^2$ obtained from simulations on pure SM bilayer[90]. These considerations indicate that the umbrella model is probably not accurate for area estimation. Getting the values for area per lipid in simulations with lipid mixtures is not a simple issue[94]. Here we have used the methodology developed by Hofsäβ et al.[71] to compare our data with some of the data published in the literature for the cholesterol/dipamitoylphosphatidylcholine (DPPC) mixture. According to Hofsäβ et al., the areas per phospholipid and cholesterol can be determined using equation (2.1):

$$a_{PL} = \frac{2A}{(1-x)N_{lipid}}[1 - \frac{xN_{lipid}V_{chol}}{V - N_w V_w}] \qquad (2.1)$$

Where N_{lipid} is the total number of lipids ($N_{PL} + N_{CHOL} = 128$), x is the concentration of CHOL: N_{CHOL}/N_{lipid}, V is the volume of the simulation box, N_w is the total number of water molecules in the system, V_w is the volume occupied per water molecule (0.0305 nm^3), and V_{chol} is the volume per cholesterol molecule taken to be 0.593 nm^3.[71] The area per cholesterol molecule can be calculated from the expression:

$$a_{chol} = \frac{2A - a_{PL}N_{PL}}{N_{chol}} \tag{2.2}$$

The calculated values of area per lipid using the above equations are a_{SOPS} = 48.8 Å^2 and a_{SM} = 46.3 Å^2. The area per cholesterol is 25 Å^2 in both systems. For comparison, notice that in the DPPC+CHOL bilayer containing 40% cholesterol[71], the DPPC area is ~55 Å^2, while the CHOL area is ~27 Å^2. So when 40% of cholesterol is added to DPPC, its area in simulations shrinks from 64 Å^2 to ~55 Å^2, while the area of SM shrinks from ~51 Å^2 to ~46 Å^2 when we add 33% of cholesterol. This indicates that in membranes containing PC a substantial condensation occurs with addition of cholesterol. Some condensation also occurs in membranes containing SM, but this condensation is not as substantial. The bilayer of pure SM is already condensed due to the presence of a strong interlipid hydrogen-bonding network. The same can be also argued about the bilayers containing PS.

Fig. 2.3 (b) shows the time evolution of the area per lipid in the asymmetric bilayer. The average area per lipid molecule is 39.5±0.2 Å^2 in this case. To find the area per lipid in case of the asymmetric bilayer we observed that the *xy* area of the simulation box for the asymmetric bilayer is the same for the two leaflets and that both leaflets have the same number of phospholipid and cholesterol molecules. Therefore, from the above-described procedure for calculating area, if one assumes that the area/cholesterol is the same in two leaflets, one obtains that the area/phospholipid is also the same for the two types of phospholipids. The area/phospholipid is found to be 47.4 Å^2 and the area/cholesterol is ~25Å^2. Thus, the area/phospholipid in the asymmetric bilayer is intermediate to the values of the

area/SOPS and the area/SM as found above.

Fig 2.4 (a) and (b) show the electron density plots for the SOPS+CHOL system.

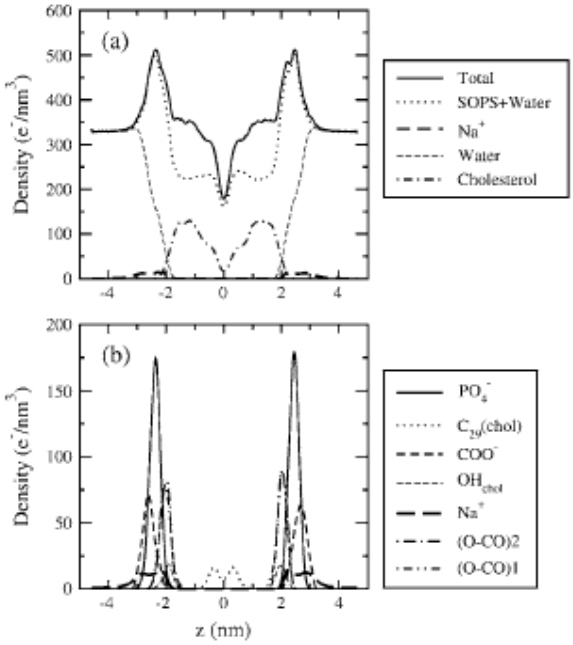

Fig 2.4 The electron density profile of the SOPS+CHOL bilayer plotted as a function of the distance long the bilayer normal.

The peak-to-peak distance, as estimated from Fig 2.4 (a), is found to be 48.2 Å. When compared to the experimental measurement of bilayer thickness of pure SOPS, which is about 42 Å[95], the SOPS+CHOL bilayer in the present work appears to be more extended. From Fig 2.4 (b), it can be seen that Na^+ ions are delocalized over the bilayer water interface. The ions are located in the region stretching from the carboxylate group to the ester-carbonyl groups denoted as (O-CO)1 for the *sn*-1 chain and (O-CO)2 for the *sn*-2 chain. No significant overlap of densities in the *z* direction is observed between hydroxyl groups of cholesterol and the Na^+ ions. However, the hydroxyl groups of cholesterol are seen to be in the proximity of ester-carbonyl

groups of SOPS. The electron density of the tail CH3 group of cholesterol shows very little overlap across the bilayer. This can be seen as a consequence of the presence of an 18-carbon tails in the SOPS molecules.

Fig 2.5 (a) and (b) are the electron density plots of the SM+CHOL system. The peak-to-peak distance in Fig. 2.5 (a) is found to be 47.6 Å. Khelasvili and Scott[77] also performed simulation of an 18:0 SM with cholesterol in a proportion which was roughly 2:1. Their areas per lipid were ~10% larger than what we found here and correspondingly their peak-to-peak distance was significantly shorter (42 Å). Note that the force field and the simulation temperatures were different in the simulations of Khelashvili and Scott. As in the case of SOPS bilayer, we did not observe any significant overlap of the tail methyl group densities (C_{29}) of cholesterol molecules in the case of the SM+CHOL bilayer (see Fig 2.5 (b)). At the same time, the density of the hydroxyl group of cholesterol is found to have a significant overlap with the density of the CO and –OH groups of the SM molecules.

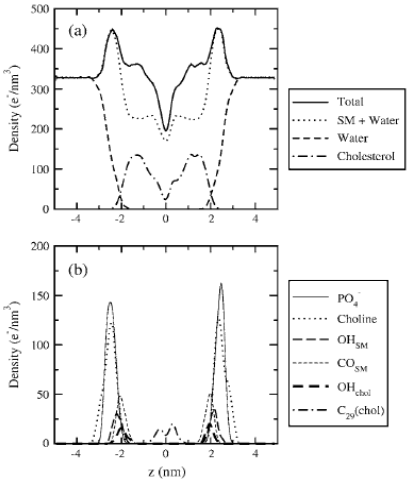

Fig 2.5 The electron density profile of the SM+CHOL bilayer plotted as a function of the distance along the bilayer normal.

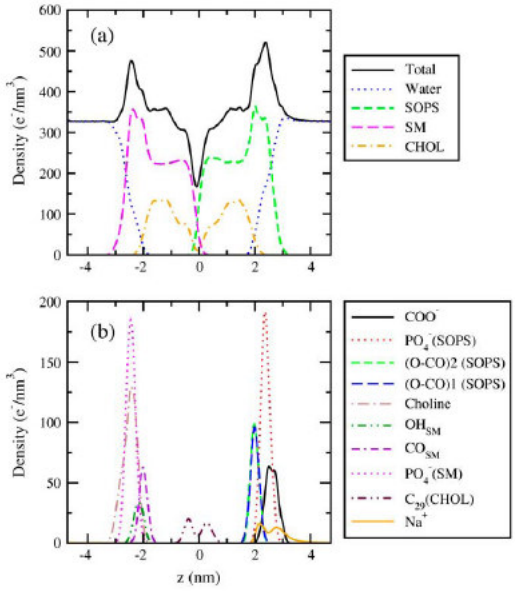

Fig 2.6 The electron density profile of the asymmetric bilayer plotted as a function of the distance along the bilayer normal

Fig 2.6 (a) and (b) are the electron density profiles of the asymmetric bilayer. The peak-to-peak distance in Fig.6 (a) is found to be 48.4 Å. The Na^+ ions were found to be located only on the side of the SOPS+CHOL leaflet, despite the presence of the three-dimensional periodic boundary conditions. These ions were delocalized over the interfacial region from carboxylate to the ester-carbonyl groups. As in the symmetric bilayers, tail groups C_{29} of CHOL were found to have a minimal overlap in the asymmetric bilayer (Fig 2.6 (b)).

We also calculated the deuterium-order parameters[96] for the lipid tails, which are defined as following:

$$S_z = \frac{3}{2} < \cos^2 \theta_z > -\frac{1}{2} \qquad (2.3)$$

Where θ_z is the angle between the z axis of the simulation box and the molecular axis

under consideration. The molecular axis is defined as the vector from C_{n-1} to C_{n+1}. The parameters S_x and S_y are defined in the same way. The brackets denote average over time and molecules. Order parameters can vary between 1 (full order along the bilayer normal) and -1/2 (full order perpendicular to the normal), with a value of zero in the case of isotopic orientation. The lipid chain order parameters are shown in Fig 2.7

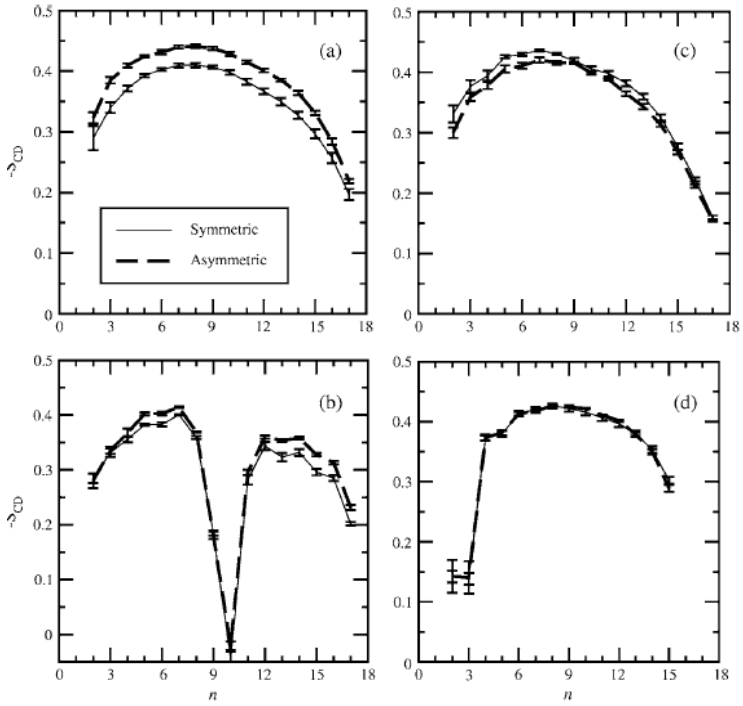

Fig 2.7 Deuterium order parameter profiles for symmetric and asymmetric bilayers. (a) Stearoyl and (b) oleoyl chains of SOPS. (c) Stearoyl and (d) sphingosine chain of SM. *n* is the number of carbon atoms along the hydrocarbon chains as shown in Fig. 2.1

The order parameters for the lipid chains in the symmetric bilayers indicate that the sphingosine and the stearoyl chains of SM are slightly more ordered than the stearoyl chain of SOPS. Fig. 2.7 (a) and (b) show that the effect of the bilayer asymmetrization in our simulation was most significant for the SOPS+CHOL leaflet. The stearoyl chain and, to some extent the oleoyl chain, of the SOPS became more

ordered in the asymmetric bilayer than in the symmetric bilayer. The ordering of the sphingosine chain of SM remained unaffected by the asymmetrization (Fig. 2.7 (d)), while the stearoyl chain became slightly less ordered (Fig. 2.7 (c)). Thus, for the symmetric bilayers, the SM+CHOL bilayer exhibits a lower average area/molecule and therefore a larger lipid chain order than the SOPS+CHOL bilayer. In the asymmetric bilayer, the SOPS+CHOL gets more ordered while the order in the SM+CHOL leaflet slightly decreases as compared to the symmetric bilayer. This is consistent with the observed trend of area/lipid (see Fig.2.3).

The results of our simulations on symmetric bilayers demonstrate that geometrical parameters such as areas per lipid and chain order parameters (for saturated bonds) are very close in cases of SOPS+CHOL and SM+CHOL bilayers. Does this mean that if the SM+CHOL bilayer is the l_o phase, the phase of the SOPS+CHOL bilayer is also an l_o phase and that SOPS and CHOL mixture if also raft-forming? If we follow a previously proposed suggestion that raft-forming tendencies are connected to cholesterol-lipid complex-forming tendencies[51], and these in turn are correlated with hydrogen-bonding network properties[72] we need to study properties of hydrogen bonding network in our bilayers.

2.3.2 Properties of Hydrogen-bonding Network

Let us consider here the interlipid hydrogen-bonding properties of the bilayers. We use the geometric criteria for the definition of the hydrogen bond between two lipids. To this end, for SOPS+CHOL bilayer, a radial distribution function was

calculated between hydrogen atoms of groups such as $NH3^+$ of SOPS and –OH of CHOL and the oxygen atoms belonging to different molecules in the system. The first minima for these radial distribution functions were found to be about 2.5 Å. Therefore, an interlipid hydrogen bond was assumed to exist when the distance between a hydrogen atom from one lipid and an oxygen atom from the other lipid, r_{HO} is <2.5 Å and the angle hydrogen-donor-oxygen, θ_{HDO}, is <30°. In the present work, the donor can either be a nitrogen or an oxygen atom. This definition of hydrogen bond is similar to the criterion used by Mukhopadhyay et al.[97], who had used $r_{HO} \leq 2.4$ Å and $\theta_{HDO} \leq 35°$. In the case of the pure SM bilayer, it was shown recently[76] that among –OH and –NH groups only the –NH group participates in the interlipid hydrogen bond. In addition to the definition of the hydrogen bond given above, to study the interlipid hydrogen bonding in the SM+CHOL mixture, we assume that the hydrogen bond can be made between the $N^+(CH_3)_3$ group of the SM and the OH group of cholesterol. The existence of such a bond was proposed when interlipid hydrogen bonding was studied for the bilayer containing dipalmitoylphosphatidylcholine (DPPC) and cholesterol mixture[72]. The concept of CH…O hydrogen bond is well established in chemistry[98-101]. The geometric criterion used for identifying this interlipid hydrogen bond for the united-atom model of CH3 was also used in the earlier work on lipid cmoplexation[82]. The hydrogen bond between $N^+(CH_3)_3$ and $-OH_{CHOL}$ is assumed to exist whenever the distance between N-CH_3-O is in the range between 79° and 139°. This criterion is also used to identify hydrogen-bonding interaction between the two charged groups, $N^+(CH_3)_3$ and PO_4-, belonging to two different SM molecules.

Fig. 2.8 shows the distributions of the number of molecules hydrogen-bonded to a given lipid in the SOPS+CHOL bilayer.

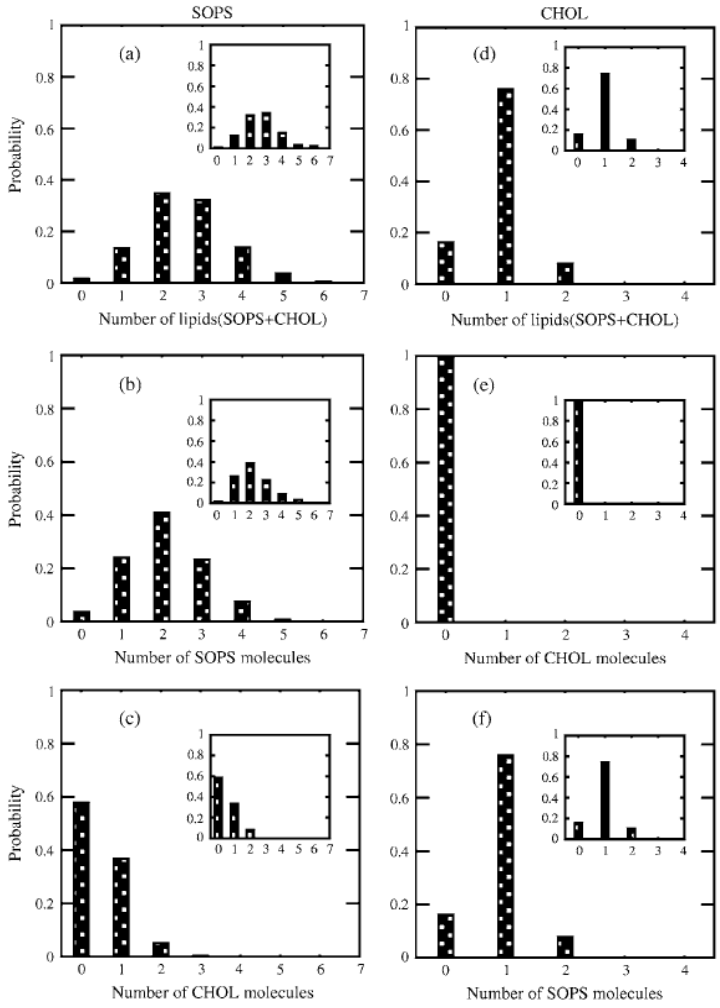

Fig 2.8 Distribution of molecules that are hydrogen-bonded to a given lipid in SOPS+CHOL bilayer. Distribution of (a) total number of lipids. (b)total number of SOPS molecules, and (c) total number of cholesterol molecules that are hydrogen-bonded to a given SOPS molecule. Panels *d-f* show the distribution of total number of lipids, total number of cholesterol molecules, total number of SOPS molecules, respectively, that are hydrogen-bonded to a given cholesterol molecule. Inset in each plot show the corresponding distribution of molecules that are

hydrogen-bonded to a given lipid in the SOPS+CHOL leaflet of the asymmetric bilayer.

As can be seen from Fig. 2.8 (a), an SOPS molecule has maximum probability to be hydrogen-bonded with two other lipids, which according to Fig. 2.8 (b) will be, with a high probability, two SOPS molecules. Also, an SOPS molecule is most likely not to be hydrogen-boned to a cholesterol, as cholesterol concentration is 33% (see Fig. 2.8 (c)). In addition, SOPS is not found to be hydrogen-bonded to more than two cholesterol molecules. The distribution for the total number of hydrogen bonds per cholesterol is dominated by the presence of just one hydrogen bond between the SOPS molecule and cholesterol (see Fig. 2.8, (d) and (f)). Notice also the absence of the cholesterol-cholesterol hydrogen bonding.

Fig.2.9 shows the distribution of a number of molecules that are hydrogen bonded to a lipid in the SM+CHOL bilayer. As in the case of SOPS+CHOL bilayer, the distribution of number of lipids hydrogen-bonded to cholesterol is dominated by the corresponding distribution of SM molecules hydrogen-bonded to cholesterol, while the probability for the cholesterol-cholesterol hydrogen bonding is insignificant (see Fig. 2.9 (d)-(f)). While the distribution of hydrogen bonds for the SOPS molecule was peaked at two hydrogen bonds, the peak of the distribution for the SM molecule is at three hydrogen bonds. Both SM-cholesterol and SM-SM hydrogen-bonding make a contribution into the total distribution (see Fig. 2.9 (a)-(c)). Also, for the SM+CHOL bilayer, a cholesterol molecule has a higher probability of participating in a hydrogen

bonding with two phospholipids, if compared to the case of the SOPS+CHOL bilayer (see Fig. 2.8 (f) and 2.9 (f)).

