Lecture 1

ATOMIC LEVEL SIMULATIONS OF MEMBRANES

A fundamental understanding of the properties of biological membranes and membrane proteins from the atomic point of view is undoubtedly of great biochemical, biophysical, and medical interest.[1] Knowledge of the structure and dynamical behavior of membranes and associated membrane proteins can help contribute to the development of pharmaceuticals, anesthetics and drug delivery agents, to name just three examples, as well as ultimately providing a picture of how the components of membrane systems behave and interact. So, how can such computer simulations help in this endeavor, especially when one considers the time- and length-scales of the multitude of processes associated with membranes? For example, these span many orders of magnitude, ranging as they do anywhere from femtoseconds for intramolecular lipid vibrations, to multi-nanosecond lipid molecule rotations, to minutes or longer for lipid molecule transbilayer flips.[2] Moreover, there is evidence that in membranes lipids exhibit collective behavior as "rafts", which are molecular non-covalent aggregates that can span several tens of nanometers or more on a cell membrane surface.[3] Current state-of-the-art molecular dynamics (MD) calculations, which integrate classical equations of motion using atomic level force fields, seldom utilize more than 150,000 atoms or yield dynamical trajectories spanning longer than a few tens of nanoseconds, which in the present context implies a basic simulation system with dimensions of roughly ten nanometers in any direction. This is a severe restriction if one considers the size of lipid rafts and the overall dimensions of a cell, whose outer membrane spans distances on the order of ten microns.

As implied above, computationally tractable MD simulations are currently restricted to probing properties that are local (e.g., a few nanometers) and relatively fast (e.g., up to a few nanoseconds). Fortunately, some important properties fall within this limit; they include the lipid alkyl chain isomerizations, lipid molecule rotational relaxation, and lateral self-diffusion. MD can be a useful tool to probe these properties in biomembranes at the atomic level and yield the relevant dynamical and structural information.

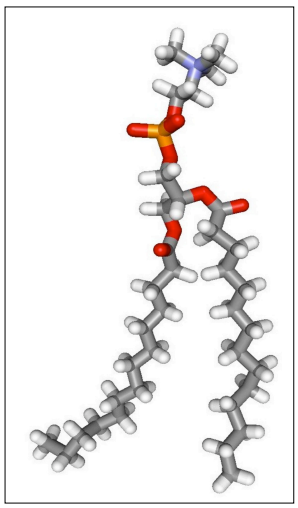

Fig. 1: Shown is an all atom representation of the phospholipids molecule dimyristolphosphatidylcholine (DMPC).

Although different lipids occur in cellular membranes (e.g., sphingolipids, glycolipids, cholesterol, etc.), those with two hydrophobic alkyl chains and a zwitterionic head groups are most abundant. The polar hydrophilic head group combines with the hydrophobic alkyl chains to provide the crucial amphiphilic character of typical phospholipids. Chain lengths of 14-24 carbons are known and chain asymmetry and unsaturated C=C bonds are also commonly present. DMPC is ubiquitous in cellular systems with fourteen carbons on each tail and a choline head (see: Figure 1). DMPC is an obvious choice as a model system for MD simulations because it is one the simplest lipid components found in cellular membranes. In addition, the pure lipid is in the physiologically relevant liquid crystalline phase at room temperature and it has been well studied experimentally.[4]

A typical way to prepare a hydrated DMPC membrane simulation is to build the MD cell by replicating a single lipid molecule 2 x 2 on a two-dimensional lattice to generate one leaflet and then adding an equivalent inverted system to produce an eight-lipid bilayer. Then water is added to "cap" the top and bottom of the bilayer; typically one uses about 25 water molecules per lipid (n_w). Since the MD cell has periodic boundary conditions, the simulation system is effectively a very tiny piece of a multi-lamellar vesicle.

The eight-lipid bilayer system is first run with MD in the (NVT) ensemble for a few picoseconds to remove any bad intermolecular contacts, after which it is again replicated 2x2 in the surface plane to yield a system of 32 lipids. The latter system was doubled in the surface plane to yield a 64 lipid bilayer, which was equilibrated for a further 100ps under NVT conditions and then for another 400ps under NPT conditions in order to relax the dimensions of the simulation cell. Once an average size for the simulation cell was established (i.e., once the simulation cell edge lengths stabilized) the system was transferred to the NVE ensemble in order to collect dynamical data. The final system was comprised of 64 DMPC molecules and 1792 water molecules, with a ratio of water molecules to lipid group (n_w) of 28. The area per head group after the simulation cell was equilibrated was 58Å^2. This method to build the simulation cell using NVT and NPT ensembles and then transfer to NVE has been shown to provide a reasonable approach to begin to probe the dynamics of membranes in the liquid crystalline phase.

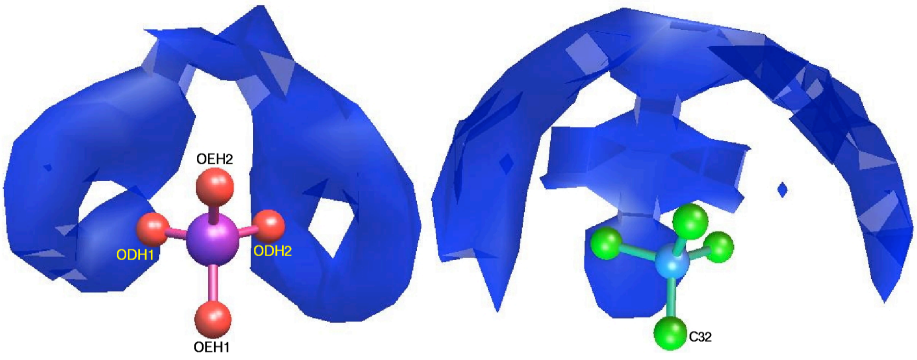

Fig. 2: Shown are the three dimensional density distributions for water oxygens around phosphate (left) and choline nitrogen (right) from a MD simulation of DMPC in its liquid crystalline phase. Dynamic properties such as hydrogen bond lifetime and decay, and diffusion of water relative to the lipid can also be determined from these types of MD simulations.

In addition to the study of the lipid dynamics in this bilayer one can also study the atomic level detail of the lipid water interface. The average number of hydrogen bonds at the lipid interface varies with the different relative position of the oxygen atom in question within the molecule. As expected, oxygen atoms at the head group will readily form one or more hydrogen bond interactions while those near the tails of the molecule will form one or none. The orientation of the interactions of the lipid headgroups was also studied and shown in Figure 2.

Lecture 2

COARSE GRAINED MODELING OF LIPID PHASES

Simulation methods in which each atom is explicitly represented are well established but have difficulty addressing many cooperative effects of current experimental and theoretical interest. There is simply too large a gap between the time and spatial scales that govern typical intramolecular events and those which are relevant for collective motions. One example is the spatial rearrangement of membrane species such as occur in the formation of a lipid raft or membrane fusion. Available simulation techniques for specific time and spatial scales are illustrated schematically in Fig. 1. These techniques take a variety of approaches to reduce the level of detail in the representation of the system under study as the time and/or length scales grow. Bridging these disparate scales is possible with multiscale modeling in which the various levels of treatment are coupled and fed back into one another.