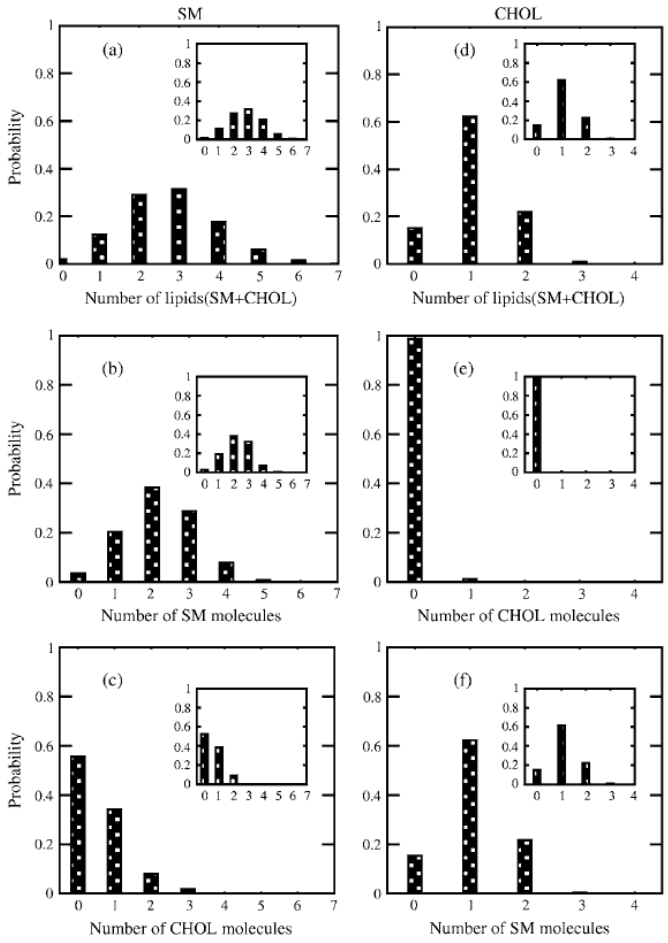

Fig 2.9 Distribution of molecules that are hydrogen-bonded to a give lipid in SM+CHOL bilayer. Distribution of (a) total number of lipids, (b) total number of SM molecules, and (c) total number of cholesterol molecules that are hydrogen-bonded to a given SM molecule. Panels *d-f* show the distribution of total number of lipids, total number of cholesterol molecules, total number of cholesterol molecules, and total number of SM molecules, respectively, that are hydrogen-bonded to a given cholesterol molecule. Inset in each plot shows the corresponding distribution of

molecules that are hydrogen-bonded to a given lipid in the SM+CHOL leaflet of the asymmetric bilayer.

Insets in Fig 2.8 and 2.9 show the distributions of molecules hydrogen-boned to a given lipid in the two leaflets of the asymmetric bilayer. While the distribution of hydrogen bonds per SOPS molecule is peaked at two bonds for the symmetric bilayer, it is peaked at three in the asymmetric bilayer. As the comparison of plots in Fig. 2.8 shows, the difference comes from a subtle change in the bonding character of SOPS with other lipids in the bilayer. Properties of the distributions for the number of lipid molecules hydrogen-bonded to a given cholesterol molecule in the SOPS+CHOL leaflet remain the same as in the case of the symmetric SOPS+CHOL bilayer. The shift of the most probable number of hydrogen bonds fro the SOPS+CHOL leaflet is consistent with the increase of order observed from calculations on area and order parameters when going from the symmetric to asymmetric bilayer. The properties of the distribution of hydrogen-bonded molecules in case of SM+CHOL leaflet of the asymmetric bilayer are similar to those in the symmetric SM+CHOL bilayer (see Fig 2.9). This is also consistent with our previous observations on order parameters (Fig. 2.7).

From the histogram shown in Fig. 2.8 and 2.9 we conclude that individual cholesterol molecules are more prone to engage in hydrogen-bonding with SM molecules compared to PS molecules. To understand the collective properties of the interlipid hydrogen-bond network, we perform an analysis of hydrogen-bonded

patches or clusters formed in each bilayer leaflet. A patch is defined as a group of lipids that share at least one hydrogen bond among them. The patch size is reported in terms of the number of lipids that constitute a patch. Fig. 2.10 shows the results for the upper leaflet of the two bilayers, the results for the lower leaflets are the same, since the bilayer are symmetric. For the SM+CHOL bilayer, the direct bonding between charge groups like $N^{+}(CH_3)_3$ and PO_4- is also taken into account. Fig 2.10 (a) and (b) show the distributions of maximum number of lipids contained in a patch for the SOPS+CHOL and SM+CHOL bilayers, respectively. These distributions were obtained from the trajectories and therefore they characterize the dynamics of the hydrogen bonding. As we see from Fig. 2.10 (a) the most probable patch contains maximum of ~30 lipids in case of the SOPS+CHOL bilayer. For the SM+CHOL bilayer the most probable patch contains ~60 lipids maximum, as Fig. 2.10 (b) shows.

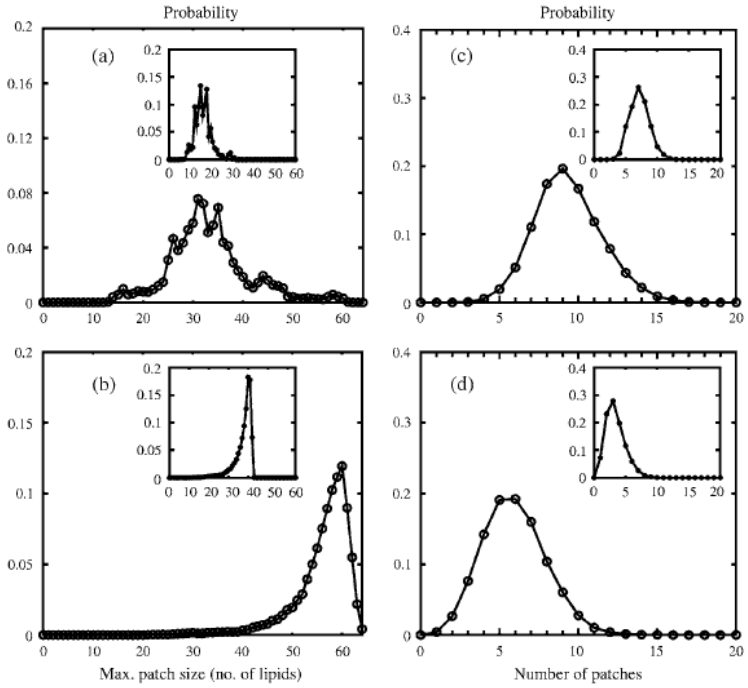

Fig 2.10 The distribution of maximum patch size in a leaflet (a) SOPS+CHOL and (b)

SM+CHOL bilayer. Patch size is measured in terms of the number of lipids that constitute the patch. The distribution of the number of patches in a leaflet of (c) SOPS+CHOL and (d) SM+CHOL bilayer. Calculation takes into account all types of interlipid bonding. Inset of each plot shows the corresponding distribution calculated using interlipid bonding among phospholipids only. In each plot the line is drawn as a guide to the eye only.

This is close to the situation when the whole leaflet in our simulation box is connected through the network of hydrogen bonds. Distributions for the number of distinct patches in a leaflet, shown in Fig. 2.10 (c) and (d), indicate the level of organization or fragmentation of the bilayer surface. It can be seen from Fig. 2.10 that the SOPS+CHOL bilayer surface is more fragmented than the SM+CHOL bilayer surface. These properties of network of interlipid bonding or patches were calculated for all types of interlipid bonding. It is important to know the contribution of the phospholipid (PL) + CHOL hydrogen bonding to the overall patch size. This can be investigated by identifying the interlipid bonding network only due to PL molecules. The corresponding results are shown as an inset in Fig.2.10. As can be seen from the inset to Fig. 2.10 (a) and (b), the influence of cholesterol on the patch size is different fro the SOPS+CHOL and SM+CHOL bilayers. For the SM+CHOL bilayer, the most probable patch contains maximum of ~40 SM molecules. This number is close to the total number of SM molecules in the leaflet. For the SOPS+CHOL bilayer, the distribution shows that most probable patches contain between 15 to 20 phospholipids.

The number of distinct SM patches is also smaller than the average number of distinct SOPS patches (see inset to Fig. 2.10 (c) and (d)). As we saw previously any given cholesterol molecule is most likely to be hydrogen-bonded to a single phospholipid molecule rather than to two or more phospholipid molecules (see Figs 2.8 (f) and 2.9 (f)). The consequences of this are manifested on a nanoscale through the results on patch sizes shown in Fig. 2.10. Based on data from Figs 2.8-2.10, we conclude that a cholesterol molecule does not act as a bridge between two patches to form a large patch that can cover nearly the whole leaflet. Rather cholesterol increases the size of the patch mostly by just connecting through hydrogen bonds to the patch that exists because of hydrogen-bonding between phospholipids. Fig. 2.11 (a) and (b) show a schematic of patches of different sizes that exist in leaflets of SOPS+CHOL and SM+CHOL bilayers, respectively. In Fig. 2.11 (a) and (b), positions of nitrogen atoms from the NH_3+ group of SOPS and the $N^+(CH_3)_3$ group of SM and the position of oxygen atoms from CHOL are plotted in the *xy* plane as representative of respective lipids. Fig 2.12 shows the distributions of hydrogen-bond network patches for the asymmetric bilayer when the PL-CHOL hydrogen-bonding is taken into account. Insets to Fig. 2.12 show the distributions calculated by considering only the phospholipid hydrogen-bond network. As can be seen from Fig. 2.12, the average properties of the patch sizes did not undergo a significant change by going from the symmetric to the asymmetric bilayer. Thus we conclude that the interleaflet coupling in the asymmetric bilayer did not induce any largescale changes in the surface organization of the two leaflets as compared to the symmetric bilayers.

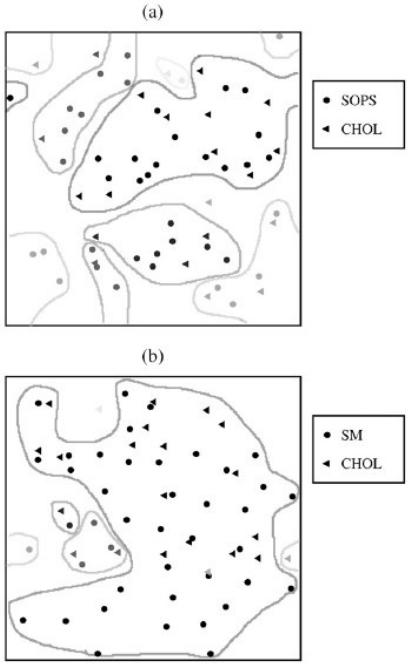

Fig 2.11 A schematic view of the different patch sizes in a leaflet of (a) SOPS1CHOL bilayer and (b) SM1CHOL bilayer. The solid circles represent positions in the x,y plane of nitrogen atoms from NH13 of SOPS and $N^{+}(CH3)3$ of SM, while the triangles represent the positions in the x,y plane of the oxygen atoms of CHOL. The line around the patches is drawn only as a guide to the eye.

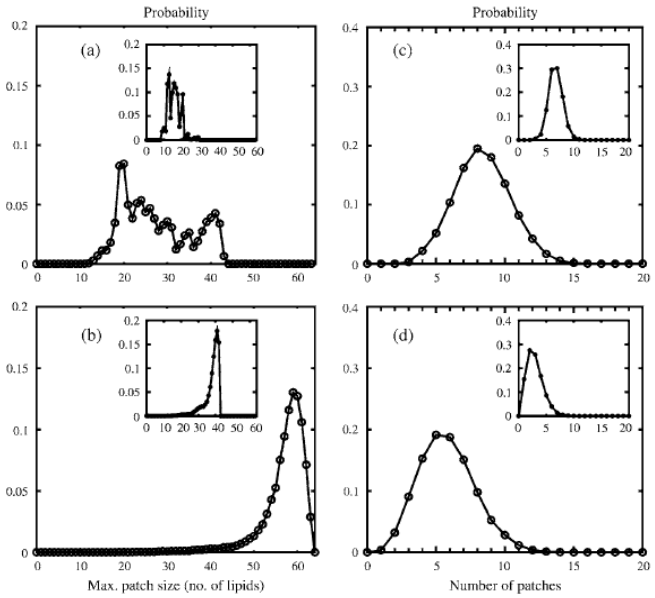

Fig 2.12 The distribution of maximum patch size in (a) SOPS1CHOL leaflet and (b) SM1CHOL leaflet of the asymmetric bilayer. Patch size is measured in terms of the number of lipids that constitute the patch. The distribution of the number of patches in (c) SOPS+CHOL leaflet and (d) SM+CHOL leaflet of the asymmetric bilayer. Calculation takes into account all types of interlipid bonding. Inset of each plot shows the corresponding distribution calculated using interlipid bonding among phospholipids only. In each plot the line is drawn as a guide to the eye only.

2.4 Conclusion

The histograms from Figs. 2.8 and 2.9 demonstrate that in the symmetric SM+CHOL and SOPS+CHOL bilayers the individual lipid molecules have nearly the same number of hydrogen-bonded nearest neighbors, although some slight differences between Figs. 2.8 and 2.9 exist. These differences practically disappear from the histograms shown in insets to Figs. 2.8 and 2.9, meaning that the hydrogen bonding per lipid is the same in the asymmetric bilayer. Nevertheless, as Figs. 2.10–2.12 show, the hydrogen-bonding network is different for the two symmetric bilayers with the SM+CHOL bilayer having more robust and SOPS+CHOL bilayer having more fractured characters. If one correlates the robustness of hydrogen-bonded network with the tendency to create l_o domains, one concludes that in the SM+CHOL mixtures the probability to observe rafts on a larger spatial scale is higher. From the simulations performed on the asymmetric bilayers we conclude that properties of the monolayers in the leaflets (with compositions specific for our simulations), do not change much

when going from a symmetric to an asymmetric bilayer, indicating that cross-leaflet interactions such as interdigitations are unimportant. The outer leaflet in an asymmetric bilayer (containing SM+CHOL) is more prone to raft creation compared to the inner leaflet. The conclusions from our simulations are consistent with the ideas expressed in the work of Devaux and Morris[58], who proposed that the sizes of the l_o domains in membrane leaflets should not be the same and that, possibly, proteins may play an important role in creation and functioning of rafts. Finally, we would like to mention that although our simulations were performed on relatively long timescales for simulations (tens of nanoseconds), the large-scale rearrangements of lipids are not possible on this timescale. Nevertheless, if the cross-leaflet interaction would be important, we would have observed its effects on the timescale of our simulation.

CHAPTER 3: ON THE INTERACTIONS OF CHOLESTEROL WITH DIFFERENT LIPIDS

3.1 Introduction

Biological membranes are very complicated entities containing self assembled mixtures of different lipids and proteins. Around 500 different lipid species are identified in biomembranes[44], but one of these species, cholesterol, plays a unique role. The importance of cholesterol is related to its influence on the physical properties of lipid bilayers. For example, cholesterol regulates elastic properties of plasma membranes and at high cholesterol concentrations the mechanical strength of this membrane is increased.[102] Cholesterol is also instrumental for the creation of "membrane rafts" or domains in membrane. The concept of a membrane or lipid raft was suggested in 1997[32] and since then the study of these rafts has undergone an explosion. It was proposed that lipid rafts participate in a multitude of cellular processes such as signal transduction [33,103,104], protein and lipid sorting [105], cellular entry by toxins and viruses and viral budding [106,107] (to name just a few). Many issues related to the nature and organization of rafts in cellular membranes are far from being clarified[61], although recently, a definition of what is a raft was established. According to the Keystone Symposium on Lipid Rafts and Cell Function in 2006, "Membrane rafts are small (10-200 nm) heterogeneous, highly dynamic, sterol– and sphingolipid-enriched domains that compartmentalize cellular processes". [39]

Since biological membranes are complicated mixtures of phospholipids, sterols and proteins, it is very difficult to analyze them and understand the role every component plays in the functioning of the assembly. As a result, many investigations are performed on model membranes containing either pure components or well-controlled mixtures of two/three components.[108] It was observed that bilayers containing ternary mixtures of cholesterol and two other phospholipids, such as a bilayer containing cholesterol, a saturated lipid and an unsaturated lipid ternary mixture of cholesterol/sphingomyelin/palmitoyloleoylphosphatidylcholine (Chol/SM/POPC) is laterally inhomogeneous and it contains domains enriched in cholesterol and SM.[53] Thus it was concluded that artificial membranes also have raft-like domains, although these domains have a different length scale. The phase diagram for this mixture was obtained and it showed regions of liquid-liquid phase coexistence. (A recent review of Veatch and Keller[56] presents an excellent introduction into the subject of phase diagrams of two and three component mixtures in general and mixtures with cholesterol specifically). Two liquid phases that appear in the bilayer correspond to the liquid ordered phase (l_o) and liquid disordered phase (l_d). In the l_o phase, the conformational order of lipid chains displays a high degree of order, while the translational order parameter shows a high degree of disorder. In the l_d phase, both conformational and translational order parameters display high degree of disorder. The l_o phase constitutes the domains that are enriched by a mixture of cholesterol and SM, while the l_d phase is enriched by POPC.

Although for natural membranes the recent definition of rafts requires presence of

the sphingomyelin in the system, it was observed that for the synthetic bilayers containing a ternary mixture of cholesterol, a saturated phosphatidylcholine with a high melting temperature such as dipalmitoylphosphatidylcholine (DPPC), and an unsaturated phosphatidylcholine with a low melting temperature, such as dioleoyl phosphatidylcholine (DOPC), i.e. for a mixture of Chol/DPPC/DOPC, the phase diagram is similar to the diagram observed for the Chol/SM/POPC ternary mixture.[56] This indicates that the phase separation and following from it lateral domain creation is quite general in model systems.

A number of different physical techniques have been used to study lateral organization of lipid domains in natural and model membranes.[53,109-117] Although the details about the structural properties of the domains in biological cells are not yet known, and it is also not clear if the domains in cell membranes (rafts) are related to the domains observed in model membranes, it has been observed that in both natural and artificial membranes domains are rich in cholesterol and saturated lipids. Therefore, it is important to understand the nature of interactions between cholesterol and other lipids.

To explain the phase diagram of the type seen for the Chol/SM/POPC or Chol/DPPC/DOPC mixtures, Radhakrishnan and McConnell proposed that the phospholipid with a higher melting temperature forms a complex[51] with cholesterol by a reversible reaction $C_p + R_q \Leftrightarrow X$ where C is cholesterol, R is a phospholipid with a high melting temperature, T_m , and $X=C_pR_q$ is the complex. Complexes can associate and exist in the oligomeric form and the l_o phase contains these oligomers.

The low T_m phospholipid is only partially miscible with the complex and therefore is considered to be "unreactive" (U) with cholesterol. The existence of complexes is hard to prove experimentally, although the interpretation of some recent nearest-neighbor recognition measurements suggested their possible presence in the bilayers.[118] No complexes were observed in the earlier molecular dynamics simulations on bilayers containing mixtures of cholesterol and phospholipids, until Chiu et al. suggested that hydrogen bonding between cholesterol and DPPC molecule is responsible for the 1:1 complex between these two lipids.[70] The idea that hydrogen bonding between phospholipids and cholesterol is responsible for complexation of cholesterol with phospholipids was advanced even further by Pandit et al.[119] who performed simulations on bilayers containing binary cholesterol and DPPC and also on bilayers containing mixtures of cholesterol with a phospholipid of a smaller T_m – DLPC. Since complexes of cholesterol and phospholipids can contain p cholesterol and q phospholipid molecules, where p and q are often assumed to be 1 and 2, it was suggested[119] that cholesterol can hydrogen bond to DPPC lipid not just through its headgroup hydrogen, but also through its oxygen, thus hydrogen bonding to the methyl group of tetramethylammonium of the DPPC. The most probable number of hydrogen bonds per cholesterol in simulations of Pandit et al. turned out to be two when the bilayer contained a mixture of cholesterol and DPPC. When the phospholipid in the bilayer was DLPC, the most probable number of hydrogen bonds per cholesterol was just one. Pandit et al. speculated that with two hydrogen bonds per cholesterol, one can create a linear network of hydrogen bonded molecules

Ph(phospholipid)-chol-Ph-chol … when Ph is a DPPC molecule. Such a network has a smaller chance to exist in the case when Ph is a DLPC molecule. Therefore, if one assumes that presence of hydrogen bonded network is correlated with the existence of oligomeric complexes, one should find more such complexes in the bilayers containing mixtures of cholesterol and DPPC compared to bilayers containing mixtures of cholesterol and DLPC. Cholesterol is most probably engaged in two hydrogen bonds with DPPC and only one bond with DLPC due to the tilt of cholesterol axis in DLPC.[119] This tilt appears so that cholesterol can be properly accommodated in the hydrophobic region of the DLPC membrane. Due to this tilt, cholesterol is not in a favorable position to engage in two hydrogen bonds with neighboring DLPC phospholipids. The phase diagrams obtained for cholesterol/DPPC and cholesterol/DLPC monolayers by McConnell and coworkers and their interpretation of these diagrams indicated that if complexes exist, they exist for the DPPC/cholesterol system and do not exist for the DLPC/cholesterol case[52], which is consistent with the interpretation of Pandit et al. obtained from the simulations performed on bilayers[119]. Although there is a qualitative agreement between the interpretation of simulations and experiment in this case, still one has to be careful comparing the results from simulations on bilayers with the experimental results obtained for monolayers, since the structural properties of monolayers and bilayers can be quite different.

Recent simulations performed on ternary mixtures containing cholesterol, POPC and SM[120,121] confirmed the previous suggestion that cholesterol/SM interaction is

different from the cholesterol/POPC interaction due to the difference in the structures of phospholipids.[122] Still, it is not clear how this difference is responsible for the shape of phase diagrams observed for the ternary mixture of these lipids. Also, simulations did not confirm the existence of the complexes between cholesterol and SM and did not clarify the role of different regions of lipids in the phospholipid/cholesterol interaction. Especially interesting will be to understand what part of the phospholipid molecule contributes mostly to this interaction and, if complexes exist, to their creation. The role of different lipid regions and their corresponding contributions into the cholesterol/phospholipid interaction was investigated by experimentalists for quite a while.[123,124]

Simulations can help in providing at least some partial answers to the above posed questions; therefore we performed a series of simulations on model lipid membranes. Three of the simulations were performed on bilayers containing binary mixtures of cholesterol with one phospholipid component: either SSM (18: 0 SM), or OSM (18:1 SM) or POPC. Three other simulations were done on bilayers containing just a pure phospholipid: SSM or OSM or POPC. By comparison between the results from our simulations, we can understand how the addition of cholesterol changes phospholipid properties in the membranes and also how a change in the structure of the headgroup or tail of the phospholipid molecule influences properties of the bilayer with cholesterol.

3.2 Computational Details

Since we want to understand how the details of the molecular structure influence cholesterol/phospholipid interactions, we performed our simulations using an atomic description of the molecules. This restricted us to simulations of bilayers that are rather limited in their size. Also, since we want to understand the relative importance of different structural regions of phospholipids in the cholesterol/phospholipid interaction, we performed simulations on different binary bilayers containing cholesterol and different phospholipids. Every bilayer in our simulation contained 128 lipid molecules, 64 in each leaflet. In bilayers containing cholesterol, there were 22 cholesterol molecules in each leaflet, resulting in a cholesterol concentration of about 34 mol %. Every lipid molecule was hydrated by 30 water molecules. The molecular structures of SSM, OSM, POPC and Chol were generated by SYBYL version 7.0 (Tripos Inc., St. Louis, MO). The initial structure of a bilayer leaflet was obtained by generating an 8x8 array of 64 lipids in the *x-y* plane by random rotation of each lipid molecule around the *z*-direction. For bilayers containing Chol, 22 of the phospholipids were randomly selected and replaced by Chol molecules. The other leaflet was obtained by reflection and translation of the first one. A water slab with 3840 water molecules was added to both sides to solvate the headgroups. The normal of the bilayer was directed along the *z*- axis.

All simulations were performed using the GROMACS package [5,6]. The LINCS algorithm was used to constrain all bonds in the systems [83] allowing an integration timestep of 2 fs. Periodic boundary conditions were applied in all three directions and long range electrostatics interaction was handled using the SPME algorithm [10] with a

real space cut off of 10 Å, 4th order interpolation and a tolerance of 10^{-5}. A 10 Å cut-off was used for all van-der Waals interactions. All simulations were done in the NPT ensemble. The temperatures in the simulations were maintained at 323 K using the Nose-Hoover scheme[84] with a thermostat oscillatory relaxation period of 0.5 ps and the Parrinello-Rahman semi-isotropic pressure coupling scheme [85] with a barostat time constant of 2.0 ps was used to keep the pressures at 1 atm. The SPC/E model of water [86] was employed in the simulations. Force field for the SM was the same as used by Niemelä et al. [90], who demonstrated that it is reasonable in reproducing experimental trends. The force field parameters of Chol were the same as used in the study by Pandit et al.[119] The force field for the POPC is obtained from Tieleman et al [125].

All the simulations except the POPC/Chol bilayer were 100 ns long, and the simulation of POPC/Chol was 95 ns long, since the area per lipid in this case behaved in a very stable way. The positions and velocities of the atoms were saved every 1 ps. The first 20 ns run was considered to be the equilibration time for all bilayers, and the last 80 ns (75 for POPC/Chol) trajectory was used for their analyses. The analyses were performed using the utilities available in GROMACS as well as programs written by us.

3.3 Results and Discussion.

3.3.1 Structural properties

We will start with some description of the structural properties of the bilayers that

we simulated. Fig. 3.1 shows the area per lipid in the six simulated bilayers as a function of time.The average values are 48.1±0.7Å^2, 53.4±0.7Å^2, and 65.0±1.2Å^2 for pure SSM, OSM, and POPC. For mixed bilayers the areas obtained from our simulations are 39.4±0.4 Å^2, 41.2±0.7 Å^2, and 42.9±0.5Å^2 for SSM/Chol, OSM/Chol and POPC/Chol bilayers, respectively. The average area values for the pure SM bilayers are somewhat smaller compared to what Niemelä et al. [90] calculated from their study, although we used the same force field for the SM lipids. This could be due to our use of the SPC/E water model instead of the SPC model that was used in their simulations. Our calculated areas per one component lipid bilayer can be compared to the values for the areas that were extracted from experiments. As it turned out, there are not that many data on the areas per SSM, OSM and POPC in the literature. The areas determined for SM lipids from the X-ray diffraction measurements were 47 Å^2 for the (16:0)SM and 55 Å^2 for the SSM lipid at 328K.[126] Other measurements, although not done on bilayers, but by using Langmuir film balance at a surface pressure of 30 mN/m and at T=303 K showed a more consistent trend: the area per (16:0)SM was 52.5 Å^2, while it decreased to 48.6 Å^2 for SSM.[48] For the OSM, the area reported by the same group was 61.5 Å^2.[48] The experimental value of the area per headgroup in POPC bilayer obtained from the X-ray measurements was reported to be 68.3 ± 1.5 A^2 at 303K.[127] To determine the area per headgroup in the bilayer from the simulation, that can be compared with the experimentally extracted, requires very careful consideration.[128,129] For mixed bilayers the areas obtained from our simulations are 39.4±0.4 Å^2, 41.2±0.7 Å^2, and 42.9±0.5Å^2 for SSM/Chol, OSM/Chol

and POPC/Chol bilayers, respectively.

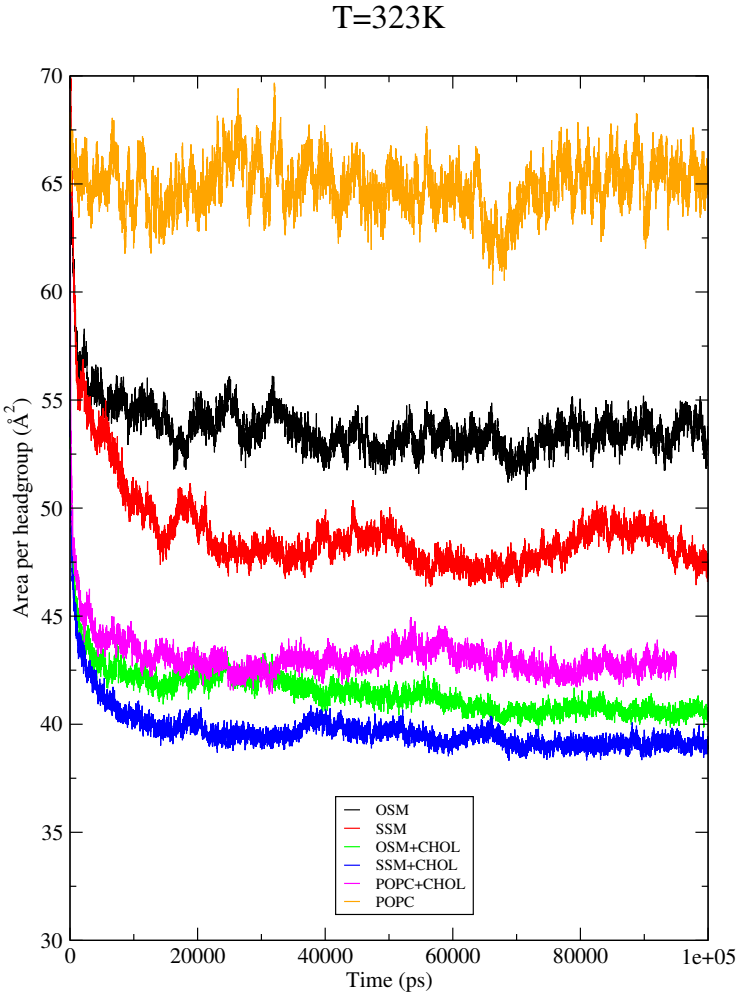

Fig 3.1 The area per molecule in the simulations: pure OSM (black), pure SSM (red), pure POPC (orange), OSM/Chol mixture (green), SSM/Chol mixture (blue), POPC/Chol (magenta).

Values for the area we calculated and that are given above called geometrical area, are obtained by using a simple geometrical construction when the area of the simulation box is divided by the number of lipids in the bilayer. Although simple, this approach allows us to perform comparison of areas per lipid in different bilayers.

Thus, we observed, that as the acyl chain in SM molecule became unsaturated, the area per lipid increased. We also observed that the area per SM molecule in the pure SM bilayer was substantially smaller than the area per POPC molecule in the pure POPC bilayer due to the intermolecular hydrogen bonds in the SM bilayer. The corresponding average areas per molecule in the bilayers with binary mixtures were reduced, with the largest reduction observed for the POPC/Chol bilayer. To see the more specific effect cholesterol molecules produce on the area per lipid molecule in the mixture bilayers, we used the method developed by Hofsäß et al [71] to evaluate the individual areas for SM and Chol molecules. According to Hofsäß et al, the area per phospholipid and Chol molecule can be calculated by the following equations:

$$a_{PL} = \frac{2A}{(1-x)N_{lipid}}\left[1 - \frac{xN_{lipid}V_{chol}}{V - N_W V_W}\right] \tag{3.1}$$

$$a_{chol} = \frac{2AV_{chol}}{V - N_W N_W} \tag{3.2}$$

where A is the cross section (x-y) area of the simulation box, N_{lipid} is the total number of molecules in the bilayer, x is the mole fraction of cholesterol: $x=N_{chol}/N_{lipid}$, V is the volume of the simulation box, N_W is the number of water molecules in the system. The volume per Chol molecule, V_{chol}, is taken to be 0.593 nm^3 and that of water, V_W, is 0.0305 nm^3. Following the above equations, the area per SSM molecule was found to be 47.6 $Å^2$ and that of Chol 23.7 $Å^2$ in the SSM/Chol bilayer; the area per OSM molecule was calculated to be 49.7 $Å^2$ and that of Chol 24.9 $Å^2$ in the OSM/Chol bilayer. The area per POPC molecule is 49.8 $Å^2$ and that of Chol is 25.2 $Å^2$ in the POPC/Chol bilayer when equations (1) and (2) are used. No significant shrinkage in

the area per SSM is observed when going from pure SSM bilayer to a bilayer with 34.4 mol% Chol. The area per OSM shrinks by 2-3 $Å^2$ with the addition of cholesterol and the area per POPC shrinks by a significant amount of 15 $Å^2$ when cholesterol is added. The reason for the observed small changes in the SM area is that SM molecules in the pure SM bilayer are already well packed, hence there is not much space left for condensation when cholesterol is added. Again, one should mention here that obtaining the area per molecule in bilayers with lipid mixtures is not a simple issue and the values derived could be dependent on the method one uses. Nevertheless, the values presented here show a clear trend in area changes with addition of cholesterol.

Fig. 3.2 shows the electron density plots from our simulations. The peak to peak distances (ppd) for the whole system estimated from the plots are 48.0 Å and 48.6 Å; 44.3 Å and 46.8 Å; 36.0 Å and 45.5 Å for pure SSM bilayer and SSM/Chol mixture; pure OSM bilayer and OSM/Chol mixture; pure POPC and POPC/Chol mixture, respectively. Perhaps more instructive will be to consider just the change in the ppd for the phospholipids, also shown on Figure 2, with the addition of cholesterol. For the SSM the ppd change is only 0.5 Å (from 45.0 Å to 45.5 Å); for the OSM it is 3 Å (from 40 Å to 43 Å); finally for the POPC it is 6 Å (from 34 Å to 40 Å). While there is a slight increase in the ppd in the case of the SSM bilayer and a moderate increase in the thickness of the OSM bilayer, the increase in the ppd of the POPC bilayer is quite large upon addition of the cholesterol. These results are consistent with the area changes in the corresponding bilayers: the larger increase in the ppd corresponds to

the larger reduction of the area per lipid molecule. Experimental values of the peak to peak distances for different SM bilayers and SM/Chol bilayer are within the range of 4.1 nm-5.6 nm [102,126,130,131]. An exact comparison between the simulation results and the experimental values are difficult due to the various experimental conditions and we consider our results to be within a reasonable range.

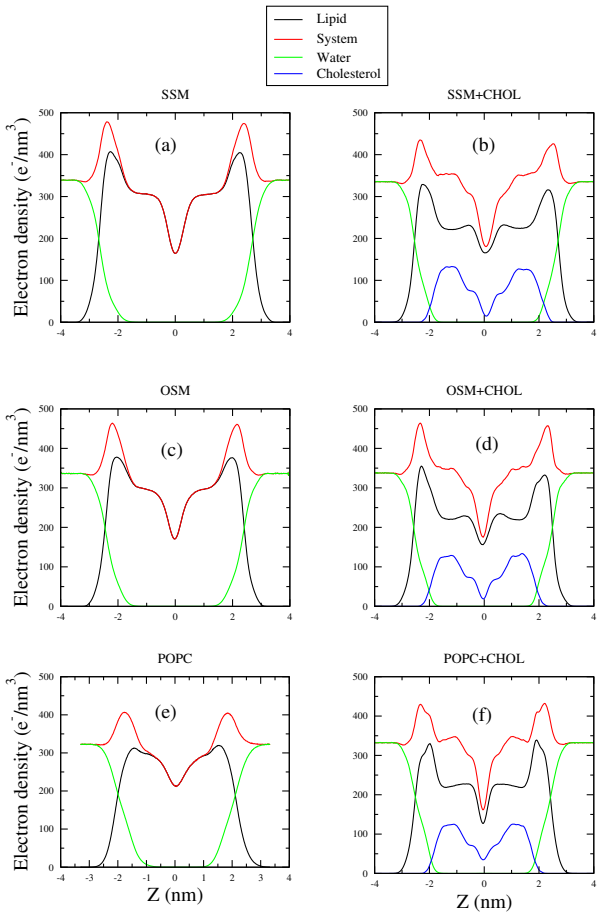

Fig 3.2 The electron density plot of pure SSM bilayer (top left), SSM/Chol bilayer (top right), pure OSM bilayer (middle left), and OSM/Chol bilayer (middle right), pure POPC bilayer (bottom left) and POPC/Chol bilayer (bottom right).

Fig 3.3 shows profiles of the deuterium order parameter for the chains of the

phospholipid molecules in our simulations. The lipid tails in pure SSM bilayer are more ordered than those in pure OSM bilayer because the cis- double bond on the acyl chain in OSM disrupts the packing of the tails. Upon the addition of Chol to the SM bilayers, the order parameters increase in both SSM and OSM bilayers. The increment is larger in the case of OSM bilayer, although the absolute value of the order parameters is higher in SSM/Chol bilayer. Overall, the data show that SSM/Chol bilayer is tightly packed. A dramatic increase in the POPC carbon tail order parameters was observed when Chol was added to a POPC bilayer, indicating that condensing effect is very strongly pronounced in this case; the implication of this will be discussed

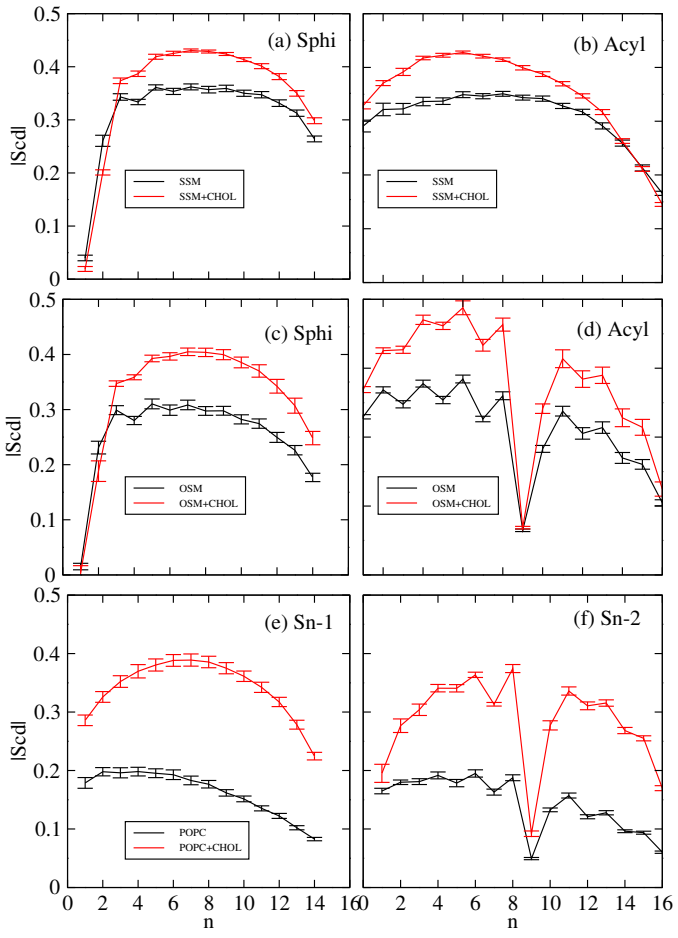

Fig 3.3 The deuterium order parameter profiles for the carbon tails of SM molecules in pure SM and POPC bilayers (black) and in SM/Chol or POPC/Chol bilayers (red).

Another revealing structural property of cholesterol in bilayers containing a mixture of cholesterol and phospholipids is the cholesterol axis tilt angle distribution with respect to the bilayer normal. This tilt angle is believed to be related to the ability of cholesterol to support a hydrogen bonding network in the bilayer and its ordering[119,121]. Fig. 3.4 shows distributions for cholesterol tilt in the SSM/Chol, OSM/Chol, and POPC/Chol bilayers. The curves are quite similar for both SM/Chol bilayers, while the distribution is broader for the POPC/Chol bilayer. Also, the average tilt angles of cholesterol in bilayers with SM are smaller than in POPC case, which is consistent with the smaller hydrophobic length of the POPC/Chol bilayer, as demonstrated in the ppd data.

3.3.2 Lipid tail distributions around cholesterol

Our results indicate that SSM/Chol and OSM/Chol bilayers display somewhat similar structural properties. Nevertheless, it is known that the structure of SM molecule does influence the properties of bilayers containing ternary mixtures such as SM/DOPC/Chol[132 133]. Therefore, we would like to understand how the replacement of the saturated acyl tail by an unsaturated one influences the properties of the SM interaction with cholesterol.

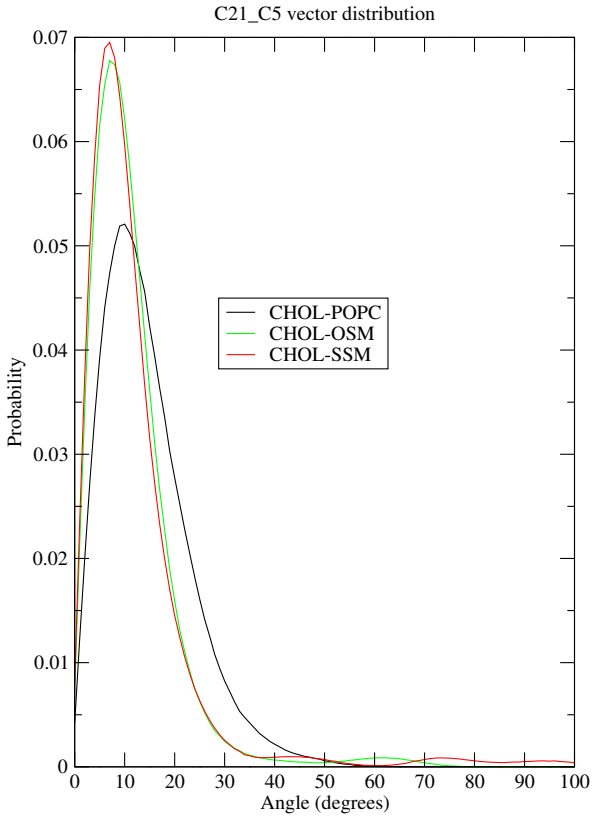

Fig 3.4 Cholesterol principal axis tilts angle distribution with respect to bilayer normal. The cholesterol principal axis is defined as the vector from carbon C21 (the carbon atom in the ring that connects to the short tail of a cholesterol molecule) to C5 (the carbon atom to which the hydroxyl group binds).

The cholesterol molecule has a smooth α-face and a rough β-face with two protruding methyl groups. This kind of a special molecular structure is believed to be important in shaping the interaction of cholesterol with different lipid molecules. In their simulation of a dioleoylphosphatidylcholine (DOPC), SSM, and cholesterol ternary mixture, Pandit et al [78] observed that the smooth α-face of cholesterol preferentially packs next to the SSM molecules. And Suits et al. from their simulation[134] observed that cholesterol prefers to be solvated by the saturated lipid tail and has a low affinity

for the polyunsaturated fatty acid. Here, we also study the organization of SM lipid tails around cholesterol molecule to see if there are any obvious differences in the SSM/Chol and OSM/Chol bilayers.

In fig. 3.5 and 3.6, we show the sphingosine chain and the acyl chain carbon atom density distribution isosurface as a function of the distance and orientation around cholesterol molecule in the bilayers. The isosurfaces were drawn using gOpenMol software by Laaksonen et al.[135,136]. The figures illustrate that the distributions for the sphingosine carbon atoms and the acyl chain atoms are not homogeneous in the case of OSM/Chol bilayer, especially when facing the smooth side of cholesterol (the down side in the plot). Next to the smooth face, one is more likely to find carbon atoms from the saturated sphingosine chain (regions shown in blue). Though carbon atoms from acyl (unsaturated) chain (regions in red) can also be found on this side, they are located more towards the

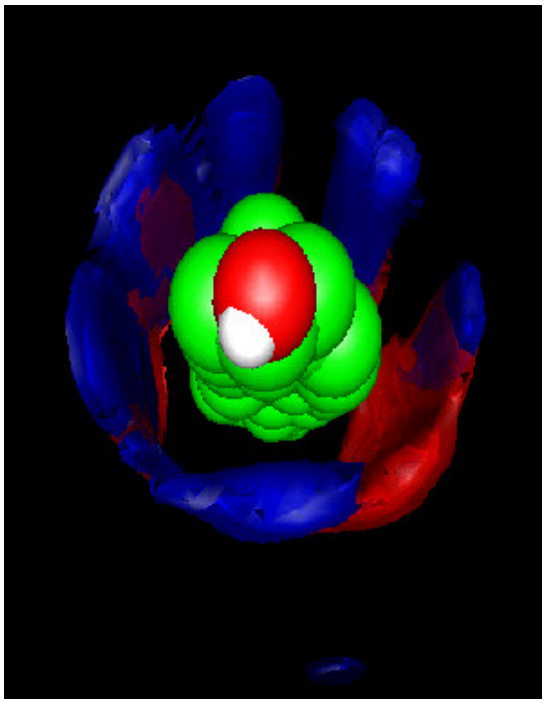

Fig 3.5 Probability density isosurface for sphingosine chain (blue) and acyl chain (red) tails of the OSM lipid around Chol molecule. There are several layers of lipid tails solvating Chol molecule and here a cut-off for the probability is applied to show only the first layer.

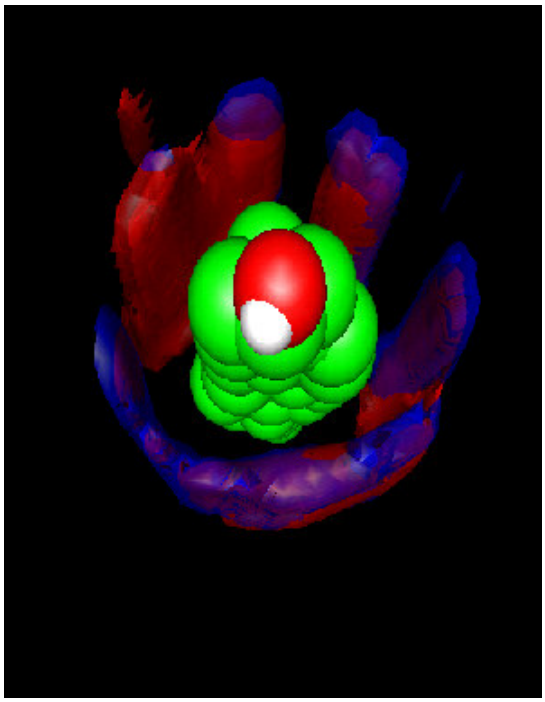

Fig 3.6 Probability density isosurface for sphingosine chain (blue) and acyl chain (red) tails of the SSM lipid around Chol molecul

edge of the cholesterol ring. While the probability to find carbon atoms next to the rough face of cholesterol molecule is small due to strong steric interaction with the methyl groups, it is more or less the same for the carbon atoms from both chains on each side of the two methyl groups. (Note that the color of a region is purple when carbon atoms from both chains can be found with similar probability in this region).

The situation in the case of SSM/Chol bilayer is different. The tail distribution for neighboring SM molecules is more homogeneous in this situation, especially in the region facing the smooth face of the cholesterol, as can be seen from the prevailing purple color in the plot. This makes intuitive sense, since in the SSM molecule both tails are saturated and there should not be any preference of cholesterol over any one of these tails. Thus, we conclude that the smooth face of cholesterol has a preference for the saturated chains of the SM molecules, while the rough face does not have a real preference of one tail over the other.

3.3.3 Interaction energy between cholesterol and phospholipids

As we have shown, there is some difference in the structures of SSM/Chol and OSM/Chol bilayers. More profound is the difference between POPC/Chol and any of the SM/Chol bilayers. Does this mean that cholesterol can make complexes with SM and not with the POPC lipid molecule? If yes, how can we recognize these complexes? Can the hydrogen bond be a signature of a complex? Since a hydrogen bond is created due to the headgroup interactions, it can characterize only the headgroup interactions. One needs to consider the full phospholipid/cholesterol interaction energy and contributions into this energy from the tail and headgroup regions to reach a conclusion about the strength of the cholesterol phospholipid interaction and the existence of a possible complex.

If there is a complex between cholesterol and the phospholipid molecule, it probably exists between neighbors and it is probably of low energy. Therefore, we

calculated the distribution of lowest interaction energies (min(U(chol,phospholipid))) between cholesterol and its phospholipid neighbors in the bilayers containing cholesterol. The lipid neighbors for a cholesterol molecule were found according to the radial distribution function of the center of mass of phosphate groups (-PO_4) in lipids with respect to the hydroxyl oxygen atom in the cholesterol molecule. According to those RDFs (not shown), a cutoff distance of 0.8 nm between the oxygen atom in a cholesterol molecule and the center of mass of –PO_4 in a phospholipid (SM or POPC) was chosen to define the neighboring lipid molecules of a cholesterol. This cutoff is somewhat larger than the first minimum position in the RDF plot, but since we are going to calculate the distribution for the lowest energies (among all the interaction energies of a cholesterol with all its neighboring lipids), a larger cutoff distance could avoid the situation when the interaction energy is lowest but the distance between cholesterol oxygen and the PO_4 is beyond a smaller cutoff. The distribution for the lowest interaction energies between lipids and cholesterol are shown in Fig. 3.7. Also shown in this figure are separate contributions to the total energy from the headgroup and tail interactions. The hydroxyl group of a cholesterol molecule is taken to be its headgroup and the rest of the cholesterol atoms represent the tail. For the phospholipids, carbon chains are considered to be the tail part and the rest of the molecule is grouped into the headgroup part for the purpose of energy calculations. We observe from the Figure that the total energy curves are very similar for the SSM/Chol and OSM/Chol bilayers, while the curve for the POPC/Chol bilayer is shifted a little bit towards the lower energy range.

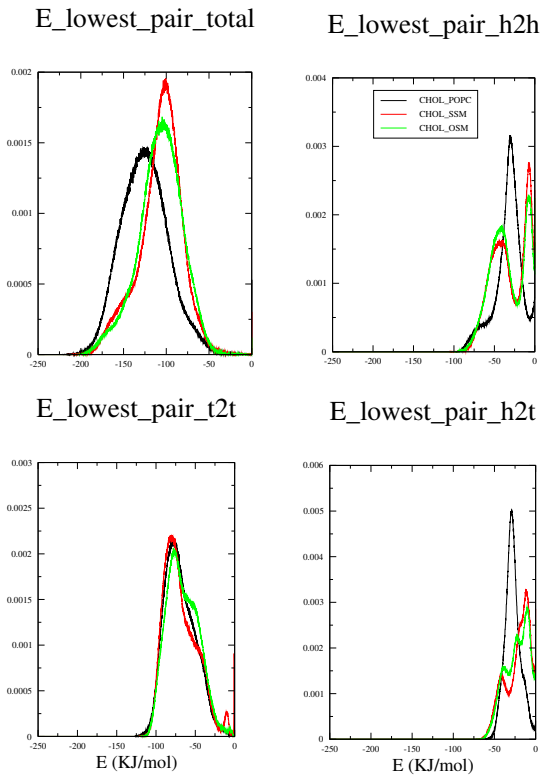

Fig 3.7 Lowest pair interaction energy distribution of Chol with its nearest lipid neighbors. (Note: h2h stands for head-to-head, t2t stands for tail-to-tail and h2t stands for head-to-tail)

Our figures show that all the interactions between cholesterol and SM (SSM or OSM) molecules are similar. The energy distributions of the headgroup-to-headgroup interaction in the SM/Chol bilayers are bimodal, while in the POPC/Chol bilayer mostly one peak is observed. One of the peaks in the SM/Chol case is situated at the lower energy than the peak in the POPC/Chol bilayer. This is a reflection of a different hydrogen bonding pattern present in our systems. SM molecules have more hydrogen bonding groups in the headgroup that could interact favorably with cholesterol. Considering that two atoms (the hydroxyl group) of cholesterol from the

headgroup of cholesterol can engage in hydrogen bonding, the interaction energies are typical for this bonding. The tail-tail interactions in all three bilayers are similar, although the bimodal character of this interaction is more pronounced in the SM/Chol bilayers, probably correlated with the different tilt of the cholesterol axis. The head to tail interaction is stronger in the POPC/Chol bilayer, and this is correlated to the larger tilt angle of cholesterol molecules in this bilayer. A larger tilt angle brings the phospholipid head and cholesterol tail group, which includes the ring system, closer to each other, thus lowering the van der Waals interaction between them. We observed that this lower head/tail interaction energy in the POPC/Chol bilayer is actually the reason for the shift towards the slightly lower total interaction energy between cholesterol and POPC. The energy distributions we obtained do not confirm the intuitive idea that the pair interaction between cholesterol and SM is stronger than the pair interaction between cholesterol and POPC. The distributions of total energy do not show a strong bimodal feature either, which can be used to identify 1:1 complexes as belonging to a population with a low interaction energy mode. Moreover from these distributions it is not possible to infer the existence of 1:1 complexes between cholesterol and SM and its absence between cholesterol and POPC. Given that the mol% of cholesterol is ~33, and that 1:2 complexes are possible, we calculated the distribution of the lowest energies of interaction of three molecules that are neighbors: one cholesterol molecule and two phospholipids. This distribution is shown in Figure 8.

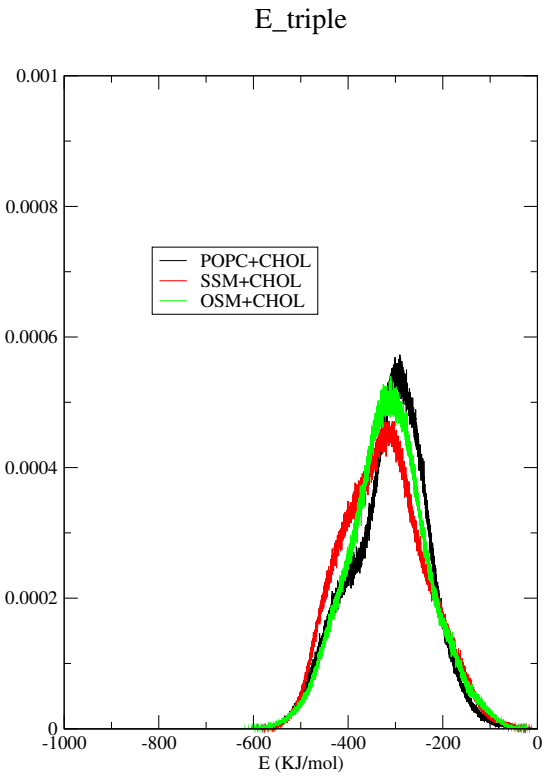

Fig 3.8 Lowest triple interaction energy distribution of CHOL with its nearest lipid neighbors. This energy also includes the interaction energy between the two lipids.

As we can see from this Figure, the distributions are very similar for all three binary bilayers. Again, we must conclude that energies of interaction we observed do not predict that 2 SM phospholipids and 1 cholesterol are engaged in a complex creation and 2 POPC molecules and 1 cholesterol are not. In addition, the triplet energy can not help us to determine what triplet of lipids is in a complex.

3.4 Conclusions

To understand the difference in the interaction between cholesterol and the two phospholipids: SSM and POPC, we performed simulations on binary bilayers

containing cholesterol and these phospholipids. Since the phospholipids are different in their tail and headgroup structures, we extended our simulations to include OSM, which represents a phospholipid with a structure somewhat transitional between POPC and SSM. The interactions of cholesterol with OSM and SSM were found to be similar in many aspects, although we observed that the smooth face of cholesterol shows some preference for the saturated tail of the OSM molecule. Also, we found, to some degree of surprise, that the distribution of the lowest interaction energy between cholesterol and POPC is similar to the one between cholesterol and SM. (We even found that the average lowest interaction energy between POPC and cholesterol is slightly lower than the one between SM and cholesterol). We studied the distribution of interaction energy of cholesterol and all nearest neighbors in all three binary bilayers and the results were similar to the one from Figure 7. The total interaction energies for the triplet of neighbors containing one cholesterol and two (same kind) lipid molecules were also very similar for the considered mixed bilayers. We conclude that our calculations of energy can not distinguish complexes from non-complexes. Also, we do not observe any energetic preference of cholesterol to interact with the SM compared to POPC. Given all this, the question that still remains to be answered is: What is the driving force for the phase separation in mixtures such as SM/POPC/Chol? We speculate here that while the interaction energy of cholesterol and POPC or cholesterol and SM is roughly the same and may be even that POPC may be slightly more favorably interacting with cholesterol compared to SM, the change in the tail ordering of POPC when cholesterol is inserted is rather dramatic

and this produces an unfavorable loss of entropy. As a result, POPC avoids mixing with cholesterol. In general, it is the delicate balance between the energy and entropy change when we mix the three components – SM, POPC and Chol that determines the phase diagram of the system. In a recent paper by Frazier et al.[137] where BSM/POPC/Chol bilayer was studied using Fluorescence Resonance Energy Transfer as an experimental technique and where MC simulations were also performed on a simple lattice, it was shown that one finds phase separations and domains in that system when all the three Flory-Huggins type parameters that describe the difference between the free energies of interactions between components in the bilayers have values of order of only 300 cal/mol. These small values indeed indicate the existence of a delicate balance between energy and entropy. Due to such balance it is possible that in ternary mixtures of sphingomyelin /phosphatidylcholine/Chol, domain structure may exist when sphingomyelin has both saturated chains like in SSM, while the domain structure may disappear when SSM is replaced by OSM. Indeed, Epand and Epand noted that this is the case when phosphatidylcholine is SOPC and attributed the change in mixing behavior to a greater miscibility of OSM with the phosphatidylcholine.[132] To understand what the driving force for the domain creation and separation in bilayers is, one needs to calculate from simulations the change in energy and entropy upon mixing. This task may be problematic, given the accuracy of the force fields we use today. Nevertheless, from simulations we can get a qualitative and semi-quantitative understanding of the role that different regions in cholesterol and phospholipid molecules play in their mutual interaction and subsequently the

ordering in the bilayers containing these lipids. Our simulations also suggest that phenomenological theories of lipid mixtures should consider in addition to compositional variables also conformational degrees of freedom reflected in the membrane width, to provide a description of free energies and subsequently phase diagrams. Such theories are already proposed.[138] Simulations can provide the values of parameters and test main assumptions made in these theories. Therefore, more computational work is needed to understand the nature of lipid rafts in both artificial and natural membranes.

CHAPTER 4: FREE ENERGY CALCULATION OF CHOLESTEROL TRANSFER BETWEEN LIPID BILYERS

4.1 Introduction.

Biological membranes are inhomogeneous mixtures containing a variety of molecular components such as phospholipids, sterols and proteins. One of the lipids in the mixture, cholesterol, plays a crucial role in promoting this inhomogenuity. It was proposed[32] that biological membranes contain domains, called "rafts", which display different from the rest of the membrane physico-chemical properties. Rafts in biological membranes contain an enhanced amount of cholesterol and sphingomyelin (SM). Although the proposed existence of the lipid rafts in biological membranes is not yet experimentally confirmed [53,61,139], the existence of such raft-like domains, possibly on a different, larger, length scale was observed in model membranes. Typical model membranes containing rafts consist of ternary mixtures of lipids, where one of the components is a phospholipid with a high melting temperature, another component is a phospholipid with a low melting temperature and finally, the third component is cholesterol[56]. In these model membranes, one can observe a liquid-liquid phase separation into the liquid ordered phase containing mostly a mixture of cholesterol and saturated phospholipid and a liquid disordered phase, containing mostly the unsaturated phospholipid[56]. A typical membrane of this kind

containing a ternary mixture of cholesterol, SM and palmytoyl-oleoyl-phosphatidylcholine (POPC) with cholesterol and SM mostly found in the raft domains. What is the reason behind the phase separation and the existence of domains in these model membranes? Is it mostly due to the difference in the interaction energy between cholesterol and different phospholipids? Indeed, phospholipids such as SM and POPC differ in their structure, and this difference should be reflected in their interaction with the cholesterol.[121] It was noticed some time ago, that, while cholesterol can engage in the hydrogen bonding between its hydroxyl headgroup and the SM's molecule headgroup as a donor and an acceptor, it can only serve as a donor when interacting with the PC's lipid headgroup.[122] Since it was also recently shown that cholesterol prefers to be in the neighborhood of a saturated chain of a phospholipid molecule[134], we expect that it will prefer the vicinity of the tails from the SM molecules. These considerations suggest that an energetic preference should exist for the cholesterol to choose an SM as its neighboring molecule and, therefore, should explain why domains with an enhanced amount of cholesterol and SM are observed. Moreover, very recently experimental work using high resolution calorimetry verified that indeed cholesterol prefers to be in the environment of SM[140]. It was determined that the free energy of transfer of cholesterol from the bilayer containing POPC and cholesterol at 30 mol% to the bilayer containing SM and 30 mol% cholesterol is equal to – 5 kJ/mol. It was also determined that the free energy of cholesterol transfer depends on the amount of cholesterol. Thus, when the concentration of cholesterol in the bilayers was 20% the free energy

of transfer was -8 kJ/mol (at 50^0C.)[140] Not only was the free energy of transfer determined in these experiments, but also the energetic and entropic components were determined. It was shown that the transfer is exothermic (with an enthalpy change of -13 kJ/mol at 30% of cholesterol and -23 kJ/mol at 20%) and therefore having an unfavorable entropic contribution upon transfer ($T\Delta S$=-8 kJ/mol at 30% cholesterol and -15 kJ/mol at 20 % cholesterol, 50 ^{0}C). Recently, we performed simulations where we compared the energetics of the cholesterol - POPC and cholesterol - SM interactions.[141] We observed from our simulations on binary bilayers containing cholesterol and SSM or cholesterol and POPC at 34 mol% cholesterol that the distribution of the interaction energy between cholesterol and the POPC is similar to the distribution of the interaction energy between SM and cholesterol. Moreover, the POPC-cholesterol average interaction energy was slightly more favorable than the SM-cholesterol average interaction energy. Superficially, this observation obtained from simulation may seem to be contradictory to the experimental result. Nevertheless, it is important to recognize that the energy of the interactions between phospholipids can change substantially with the cholesterol transfer due to the condensing effect caused by the presence of the cholesterol. As a result the total energy contributions from the cholesterol-phospholipid and the phospholipid-phospholipid interactions may change upon cholesterol transfer in any direction, depending on the balance of these interactions.

To obtain a better understanding of the balance of entropy and energy in the phospholipid-cholesterol interaction and also to get a quantitative measure of such a

balance, we performed simulations where we calculated the shape of the potential of mean force (pmf) for the cholesterol removal from bilayers containing SM or POPC lipids. We chose to perform our simulations at limited cholesterol dilution, so that we can have a clear understanding of how cholesterol-phospholipid interactions affect the free energy of cholesterol transfer from the POPC to the SM bilayer.

4.2 Computational Model

To calculate the relative affinity of cholesterol for different bilayers, such as 18:0 SM (SSM) or POPC we need to perform free energy calculations. We calculate the free energies of cholesterol removal from the POPC (ΔG_1) and SSM (ΔG_2) bilayers and determine the free energy of cholesterol transfer as a free energy difference ($\Delta\Delta G$) between these quantities.

We use the umbrella sampling method[142,143] to calculate the potential of mean force (pmf), i.e. the free energy of the cholesterol removal, as a function of distance between the cholesterol center-of mass and the center of a phospholipid bilayer, as cholesterol is slowly removed from the bilayer. Umbrella sampling technique was successfully employed to study the pmf for the insertion of liquid crystal-forming molecules into phospholipid bilayers[144], a problem similar to ours from the computational point of view. The difference in the pmf values when cholesterol is outside the bilayer and when it is inside the bilayer (for more accurate definition of this difference, see below) determines the free energy of cholesterol removal. Since we need to perform a large number of runs due to different windows in the umbrella

sampling, the size of the system we chose was not that large. Thus we considered systems containing only 35 phospholipid molecules and one cholesterol molecule in each leaflet of the bilayers, resulting in the presence of 72 lipid molecules in each bilayer solvated with 3600 water molecules. To avoid artifacts that could be created due to asymmetric amount of cholesterol in our small sized systems we placed one cholesterol molecule in every leaflet of the bilayer. The SPC/E water model [86] was employed in the simulations. The force field for the SM we used in our simulations was the same as used by Niemelä et al[90]. Force field parameters for the cholesterol were the same as used in the study by Pandit et al[72], and the force field for the POPC was obtained from Hoff et al.[125] Prior to using the umbrella sampling technique, 40ns runs were done to obtain the equilibrated initial configurations. All the simulations were performed using the GROMACS package[5,6]. The LINCS algorithm[83] was used to constrain all bond length in the systems, allowing an integration timestep of 2 fs. Periodic boundary conditions were applied in all three directions and the long range electrostatics was handled using the SPME algorithm [10] with a real space cut off of 10 Å, 4th order interpolation and a tolerance of 10^{-5} . A 10 Å cut-off was used for the van-der Waals interactions. All simulations were done in the NPT ensemble. We performed two series of simulations at two different temperatures, 319 K and 329 K. The temperatures in the simulations were maintained using the Nose-Hoover scheme[84] with a thermostat oscillatory relaxation period of 0.5 ps. The Parrinello-Rahman pressure coupling scheme[85] with a barostat time constant of 2.0 ps was used to keep the pressures at 1 atm. For the equilibrium

process, the semi-isotropic pressure coupling scheme was applied to give the correct area per headgroup in the systems and for the removal simulations, the isotropic pressure coupling was used. We found that this arrangement was needed, since it prevented the bilayer from falling apart during the cholesterol removal procedure.

The umbrella potential we used had a simple harmonic form:

$$U_{umb} = k_{umb}(h - h_0)^2 \tag{4.1}$$

with k_{umb} = 500.0 kJ/(mol*nm^2). In equation (1) h is the distance between the center of mass of the bilayer and the center of mass of the cholesterol molecule. The values of h_0 ranged from 0 nm to 3.8 nm with an increment of Δh_0 = 0.2 nm. We considered 20 windows for each system at a specified temperature, therefore performing 80 separate simulations in total. In each window we simulated the system for 30 ns, therefore the total simulation time in windows was 2.4 μs. The first 10 ns of every window simulation was considered to be the equilibrating time and was discarded, while the last 20 ns run was used for the free energy calculation. The potential of mean force curves for each bilayer were calculated using the weighted histogram analysis method (WHAM) formulas. [145]

4.3 Results.

We calculated the free energies for the process of cholesterol removal from the bilayers and show the results for the corresponding pmf's at temperature T=319 K in Fig. 4.1. In addition to pmf plots we also show the densities for the different components and groups in the bilayers, to aid in the understanding of these pmf's. Fig.

4.1 (a) and (b) show the density profiles of the various components in the SSM and POPC bilayers respectively. Every bilayer contains two cholesterol molecules in equilibrium, one in each leaflet. The density of cholesterol has been magnified on the Figure by a factor of 20 for clarity (therefore the asymmetry in cholesterol density that is observed in Fig. 4.1(b) is not as strong in reality). The density maxima for cholesterol in SSM bilayer are located around h = ±1.5 nm and those in POPC bilayer are around h = ±1.0 nm. These maxima are consistent with the previously observed stronger tilt of cholesterol axis in POPC[141]. Fig. 4.1 (c) and (d) display the shapes for the potentials of mean force curves obtained for the removal of cholesterol from the SSM and POPC bilayers, respectively.

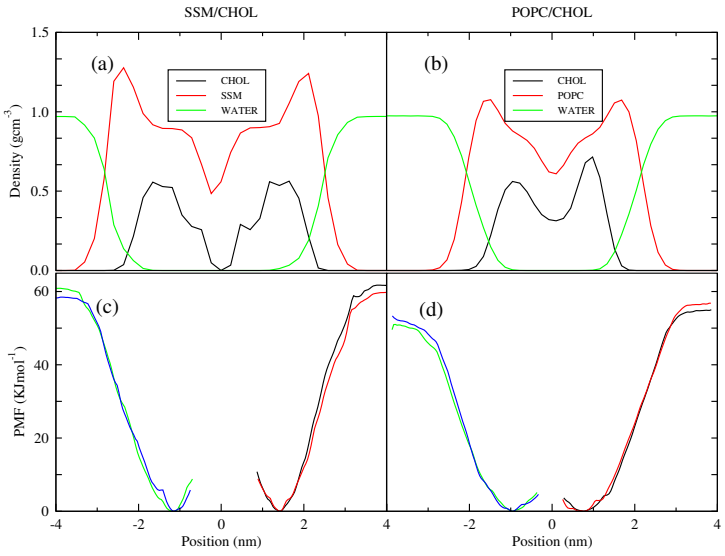

Fig 4.1 (a) Density profiles for different components in SSM/cholesterol bilayer. Dotted line: cholesterol; Solid line: SSM; Dash-dotted line: water. The density magnitude of cholesterol has been multiplied by a factor of 20 for clarity. (b) Same as (a), only for the POPC/cholesterol bilayer. (c) Potential of mean force (pmf) curves for the removal of a cholesterol molecule from the SSM bilayer. Left and right curves

are corresponding to cholesterol removal from one of the two leaflets of the bilayer. Each pmf is calculated twice by dividing the simulations in every window into two parts (each 10 ns long), thus producing a total of four pmf curves (d) Like (c) only for the removal of cholesterol from the POPC bilayer.

We observe that the minima in these curves are consistent with the positions of the maxima in cholesterol density profiles. From the pmf curves we calculated the free energy costs for removing a cholesterol molecule from an SSM bilayer or from a POPC bilayer as $\Delta G = G_{max} - G_{min}$, where G_{max} is the free energy value where the curve is flat in water and G_{min} is the minimum value on the curve which has been arbitrarily set to zero. The ΔG value in the SSM case is 60.5±0.6 kJ/mol, and the value in the POPC case is 54.0±1.0 kJ/mol. The error bars were obtained from four pmf curves for each bilayer. These four curves were produced by removing cholesterol from each leaflet of the bilayer and each removal pmf was calculated twice by dividing simulations into two halves, each half of 10 ns length.

According to our calculation, the difference in ΔG, $\Delta\Delta G$ that shows the free energy of cholesterol transfer from POPC to SSM at the limited cholesterol dilution is only -6.5±1.6 kJ/mol. This number is the measure of the relative affinity cholesterol has to SSM compared to POPC. Also, according to our calculations, cholesterol prefers the SSM bilayer to POPC bilayer.

To separate the free energy difference into its energetic and entropic parts, one can use numerical approximations to thermodynamic relationships. For this purpose,

one needs to calculate the free energy of cholesterol removal from the bilayer at few different (but close to each other) temperatures. We decided to perform our calculations at two different temperatures: 319K and 329 K, and use a simple approximation to the derivative that determines entropy at temperature T

$$-S(T) = \left(\frac{\partial G}{\partial T}\right)_P \approx \frac{G(T+\Delta T) - G(T)}{\Delta T} \tag{4.2}$$

Usually a central difference is used to approximate the derivative

$$\left(\frac{\partial G}{\partial T}\right)_P \approx \frac{G(T+\Delta T) - G(T-\Delta T)}{2\Delta T} \tag{4.3}$$

The use of central difference is numerically advantageous, but it requires calculations of free energy at two different temperatures T+ΔT and T-ΔT, in addition to the calculation of the free energy at temperature T. Since the calculations of free energies of transfer are expensive, we approximate the derivative as a one sided difference, i.e. we use eq. (4.2). In addition, since we want to find the energy/entropy contributions at T=319 K, which is close to the main transition temperature of the SSM bilayer, we did not want to perform calculations at temperatures below 319K.

We obtained from our simulations that at temperature T=329 K, the ΔG of cholesterol removal from the SSM is 55.0 ±0.8 kJ/mol, and the value for the removal from the POPC is 50.2±0.6 kJ/mol. Therefore, the difference in ΔG, ΔΔG, which shows the free energy of cholesterol transfer from POPC to SSM at limited cholesterol dilution, is -4.8±1.4 kJ/mol at T=329 K If we calculate the entropic contribution to the free energy, given by the term TΔS by using eq. (4.2), we get that it

is equal to –(121.2 ± 51.0) kJ/mol for the insertion of cholesterol into the POPC bilayer and it is –(175.5 ± 45.2) kJ/mol for cholesterol insertion into the SSM bilayer. This results in the unfavorable entropic contribution of –(54.3 ± 96.2) kJ/mol to the free energy when cholesterol is transferred from POPC to SSM. Here we observe that large error bars are present in the calculation of the free energy difference and entropy difference. These large error bars appear due to the accumulation of errors in the calculation of a difference between two similar in value numbers. Thus, although the error in the free energy calculation for the cholesterol removal (ΔG) is less than 1%, the error in the difference of the free energies (ΔΔG) is already much larger and it is around 30%. Use of equation (4.2) produces even larger relative error.

Since the free energy of cholesterol transfer from POPC to SSM is favorable, but entropy is not, this means that the transfer is promoted by the energetic component, i.e. transfer is exothermic. This is in a qualitative agreement with the recent results from the work of Heerklotz and collaborators.[140]

As we already mentioned, in our earlier simulations we observed that cholesterol interacts somewhat stronger with POPC than with SSM. Indeed, we calculated (at T=319K) the average interaction potential energy of a cholesterol molecule with the SSM bilayer and with the POPC bilayer respectively for the equilibrium positions of cholesterol in these bilayers. We observed that the average interaction potential energy of a cholesterol molecule with an SSM bilayer is equal to -(356.7 ± 0.7) kJ/mol, and that of a cholesterol with a POPC bilayer is –(360.8 ± 0.8) kJ/mol. (These energy values were calculated with a cut-off distance of 10 Å for the van der Waals

interaction and a cut-off distance of 18 Å for the electrostatic interaction. The last 20 ns of the equilibrating run were used for these calculations). Based just on the interaction energy between cholesterol and phospholipids we would conclude that transfer is endothermic and it should be driven by entropy. The fact that the transfer is exothermic is due to a large change in the interaction energy between the phospholipid molecules when cholesterol is removed from or inserted into the bilayer. Moreover, since we deal with the transfer process, it is the delicate balance of the *difference* in the removal energies and entropies that determines if the transfer is exothermic or endothermic. Indeed, it is possible that there are cases where favorable transfer is dominated by entropy.

4.4 Discussion.

At this point, let us try to understand why the transfer of cholesterol from the POPC to SSM bilayer at limited cholesterol concentration produces favorable change in energy and unfavorable entropy change. With this in mind, we first consider the chain order parameters S_{CD} for lipids that are next to cholesterol and that are away from cholesterol upon cholesterol insertion. These are shown in Fig. 4.2. It is also useful to compare these order parameters with the ones we calculated from our previous simulations performed without cholesterol and at concentration of 34 mol%[141]. As we observed in our previous simulations, the change in the S_{CD} order parameters for the tails of POPC molecules was rather large when we added a large amount of cholesterol.

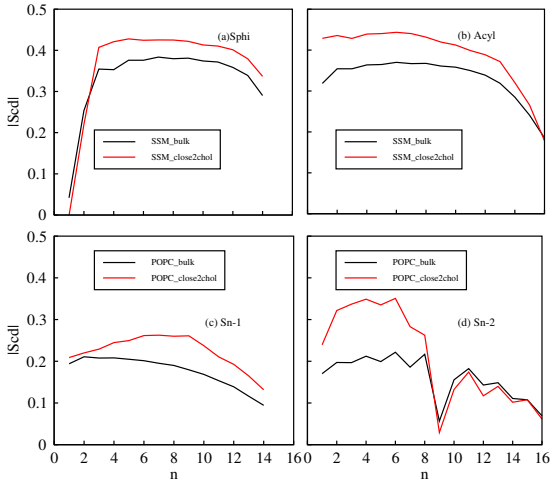

Fig 4.2 Deuterium order parameters of lipid tails in SSM and POPC bilayer with one cholesterol molecule in each leaflet. Order parameter profiles of lipid tails in bulk are represented by solid curves while dashed curves are for the order parameter profiles of lipid tails close to cholesterol molecule.

At limited cholesterol dilution, the change is not that large for the POPC molecules neighboring cholesterol and, for the phospholipids far away from cholesterol, the order parameters are the same as in a bilayer containing pure phospholipids. Thus the change in order parameters of POPC tails depends on cholesterol concentration. The change in the order parameters for the SSM lipid next to cholesterol in our present simulation is nearly the same as in our previous simulation, i.e. we observe that the change in S_{CD} does not depend on the cholesterol concentration in the case of a SSM bilayer. In addition, we observed that cholesterol orients itself in the phospholipid bilayer in such a way as to produce a small interference. In the case of SSM bilayer, cholesterol can achieve this by inserting itself in an almost vertical position into the bilayer, where it fits nicely due to its matching hydrophobic length.

In the case of the POPC bilayer, cholesterol will try to orient itself to avoid the misfit of hydrophobic lengths and it can do this, since the free volume in POPC is larger than in SSM. By orienting itself at an angle to the bilayer normal, cholesterol produces smaller interference on the tail conformational motion of neighboring POPC lipids and, therefore, causes a smaller reduction of the lipid chain conformational entropy. In Fig. 4.3, we present the distribution for the orientational angle, which is defined as the ang bilayer.

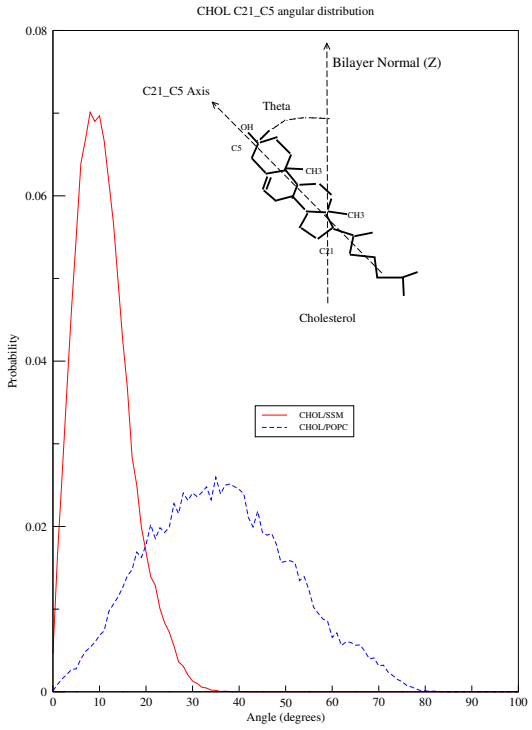

Fig 4.3 Cholesterol principal axis tilt angle distributions with respect to bilayer normal in SSM (solid) and POPC (dotted) bilayers. The cholesterol principal axis is defined as the vector connecting carbon C21 (the carbon atom in the ring to which the short tail of a cholesterol molecule is attached) and C5 (the carbon atom to which the hydroxyl group binds).

While the distribution is narrow for cholesterol in SSM, it is broad for cholesterol in POPC, much broader compared to the distribution obtained when cholesterol concentration was 34 mol%.[141] The broadness of the distribution in the POPC that we observe at limited cholesterol dilution indicates that cholesterol undergoes rotational (librational) motion in this bilayer, when its concentration is small. Indeed, this can be confirmed by considering the change of the cholesterol tilt angle as a function of time (plot is not shown). Fig. 4.4 presents snapshots from our simulations and the difference in the orientation of cholesterol with respect to the bilayer normal in POPC and in SSM bilayers is clearly seen from this Figure.

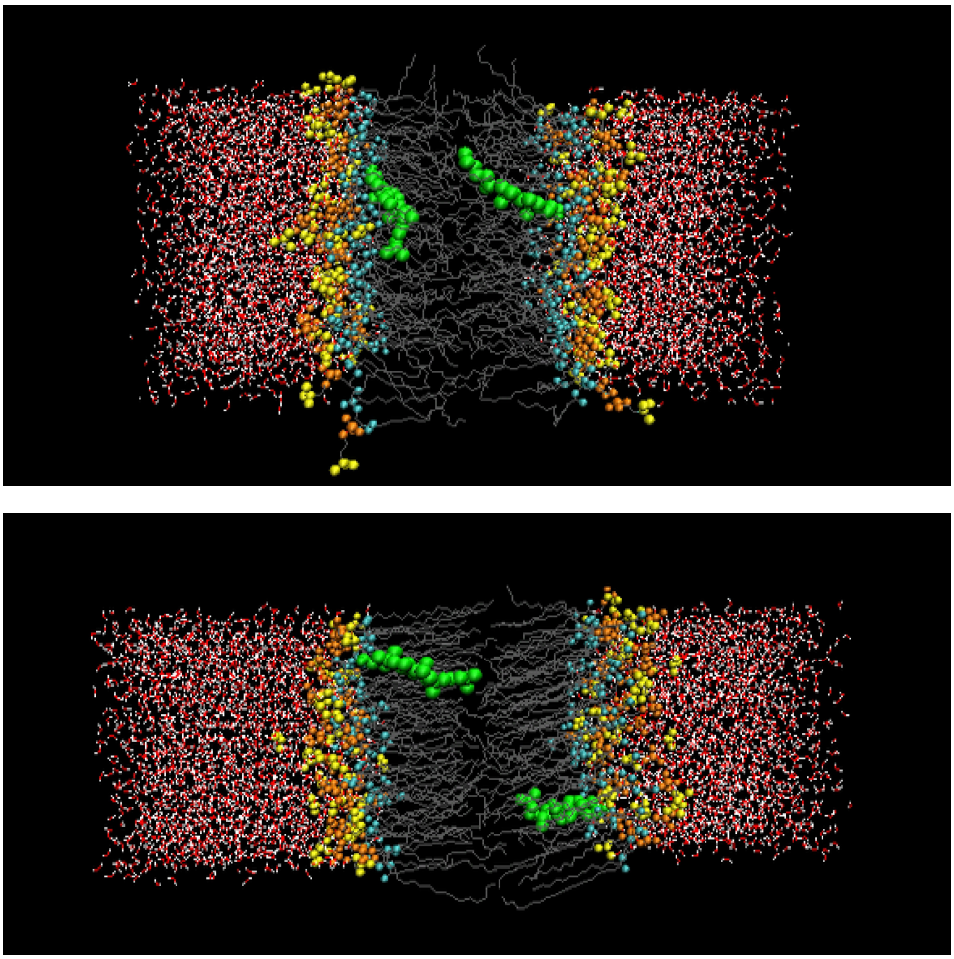

Fig 4.4 Snapshots of the lipid bilayers used in our study. The upper one is depicting the POPC bilayer with two cholesterol molecules (green) in it (water: red and white; choline: yellow; phosphate: orange; glycerol and carbonyl: blue; carbon tails: grey). The lower one is of the SSM bilayer with two cholesterol molecules (green) in it (water: red and white; choline: yellow; phosphate: orange; sphingosine: blue; carbon tails: grey).

Thus, when cholesterol is transferred from the POPC to SSM, its rotational entropy is reduced and, perhaps, a small relative change in the conformational entropy of the surrounding lipids is produced. As a result we loose in total entropy. The balance in total energy is also dependent on the inclination of cholesterol, since it perturbs the tail ordering of neighboring lipids and subsequently the van der Waals interactions between them. To determine the values of the change in the conformational and rotational entropies and cholesterol-phospholipid and phospholipid-phospholipid interactions as a result of cholesterol transfer, one will need to perform very careful and detailed calculations. At this stage, it is hard to perform such, since the methodology of calculations of free energy components is not well developed and the calculations by using the derivatives of free energy suffer from large uncertainties in the results. In addition, the results of the calculations will be very sensitive to the force fields used. Nevertheless, as we can see, the present calculations are in qualitative agreement with the experiment by determining that the transfer of cholesterol from the POPC bilayer to SSM bilayer has an exothermic

character with a loss of entropy. Although the experiment was not performed at limited cholesterol dilution, it showed a dependence on the cholesterol concentration. Thus, from experiment, the free energy change for a process of cholesterol transfer from the POPC bilayer to SM at 30 mol% cholesterol was -5 kJ/mol and it decreased to -8 kJ/mol when cholesterol concentration decreased to 20 mol%. We do not know if the measured free energy change decreases monotonically with the decrease in cholesterol concentration, but if it is, the calculated free energy change is having a value that is below the experimental value, and a refinement of the force field for cholesterol interactions with phospholipid molecules will be required. This indicates the need for further experiments and calculations that will produce consistent results. At this stage we want to point out that our calculations and available experimental data show that the free energy of cholesterol transfer from POPC to SM is only ~5-10 kJ/mol. Assuming that this energy is equally distributed between 4-6 phospholipid neighbors interacting with the cholesterol we get a change of ~ 1-2 kJ/mol per interacting pair upon cholesterol transfer. The same order of magnitude in free energy change was also obtained in the resent experiments performed by Frazier et al.[137] who studied BSM/POPC/cholesterol bilayer using Fluorescence Resonance Energy Transfer technique.

At this stage we believe that simulations can provide a qualitative and even semi-quantitative insight into the complicated (and system sensitive) energetics of cholesterol transfer from one lipid bilayer to another. This can be very helpful for the understanding of lipid raft formation.

CHAPTER 5: ORIENTATIONAL DYNAMICS OF WATER IN PHOSPHOLIPID BILAYERS WITH DIFFERENT HYDRATION LEVELS

5.1 Introduction

Biological molecules such as proteins and DNA or biological assemblies such as cell membranes need to be hydrated by water for their proper functioning. Therefore, it is important to understand the structural and dynamical properties of water molecules located at the interface with bio-molecules. Often this interfacial water is called "biological water". Different experimental techniques were employed to study the structure and dynamics of biological water, techniques such as NMR[146-148], time-resolved fluorescence spectroscopy[149-153], neutron scattering[148,154,155], ultrafast vibrational spectroscopy and IR absorption spectroscopy[156]. Computer simulation techniques such as molecular dynamics and Monte Carlo were also applied to study properties of biological water molecules[157], including simulations that studied the interface of water with model membranes or micelles[155,158-160]. Experimental and computational studies showed that biological water is characterized by slowed down orientational dynamics, although the mechanism for such a slowing-down is not clarified yet. Even for bulk water, the mechanism of reorientational relaxation was clarified only recently. Previously it was assumed that reorientational relaxation

occurs via diffusion; recent work based on a series of computer simulation studies of water in bulk or in the hydration layer of an anionic solute[161-163] demonstrated that water reorientation occurs via molecular jump mechanism or MJM. In MJM water molecules reorient mainly through the switch of hydrogen bond acceptor during which the water OH bond experiences a large angular jump. Using the results from molecular dynamics simulations Jana[157] et al. studied the H-bond breaking mechanism and water reorientation dynamics in the hydration layer of lysozyme protein and found that a large angular jump was common for all the reorientation processes in this case also.

Below we present the results obtained from a molecular dynamics simulation study of water next to phospholipid bilayers with different hydration levels. The simulations were devised according to a recent experimental study by Zhao et al[156]. We studied water reorientational relaxation and compared the results to the experimental ones. We also discuss possible mechanisms for the slowing down of the orientational relaxation of water molecules located in the hydration layer of lipid bilayers.

5. 2 Computational Details

We performed simulations on six different systems: each containing a dilauroyl -phosphatidylcholine (DLPC) lipid bilayer but with a different hydration level. The amount of water (x) was 2, 4, 6, 8, 16 and 32 molecules per lipid. Each bilayer in the simulation cell contained 128 lipids, with 64 lipids in each leaflet. We also simulated

bulk water with a simulation cell containing 2180 water molecules. The force field for the DLPC we used was the same as that of dipalmitoyl-phosphatidylcholine (DPPC)[87], except that the last four united carbon atoms in each lipid tail were removed. The SPC/E[86] model was employed to describe water molecules. All simulations were done with the GROMACS package[5,6]. The LINCS algorithm[83] was used to constrain all bond lengths in the system. Periodic boundary conditions were applied in all three dimensions and long range electrostatics was handled with SPME algorithm[10] with a real space cutoff of 10 Å. A 10 Å cut-off was used for Van der Waals' interactions. All simulations were done in NPT ensemble. The temperature in the simulations was maintained at 310 K using the Nose-Hoover scheme[84] with a thermostat oscillatory relaxation period of 0.5ps. The Parrinello-Rahman semi-isotropic pressure coupling[85] scheme with a barostat time constant of 2.0ps was used to keep the pressures at 1 atm

The simulations for the bilayers were equilibrated for 15ns, with a time step of 2 fs. After that a 1ns production simulation was done for each system with the same time step but with configurations saved every 10 fs. The first 300 ps of these 1ns trajectories were used for the calculation of the OH bond reorientation correlation functions that were evaluated until t= 5 ps. Following the 1 ns production simulation, a 100 ps simulation was done for each system with a time step of 0.5 fs and configurations were saved every step to capture the hydrogen bonding switching events. Another 1ns simulation was also done for the 2 -16 water/lipid case with coordinates saved every 100 fs to study the OH bond reorientation decay over longer time lengths. Bulk water simulation analysis was done in a similar way, except that

the equilibrium time was 5 ns, instead of 15 ns.

5.3 Results and Discussion

5.3.1 Water OH Bond Orientational Relaxation

The anisotropy decay measured in experiment is connected to the orientational relaxation of the OH bond in water by the following equation:

$$R(\mathrm{t}) = 0.4 * P_2(\mu(t)*\mu(0)) \quad (1)$$

In eq. (1) P_2 is the second order Legendre polynomial, and μ is the unit vector along the OH bond in water molecules. The orientational relaxations calculated from the simulation data are shown in Fig. 5.1. The relaxation curves from simulations are in a good qualitative agreement with the experimental findings of Zhao et al[156], though the latter investigated the OD orientation relaxation in water. The trend that the relaxation slows down as the hydration level decreases was well reproduced in the simulations. Even when the hydration level is at 32 water molecules per lipid (this water/lipid case was absent in experiment), the relaxation is still slower than in the bulk water, although the hydration level in this case is above the full hydration level of the bilayer, indicating a significant influence of the lipids on the dynamics of water molecules. As in the experiment, we fitted all the relaxation curves in bilayers with biexponential decay and the relaxation curve for bulk water with a single exponential decay. Since the experimental conditions do not allow to register the data for anisotropy decay at short times of order 0.1-0.2 ps we also fitted the data from the simulations over time interval t = 0.2 ps to t = 5 ps. The biexponential fitting was done using the following

form:

$$R(t) = A_1*\exp(-t/\tau_1) + A_2*\exp(-t/\tau_2) \quad (2)$$

with the single exponential fit done in a similar manner but using just one exponent.

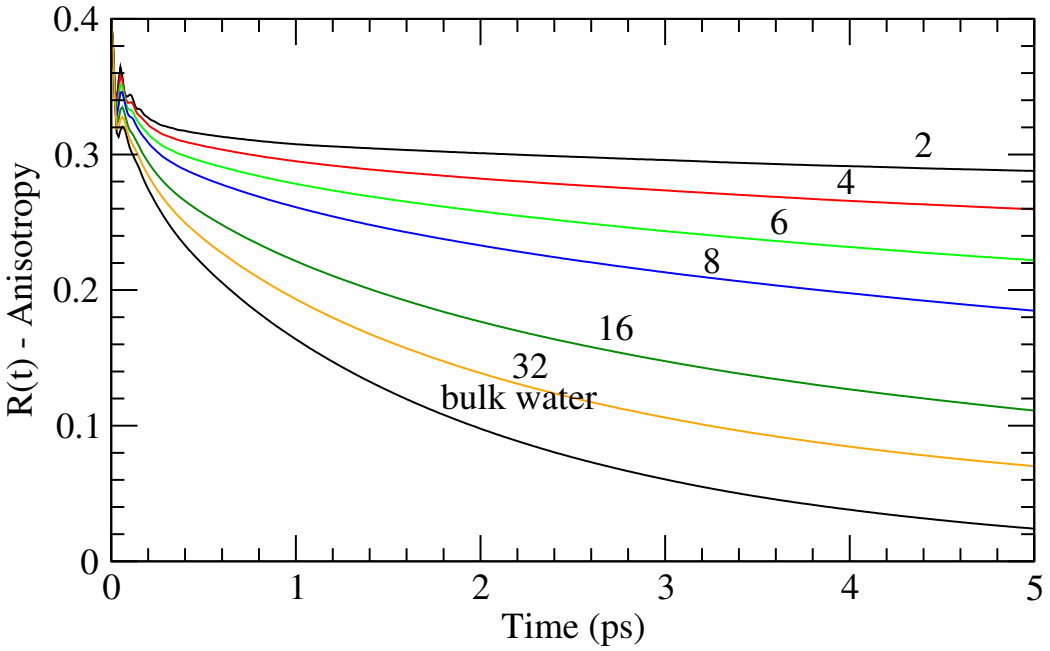

Fig. 5.1 Anisotropy decays (orientational relaxation) of the OH bond in water next to hydrated bilayers containing different numbers of water molecules per lipid and in bulk water. The correlation function is calculated over time period of 5 ps.

Sample	A_1	τ_1 (ps)	A_2	τ_2 (ps)
X=2	0.024	0.466	0.310	65.70
X=4	0.035	0.600	0.298	34.98
X=6	0.046	0.740	0.285	19.49
X=8	0.062	0.800	0.269	13.46
X=16	0.093	0.880	0.220	7.19
X=32	0.123	0.900	0.180	4.79
Bulk Water	0.287	1.900	-------	-------

Table 5.1 Biexponential fit parameters for the orientational relaxation R(t) when the correlation function is calculated over time period of t=5 ps.

While the shape and the qualitative behavior of the curves obtained from experiments and from simulations look quite similar, there is a quantitative difference between the simulation and experiment. Thus in experiment the short time decay exponents (τ_1) had a nearly constant value of ~0.4 ps; in our simulations they constantly increased as the degree of hydration increased. The longer time relaxation exponent decreased both in the simulation and in the experiment. From the data in experiment it was assumed that the τ_2 value in the case of x= 2 is practically equal infinity, while in our simulation it has a final value of ~ 66 ps. The values of τ_2 at x=8 and 16 we obtained are close to the ones reported in the experiment.

Zhao et al[156], based on a theoretical analysis[164] as well as on the previous MD simulations that studied water in reverse micelles[155,159,160], excluded the possibility that the biexponential decay arises from two single exponential decays associated with two populations of water in distinct environments. In those simulations it was found that water molecules both inside the headgroup region and away from that region displayed a decay that was fitted to a biexponential function. In our previous simulations on water next to phospholipid membranes we also observed the same phenomenon[158,165]. Using the data from the present simulations we have calculated the OH orientation decay for water molecules inside the headgroup region and outside of that region in the bilayer with 16 water molecules per lipid, and the results are shown in Fig. 5.2.

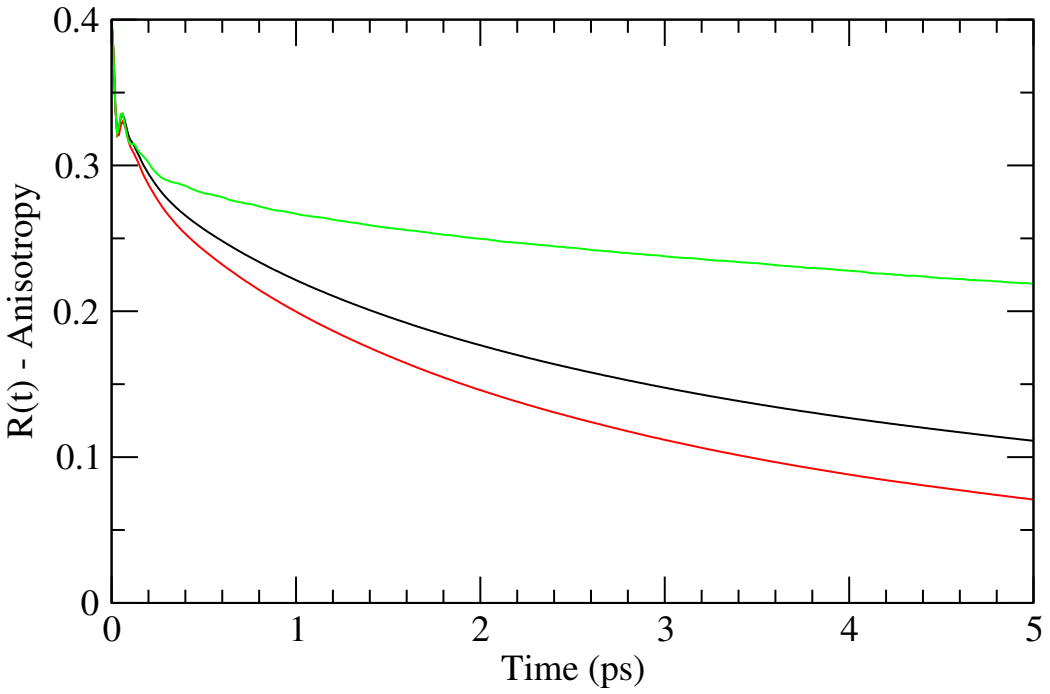

Fig. 5.2 Anisotropy decays of the OH bond in the bilayer with x=16. The black curve is for all water molecules, green for water inside the headgroups, red outside headgroups.

We observed that the curves for the total relaxation of the whole population of water in the system., for the slow relaxation of water molecules that stayed inside the headgroup region during the calculation time and for the fast relaxation of water molecules that stayed outside of the headgroup region, all these three curves from Fig. 5.2 follow biexponential decay (the fitting parameters are not shown). Moreover, for the cases when the hydration level is low, the biexponential decay cannot arise from two water populations since there are no two water populations in the system. For example, in the case of 2 water molecules per lipid, the maxima in the density distributions (data not shown) of water and phosphate group overlap, indicating that

there is no bulk-like water in this case.

What is the origin of the biexponential decay in the correlation functions? Zhao et al. suggested that the biexponential orientational relaxation arises from a fast wobbling-in-the-cone[166] motion followed by a complete orientational randomization. The wobbling-in-the-cone motion is responsible for the decay of the correlation function on a ~ 0.3-0.4 ps time scale, and corresponds to a tumbling motion of single water molecules. While the simulations show that there is also an initial fast orientational relaxation decay on a ~100 fs scale corresponding to the fast librational motion of water molecules, this decay is not observed in experiments of Zhao et al. due to the inability of capturing it. Therefore the experimental data were collected for the time intervals beyond 200 fs and all the exponential fits in that work were done with starting time at 200 fs. The long time decay, according to Zhao et al., corresponds to the collective motion of the water molecules, producing the restructuring of the hydrogen bonding network. The MJM mechanism, mentioned in the Introduction, may be used to explain this long time decay.

Are the time constants obtained from the biexponential fits to experimental data and simulations reflecting some physical reality? To answer this question, we calculated for the case of 16 waters/lipid, the decay of the correlation function up to 10, 20, 30, 40 and 50 ps from the trajectories saved every 100 fs, as mentioned in the computational details section. We fitted these curves with biexponential as well as three-exponential decays to see if the fitting parameters depend on the time interval over which we calculated the correlation function. The fitting parameters we obtained

are listed in Table 5.2.

(a) Three Exponential Fit							
Time(ps)	A_1	τ_1 (ps)	A_2	τ_2 (ps)	A_3	τ_3 (ps)	R
0.2-10	0.053	0.261	0.143	2.13	0.14	13.93	0.9999
0.2-20	0.061	0.513	0.165	3.17	0.096	24.14	0.9999
0.2-30	0.071	0.687	0.165	3.74	0.082	30.38	0.9999
0.2-40	0.083	0.887	0.158	4.33	0.072	36.12	0.9999
0.2-50	0.094	1.058	0.151	4.87	0.066	40.80	0.9999
(b) Biexponential Fit							
Time (ps)		A_1	τ_1 (ps)	A_2	τ_2 (ps)	R	
0.2-10		0.134	1.448	0.172	10.882	0.9996	
0.2-20		0.175	2.228	0.120	18.359	0.9994	
0.2-30		0.189	2.626	0.101	23.940	0.9992	
0.2-40		0.196	2.900	0.090	28.488	0.9989	
0.2-50		0.201	3.090	0.080	32.028	0.9987	
(c) Stretched Exponential Fit							
Time (ps)		A_0	τ (ps)	β	R		
0.2-10		0.37	3.58	0.52	0.9998		
0.2-20		0.42	2.68	0.42	0.9991		
0.2-30		0.46	2.06	0.38	0.9987		
0.2-40		0.42	2.68	0.41	0.9980		
0.2-50		0.53	1.41	0.34	0.9983		

Table 5.2 Parameters for the multiexponential fits of the orientational relaxation R(t) for water molecules in the case of 16 water molecules per lipid bilayer when the correlation functions have been calculated over different time intervals.

It can be seen that none of the values of the parameters are preserved when the fittings are done over different time spans or when a three exponential formula is used instead of a biexponential one, even when the fittings are done for the same time span. Each individual fitting has quite a good correlation coefficient R, though the three exponential fittings are better in all cases. Some of the correlation functions are fitted to curves having a stretched exponential form, i.e.

$$R(t) = A_0 * \exp(-t/\tau)\beta \quad (3)$$

We also fitted the correlation functions in x=16 case to the stretched exponental curves and the data are included in Table 2. Again, as in the case of multiexponential fits, the values of the parameters are not preserved when the fittings are done over different time spans. In addition, we observed that the quality of the multiexponential fits was higher than the stretched exponential. Furthermore, we extended the length of our simulations and obtained the correlation functions for the time period of up to

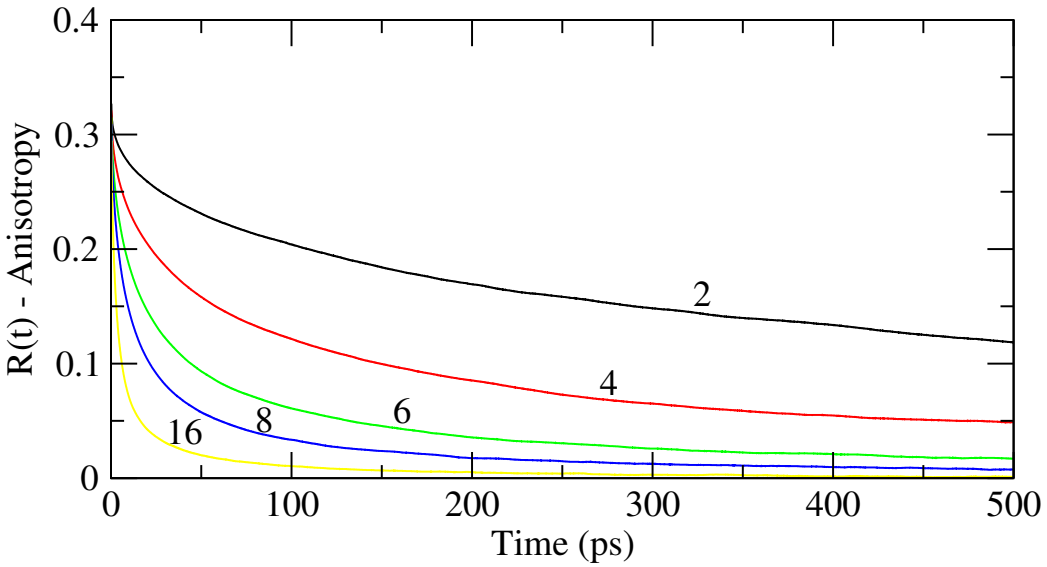

Fig. 5.3 Anisotropy decays (orientational relaxation) of the OH bond in water next to hydrated bilayers containing different numbers of water molecules per lipid. The correlation function is calculated over time period of 500 ps.

We fitted the correlation functions to biexponential curves and also to three-exponential curves and again, in all cases, obtained better results for the fits

using three exponents, although the biexponential fits produce sets with more consistent behavior, as the level of hydration was changing. The results of the fits are given in Table 3. As we can see from the comparison of data in Tables 1-3, the characteristic time scales that appear in the fits depend on the time period for which we know the values of the correlation function. If the correlation function is observed for the longer time period, the fitted curves contain exponents with longer time decay.

(a) Three Exponential Fit							
Sample	A_1	τ_1 (ps)	A_2	τ_2 (ps)	A_3	τ_3 (ps)	R
X=2	0.037	7.25	0.075	83.1	0.20	945.8	0.9998
X=4	0.088	13.90	0.137	116.6	0.07	1138.8	0.9997
X=6	0.108	6.85	0.128	46.75	0.06	380.7	0.9996
X=8	0.134	5.14	0.120	34.39	0.036	297.9	0.9995
X=16	0.190	2.72	0.08	18.5	0.02	157.7	0.9995
(b) Biexponential Fit							
Sample		A_1	τ_1 (ps)	A_2	τ_2 (ps)	R	
X=2		0.079	52.8	0.21	837.3	0.9990	
X=4		0.138	39.8	0.13	442.6	0.9975	
X=6		0.178	23.8	0.081	274.4	0.9966	
X=8		0.194	15.6	0.057	195.2	0.9950	
X=16		0.216	4.89	0.044	77.6	0.9950	

Table 5.3 Parameters for the multiexponential fits of the orientational relaxation R(t) for water molecules when the correlation functions have been calculated over 500 ps time interval.

Is this decay a reflection of some collective process over longer length scale and therefore requiring longer time to happen? The answer to this question can be given by considering a detailed model of the correlation function and a construction of such a model and its testing should be pursued. Nevertheless, from Tables 5.1- 5.3 we

observe that longer time of the observation of the correlation function allows the observation of the relaxation component that has longer time decay. At the same time, the decay with a short time component gets suppressed.

Finally, some more words of caution. We also observed previously[158,165] that the obtained values for the exponents depend on the presumed behavior of the correlation function over very long times. If we presume that the correlation function does not decay to zero, but instead decays to a plateau, the exponent with very long time decay changes substantially[165]. Therefore, one should be careful when one assigns a physical interpretation to the time exponents obtained from the fits, especially to the long time exponents.

5.3.2 Water Hydrogen Bonding

In bulk liquid water molecules are hydrogen bonded to each other and these hydrogen bonds form and break all the time. In hydrated lipid bilayers some water molecules penetrate into the headgroup region and hydrogen bond with oxygens of the phosphate groups and to a lesser extent with oxygens of the carbonyl groups. Previously we observed that these hydrogen bonds are stronger[158] than water-water hydrogen bonds. Moreover, lipid molecules are much heavier than water molecules and therefore diffuse slower. Therefore we expect that hydrogen bonds between water and lipid headgroups break less frequently. This can be quantified by calculating the average lifetimes of hydrogen bonds in bilayers with different hydration levels and comparing these data with the average lifetime of hydrogen bond in bulk water. Table 4 shows the average hydrogen bond lifetimes in different cases. (We use the geometric

criterion for the hydrogen bond definition: Roo < 3.5 Å, θ_{HOO} < 30°, where Roo is the distance between the donor and acceptor oxygen atoms, and θ_{HOO} is the angle between the OH bond and the OO vector. The hydrogen bonding lifetime is defined as the time during which the hydrogen bond exists continuously without any interruptions.)

Hydration Level (water/lipid)	Hydrogen bonding lifetime (fs)
2	413
4	342
6	295
8	264
16	214
Bulk Water	179

Table 5.4 Hydrogen bonding lifetime of water in bilayers and in bulk water.

The table shows that the average lifetime of a hydrogen bond for water in bilayers increases as the hydration level decreases. Thus, the average lifetime in the bilayer with two water molecules per lipid is larger by an order of magnitude compared to bulk water. The increase in hydrogen bonding lifetime is consistent with the slowing down of water OH bond reorientational relaxation in each case: the longevity of hydrogen bonding means slow relaxation.

As we can see from Table 5.4, hydrogen bonds live longer in lipid bilayers with low hydration level, but still they break up and form again. It was shown recently that the dynamics of hydrogen bond switching in bulk water follows a so called “jump” model: the O-H bond vector will have a fast and large angular jump during a hydrogen bond switch event[161]. To see if this kind of a quick switch also happens in our simulations, when a water molecule changes its hydrogen bonding acceptor, we

first monitored the hydrogen bonding switch events in the bilayer with the lowest hydration level with x=2, and monitored the directional change of the OH bond vector as a function of time. For our study we considered one of the OH bonds from each of 256 water molecules as hydrogen bonding donor and monitored the acceptors' change during a 20ps simulation. Of the 256 H-bonding donor groups, 35 (14%) of them never H-bonded to any acceptor as a donor during 20 ps observation time; 86 (34%) of them H-bonded to the same acceptor during this time span as a donor (these H-bonds do not have to exist continuously). The rest 52% of OH bonds H-bonded to more than one acceptor during this time. From all the H-bonding switches (more than one switch can happen to an OH donor during 20 ps) we were able to observe only 15 clear jump motions of the OH bond orientation, which is less than 10% of all the switches. Fig. 5.4 depicts such an event. (The angle is defined in the same way as in the original jump model, namely, between the projection of the OH vector on the OOaOb plane and the bisector of the OaOOb angle. Atom Oa and Ob are the acc

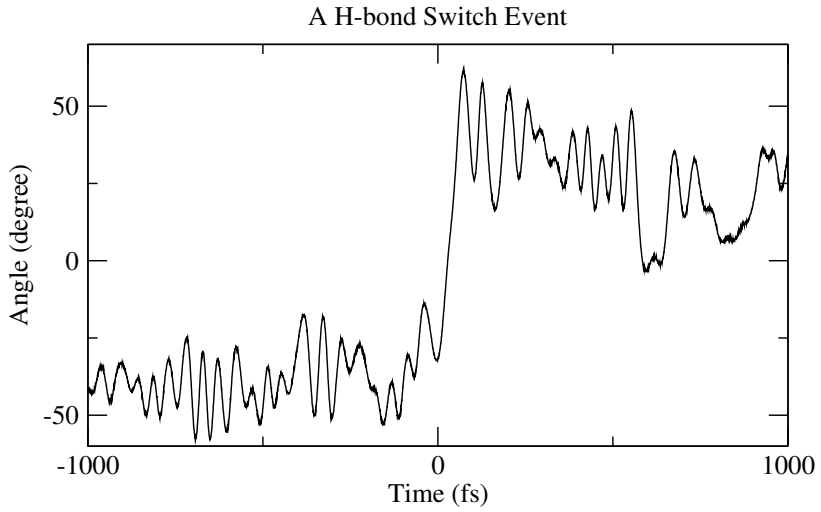

Fig 5.4 The angle (see the text for the definition) as a function of time during a

H-bond switching event.

We also observed other H-bonding switches when no clear jump was present. These usually happened when the switch was not complete and the OH donor group was just more or less shared by two acceptors, e.g. two oxygen atoms from the same phosphate group in a lipid. According to our observations, a water molecule as the old or new acceptor or two water molecules as both old and new acceptors were usually found to be involved in a jump. We observed that a clear jump motion of an OH bond is rare in bilayers with small hydration level due to the lack of water molecules. As the hydration level increased, the number of available waters for the jump mechanism to take place also increased. This explains why the OH bond relaxation becomes faster as the hydration level of the bilayer increases: there are more water molecules moving around to break and form hydrogen bonds and as a result, facilitate the orientation of OH bond in water molecules. We studied the H-bonding switch in the case of 8 water molecules per lipid. In this case H-bonding switch happened much more frequently due to the higher level of hydration. We observed that when a clear switch happened, which means the H-bonds before and after the switch lived for a long time, the OH bond orientation always followed a jump motion.

5.4 Conclusions

Following the conditions from a recent experimental work[156], we simulated DLPC lipid bilayers with hydration levels of 2 to 16 water molecules for each lipid

and also considered bilayers with 32 water molecules per lipid. The water OH bond reorientation relaxation in these different bilayers were calculated and found to agree with experimental results qualitatively and even semi-quantitatively, when the data from our calculations and experiment were compared over the same time period. Our results also showed that the orientational relaxation can be fitted nicely by curves containing a sum of two or three exponentials. We observed that the values of time constants from the fits depend on the amount of the information we have for the correlation function, i.e. on the length of time for which the correlation is measured and also on the resolution. If we measured the correlation over longer times we observed that multi-exponent curves contain components with longer time decays, although the relative contribution of these components is small. Moreover, some of the long time exponents may be an artifact due to the extrapolation of the data, in cases when the data actually reach a plateau value in longer times. Therefore we would not recommend emphasizing either the functional form of the fit to the correlation function or the time constants obtained from such a fit. We tried different fits for some of the relaxation curves and conclude that interpretation of the physical process according to a particular fitting formula can be misleading. We also would like to suggest that "wobbling in the cone" model is perhaps not a correct model, since it is hard to define the physics of the wobbling motion and the "cone". We observed that the hydrogen bonding lifetime of water increases, as the hydration level decreases in bilayers, and this is correlated with the slower reorientation relaxation of water OH bond. The simulations show that hydrogen bonding switch follows the

molecular jump model[161] and we confirm that such jumps are more difficult to accomplish at low hydration level, as was suggested by Zhao et al[156]. Interestingly, we observed very long relaxation decay times for very low hydration levels of x=2 and 4 waters/lipid only after we calculated the relaxation function for 500 ps, while the long relaxation decay in experiment was already observed with the data gathered for up to 5 ps. This may be due to the difference in the conditions in the experiments and simulations.

REFERENCES

(1) Alder, B. J.; Wainwright, T. E. *Journal of Chemical Physics* 1957, *27*, 1208.

(2) McCammon, J. A.; M, K. *Nature* 1977, *268*, 765.

(3) Freddolino, P. L.; Arkhipov, A. S.; Larson, S. B.; McPherson, A.; Schulten, K. *Structure* 2006, *14*, 437.

(4) Jayachandran, G.; Vishal, V.; Pande, V. S. *Journal of Chemical Physics* 2006, *124*, 164902.

(5) Berendsen, H. J. C.; Van der Spoel, D.; Van Drunen, R. *Comp. Phys. Comm.* 1995, *91*, 43.

(6) Lindahl, E.; Hess, B.; Van der Spoel, D. *J. Mol. Mod.* 2001, *7*, 306.

(7) Ryckert, J. P.; Bellemans, A. *Chemical Physics Letters* 1975, *30*, 123.

(8) Ewald, P. P. *ANN. Phys.* 1921, *64*, 253.

(9) Darden, T.; York, D.; Pedersen, L. G. *J. Chem. Phys* 1993, *98*, 10089.

(10) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys* 1995, *103*, 8577.

(11) Morse, P. M. *Phys. Rev.* 1929, *34*, 57.

(12) Verlet, L. *Phys. Rev.* 1967, *159*, 98.

(13) Hochney, R. W.; Eastwood, J. W. *Computer Simulations Using Particles*; McGraw-Hill: London, 1981.

(14) Rahman, A. *Phys. Rev.* 1964, *136*.

(15) Rahman, A.; Stillinger, F. H. *J. Chem. Phys* 1971, *55*, 3336.

(16) Schofield, P. *Comp. Phys. Comm* 1973, *5*, 17.

(17) Berendsen, H. J. C.; Van Gunsteren, W. F. *Molecular-Dynamics Simulation of Statistical-Mechanical Systems*; North-Holland: Amsterdam, 1986.

(18) Warshel, A.; Levvit, M. *J.Mol.Biol* 1976, *103*, 227.

(19) McQuarrie *Statistical Mechanics*; Harper & Row: New York, 1976.

(20) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: Cambridge, UK, 1989.

(21) Brooks, C. L.; M, K.; Pettitt, M. *A Theoretical Perspective of Dynamics, Structure and Thermodynamics, Advances in Chemical Physics*; John Wiley & Sons: New York, 1989; Vol. LXXI.

(22) Bash, P. A.; Singh, U. C.; Langridge, R.; Kollman, P. A. *Science* 1987, *236*, 564.

(23) Northrup, S. H.; Pear, M. R.; Lee, C. Y.; McCammon, J. A.; M, K. *Proceedings of the National Academy of Sciences* 1982, *79*, 4035.

(24) Pangali, C.; Rao, M.; Berne, B., J. *J. Chem. Phys* 1979, *71*, 2975.

(25) Valleau, J. P.; Card, D. N. *J. Chem. Phys* 1972, *57*, 5457.

(26) Patey, G. N.; Valleau, J. P. *J. Chem. Phys* 1975, *63*, 2334.

(27) Mezei, M.; Mehrotra, P. K.; Beveridge, D. L. *J. Am. Chem. Soc* 1985, *107*, 2239.

(28) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comp. Chem* 1992, *13*, 1011.

(29) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Letts* 1989, *63*, 1195.

(30) Simons, K.; Van Meer, G. *Biochemistry* 1988, *27*, 6197.

(31) Van Meer, G.; Simons, K. *J. Cell. Biochem* 1988, *36*.

(32) Simons, K.; Ikonen, E. *Nature* 1997, *387*, 569.

(33) Draber, P. *Trends in Glycoscience and Glycotechnology* 2001, *13*, 261.

(34) Simons, K.; Toomre, D. *Nature Reviews Molecular Cell Biology* 2000, *1*, 31.

(35) Baird, B.; Sheets, E. D. *Biophysical Chemistry* 1999, *82*, 109.

(36) Simons, K.; Toomre, D. *Science* 2000, *290*, 1721.

(37) Brown, D. A.; London, E. *Annu. Rev. Cell Dev. Biol.* 1998, *14*, 111.

(38) Brown, D. A.; London, E. *Journal of Membrane Biology* 1998, *164*, 103.

(39) Pike, L. J. *Journal of Lipid Research* 2006, *47*, 1597.

(40) Carter, W. G.; Hakomori, S. *J. Bio. Chem* 1981, *256*, 6953.

(41) Okada, Y.; Mugnai, G.; Bremer, E. G.; Hakomori, S. *Exp. Cell. Res* 1984, *155*, 448.

(42) Brown, D.; Rose, J. *Cell* 1992, *68*, 533.

(43) Parton, R. G.; Simons, K. *Science* 1995, *269*, 1398.

(44) Varma, R.; Mayor, S. *Nature* 1998, *394*, 802.

(45) Friedrichson, T.; Kurzchalia, T., V. *Nature* 1998, *394*, 802.

(46) Harder, T.; Scheiffele, P.; Verkade, P.; Simon, K. *J. cell. Biol* 1998, *141*, 929.

(47) Pralle, A.; keller, P.; Florin, E. L.; Simon, K.; Horber, J. K. 2000, *148*, 997.

(48) Wilson, B. S.; Pfeiffer, J. R.; Oliver, J. M. *J. cell. Biol* 2000, *149*, 1131.

(49) Ahmed, S. N.; Brown, D. A.; London, E. *Biochemistry* 1997, *36*, 10944.

(50) Schroeder, R. J.; Ahmed, S. N.; Zhu, Y.; London, E.; Brown, D. A. *J. Biol. Chem* 1998, *273*, 1150.

(51) McConnell, H. M.; Radhakrishnan, A. *Biochim, Biophys. Acta* 2003, *1610*, 159.

(52) McConnell, H. M.; Vrljic, M. *Annual Review of Biophysics and Biomolecular Structure* 2003, *32*, 469.

(53) Dietrich, C.; Bagatolli, L. A.; Volovyk, Z. N.; Thompson, N. L.; Levi, M.; Jacobson, K.; Gratton, E. *Biophys. J* 2001, *80*, 1417.

(54) Samsonov, A. V.; Mikalyov, I.; Cohen, F. S. *Biophys. J* 2001, *81*, 1586.

(55) Gandhavadi, M.; Allende, D.; Vidal, A.; Simon, S. A.; McIntoch, T. J. *Biophys. J* 2002, *82*, 1469.

(56) Veatch, S. L.; Keller, S. L. *Biochim, Biophys. Acta* 2005, *1746*, 175.

(57) Katsaras, J.; Gutberlet, T. *Lipid Bilayers: Structure and Interactions*; Springer Verlag: Berlin, 2001.

(58) Devaux, P. F.; Morris, R. *Traffic* 2004, *5*, 241.

(59) Nelson, D. L.; Cox, M. M. *Lehninger: Principles of Biochemistry, 4th Ed*; W. H. Freeman and Company: New Yourk, 2005.

(60) Simon, K.; Vaz, W. L. C. *Annu. Rev. Biophys. Biomol. Struct* 2004, *33*, 269.

(61) Hancock, J. F. *Nat. Rev. Mol. Cell Biol* 2006, *7*, 456.

(62) McMullen, T. P. W.; McElhaney, R. N. *Curr. Opin. Colloid Interface Sci* 1996, *1746*, 172.

(63) Holopainen, J. M.; Mesto, A., J; Mattila, J. P.; Jutila, A.; Kinnunen, P. K. J. *Biophys. J* 2004, *88*, 1510.

(64) Terova, B.; Heczko, R.; Slotte, J. P. *Biophys. J* 2005, *88*, 2661.

(65) Fillippov, A.; Oradd, G.; Lindblom, G. *Biophys. J* 2006, *90*.

(66) de Almeida, R. F.; Fedorov, A.; Prieto, M. *Biophys. J* 2003, *85*, 2406.

(67) Tu, K.; Klein, M. L.; Tobias, D. J. *Biophys. J* 1998, *75*, 2147.

(68) Smondyrev, A. M.; Berkowitz, M. L. *Biophys. J* 1999, *77*, 2075.

(69) Rog, T.; Pasenkiewicz-Gierula, M. *Biophys. J* 2001, *81*, 2190.

(70) Chiu, S. W.; Jakobsson, E.; Mashl, R. J.; Scott, H. L. *Biophys. J* 2002, *83*, 1842.

(71) Hofsäß, C.; Lindahl, E.; Edholm, O. *Biophys. J* 2003, *84*, 2192.

(72) Pandit, S. A.; Bostick, D. L.; Berkowitz, M. L. *Biophys. J* 2004, *86*, 1345.

(73) Chiu, S. W.; Vasudevan, S.; Jakobsson, E.; Mashl, R. J.; Scott, H. L. *Biophys. J* 2003, *85*, 3624.

(74) Hyvönen, M. T.; Kovanen, P. T. *J. phys. Chem. B* 2003, *107*, 9102.

(75) Mombelli, E.; Morris, R.; Taylor, W.; Fraternali, F. *Biophys. J* 2003, *84*, 1507.

(76) Niemela, P. S.; Hyvönen, M. T.; Vattulainen, I. *Biophys. J* 2004, *87*, 2976.

(77) Khelashvili, G. A.; Scott, H. L. *J. Chem. Phys* 2004, *120*, 9841.

(78) Pandit, S. A.; Jakobsson, E.; Scott, H. L. *Biophys. J* 2004, *87*, 3312.

(79) Pandit, S. A.; Vasudevan, S.; Chiu, S. W.; Mashl, R. J.; Jacobsson, E.; Scott, H. L. *Biophys. J* 2004, *87*, 1092.

(80) Fridriksson, E. K.; Shipkova, P. A.; Sheets, E. D.; Holowka, D.; Baird, B.; McLafferty, F. W. *Biochemistry* 1999, *38*, 8056.

(81) Pandit, S. A.; Berkowitz, M. L. *Biophys. J* 2002, *82*, 1818.

(82) Pandit, S. A.; Bostick, D. L.; Berkowitz, M. L. *Biophys. J* 2003, *85*, 3120.

(83) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaige, J. G. E. *J. Comp. Chem* 1997, *18*, 1463.

(84) Nose, S.; Klein, M. L. *Mol. Phys* 1983, *50*, 1055.

(85) Parrinello, M.; Rahman, A. *J. Appl. Phys* 1981, *52*, 7182.

(86) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem* 1987, *91*, 6269.

(87) Berger, O.; Edholm, O.; Jahnig, F. *Biophys. J* 1997, *72*, 2002.

(88) van Gunsteren, W. F.; Billeter, A. A.; Eising, A. A.; Hunenberger, P. H.; Kruuger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulations: the GROMOS96 Manual and User Guide.*; Vdf Hochschulverlag AG and der ETH: Zurich, Switzerland, 1996.

(89) Tieleman, D. P.; Berendsen, H. J. C. *Biophys. J* 1998, *74*.

(90) Niemelä, P. S.; Hyvönen, M. T.; Vattulainen, I. *Biophys. J* 2006, *90*, 851.

(91) Nagle, J. F.; Tristram-Nagle, S. *Biochim, Biophys. Acta Rev. Biomembr* 2000, *1469*, 159.

(92) Huang, J.; Feigenson, G. W. *Biophys. J* 1999, *76*, 2142.

(93) Petrache, H. I.; Tristram-Nagle, S.; Gawrisch, K.; Harries, D.; Parsegian, V. A.; Nagle, J. F. *Biophys. J* 2004, *86*, 1574.

(94) Edholm, O.; Nagle, J. F. *Biophys. J* 2005, *89*, 1827.

(95) Brzustowicz, M. R.; Brunger, A. T. *J. Appl. Crystallogr* 2005, *38*, 126.

(96) Douliez, J. P.; Leonard, A.; Dufourc, E. J. *Biophys. J* 1995, *68*, 1727.

(97) Mukhopadhyay, P.; Monticelli, L.; Tieleman, D. P. *Biophys. J* 2004, *86*, 1601.

(98) Desiraju, G. R. *Acc. Chem. Res* 1991, *24*, 290.

(99) Jeffrey, G. A. *An Introduction to Hydrogen Bonding.* ; Oxford University Press: New York, 1997.

(100) Gu, Y.; Kar, T.; Scheiner, S. *J. Am. Chem. Soc* 1999, *121*, 9411.

(101) Raveendran, P.; Wallen, S. L. *J. Am. Chem. Soc* 2002, *124*, 12590.

(102) McIntoch, T. J.; Simon, S. A.; Needham, D.; Huang, C. *Biochemistry* 1992, *31*, 2020.

(103) Simons, K.; Toomre, D. *Nature Reviews Molecular Cell Biology* 2000, *1*, 31.

(104) Baird, B.; Sheets, E. D.; Holowka, D. *Biophysical Chemistry* 1999, *82*, 109.

(105) Simons, K.; Ikonen, E. *Science* 2000, *290*, 1721.

(106) Brown, D. A.; London, E. *Annual Review of Cell and Developmental Biology* 1998, *14*, 111.

(107) Brown, D. A.; London, E. *Journal of Membrane Biology* 1998, *164*, 103.

(108) Simons, K.; Vaz, W. L. C. *Annual Review of Biophysics and Biomolecular Structure* 2004, *33*, 269.

(109) Gaus, K.; Gratton, E.; Kable, E. P. W.; Jones, A. S.; Gelissen, I.; Kritharides, L.; Jessup, W. *Proceedings of the National Academy of Sciences of the United States of America* 2003, *100*, 15554.

(110) Burgos, P.; Yuan, C. B.; Viriot, M. L.; Johnston, L. J. *Langmuir* 2003, *19*, 8002.

(111) Burns, A. R. *Langmuir* 2003, *19*, 8358.

(112) Giocondi, M. C.; Milhiet, P. E.; Dosset, P.; Le Grimellec, C. *Biophysical Journal* 2004, *86*, 861.

(113) Ianoul, A.; Burgos, P.; Lu, Z.; Taylor, R. S.; Johnston, L. J. *Langmuir* 2003,

19, 9246.

(114) Kahya, N.; Scherfeld, D.; Bacia, K.; Schwille, P. *Journal of Structural Biology* 2004, *147*, 77.

(115) Nicolini, C.; Thiyagarajan, P.; Winter, R. *Physical Chemistry Chemical Physics* 2004, *6*, 5531.

(116) Subczynski, W. K.; Kusumi, A. *Biochimica Et Biophysica Acta-Biomembranes* 2003, *1610*, 231.

(117) Yuan, C. B.; Furlong, J.; Burgos, P.; Johnston, L. J. *Biophysical Journal* 2002, *82*, 2526.

(118) Zhang, Z. C.; Bhide, S. Y.; Berkowitz, M. L. *J. phys. Chem. B* 2007, *111*, 12888.

(119) Pandit, S. A.; Bostick, D.; Berkowitz, M. L. *Biophysical Journal* 2004, *86*, 1345.

(120) Niemela, P. S.; Ollila, S.; Hyvonen, M. T.; Karttunen, M.; Vattulainen, I. *Plos Computational Biology* 2007, *3*, 304.

(121) Aittoniemi, J.; Rog, T.; Niemela, P.; Pasenkiewicz-Gierula, M.; Karttunen, M.; Vattulainen, I. *The Journal of Physical Chemistry B* 2006, *110*, 25562.

(122) Ohvo-Rekila, H.; Ramstedt, B.; Leppimaki, P.; Slotte, J. P. *Progress in Lipid Research* 2002, *41*, 66.

(123) Kan, C. C.; Ruan, Z. S.; Bittman, R. *Biochemistry* 1991, *30*, 7759.

(124) Smaby, J. M.; Brockman, H. L.; Brown, R. E. *Biochemistry* 1994, *33*, 9135.

(125) Hoff, B.; Strandberg, E.; Ulrich, A. S.; Tieleman, D. P.; Posten, C. *Biophys. J* 2005, *88*, 1818.

(126) Maulik, P. R.; Shipley, G. G. *Biochemistry* 1996, *35*, 8025.

(127) Kucerka, N.; Tristram-Nagle, S.; Nagle, J. F. *Journal of Membrane Biology* 2006, *208*, 193.

(128) Pandit, S. A.; Chiu, S. W.; Jakobsson, E.; Grama, A.; Scott, H. L. *Biophys. J* 2007, *92*, 920.

(129) Klauda, J. B.; Kucerka, N.; Brooks, B. R.; Pastor, R. W.; Nagle, J. F. *Biophysical Journal* 2006, *90*, 2796.

(130) Maulik, P. R.; Sripada, P. K.; Shipley, G. G. *Biochim, Biophys. Acta* 1991, *1062*, 211.

(131) Maulik, P. R.; Shipley, G. G. *Biophys. J* 1995, *69*, 1909.

(132) Epand, R. M.; Epand, R. F. *Chem.Phys.Lipids* 2004, *132*, 37.

(133) Lindblom, G.; Orädd, G.; Fillippov, A. *Chem.Phys.Lipids* 2006, *141*, 179.

(134) Pitman, M. C.; Suits, F.; Mackerell, A. D.; Feller, S. E. *Biochemistry* 2004, *43*, 15318.

(135) Laaksonen, L. *J.Mol. Graphics* 1992, *10*, 33.

(136) Bergman, D. L.; Laaksonen, L.; Laaksonen, A. *J.Mol. Graphics* 1997, *15*, 301.

(137) Frazier, M. L.; Wright, J. R.; Pokorny, A.; Almeida, P. F. F. *Biophysical Journal* 2007, *92*, 2422.

(138) Komura, S.; Shirotori, H.; Olmsted, P. D.; Andelman, D. *Europhysics Letters* 2004, *67*, 321.

(139) Munro, S. *Cell* 2003, *115*, 377.

(140) Tsamaloukas, A.; Szadkowska, H.; Heerklotz, H. *Biophys. J* 2006, *90*, 4479.

(141) Zhancheng, Z.; Bhide, S. Y.; Berkowitz, M. L. *Journal of Physical Chemistry B* 2007, *In Press*.

(142) G. N. Patey and J. P. Valleau. *Chem, Phys. Lett* 1973, *21*.

(143) Torrie, G. M.; Valleau, J. P. *J.Comput. Phys* 1977, *23*, 187.

(144) Kim, E. B.; Lockwood, N.; Chopra, M.; Guzman, O.; Abbott, N. L.; de Pablo, J. J. *Biophysical Journal* 2005, *89*, 3141.

(145) Shankar Kumar, D. B., Robert H. Swendsen, Peter A. Kollman, John M. Rosenberg. *J. Comp. Chem* 1992, *13*, 1011.

(146) Halle, B. *Philosophical Transactions of the Royal Society B* 2004, *359*, 1207.

(147) Otting, G. *Progress in Nuclear Magnetic Resonance Spectroscopy* 1997, *31*, 259.

(148) Zanotti, J. M.; Bellissent-Funel, M. C.; Parello, J. *Biophys. J* 1999, *76*, 2390.

(149) Krishna, M. M. G. *J. Phys. Chem. A* 1999, *103*, 3589.

(150) Pal, S. K.; Sukul, D.; Mandal, D.; Bhattacharyya, K. *j. phys. Chem. B* 2000, *108*, 4529.

(151) Pal, S. K.; Zewail, A. H. *Chemical Reviews* 2004, *104*, 2099.

(152) Sykora, J.; Kapusta, P.; Fidler, V.; Hof, M. *Langmuir* 2002, *18*, 571.

(153) Sykora, J.; Jurkiewicz, P.; Epand, R. M.; Kraayenhof, R.; Langer, M.; Hof, M. *Chemistry and Physics of Lipids* 2005, *135*, 213.

(154) Russo, D.; Murarka, R. K.; Hura, G.; Verschell, E.; Copley, J. R. D.; Head-Gordon, T. *J. Phys. Chem. B* 2004, *108*, 19885.

(155) Harpham, M. R.; Ladanyi, B. M.; Levinger, N. E. *J. Chem. Phys* 2004, *121*, 7855.

(156) Zhao, W.; Moilanen, D. E.; Fenn, E. E.; Fayer, M. D. *J. Am. Chem. Soc* 2008, *130*, 13927.

(157) Jana, B.; Pal, S.; Bagchi, B. *J. Phys. Chem. B* 2008, *112*, 9112.

(158) Bhide, S. Y.; Berkowitz, M. L. *J. Chem. Phys* 2005, *123*, 224702.

(159) Faeder, J.; Ladanyi, B. M. *J. Phys. Chem. B* 2000, *104*, 1033.

(160) Faeder, J.; Ladanyi, B. M. *J. Phys. Chem. B* 2001, *105*, 11148.

(161) Laage, D.; Hynes, J. T. *Science* 2006, *311*, 832.

(162) Laage, D.; Hynes, J. T. *Chem, Phys. Lett* 2006, *433*, 80.

(163) Laage, D.; Hynes, J. T. *Proceedings of the National Academy of Sciences* 2007, *104*, 11167.

(164) Piletic, I. R.; Moilanen, D. E.; Spry, D. B.; Levinger, N. E.; Fayer, M. D. *J. Phys. Chem. A* 2006, *110*, 4985.

(165) Bhide, S. Y., Berkowitz, Max L. *J. Chem. Phys* 2006, *125*, 94713.

(166) Tan, H. S.; Piletic, I. R.; Fayer, M. D. *J. Chem. Phys* 2005, *122*, 174501.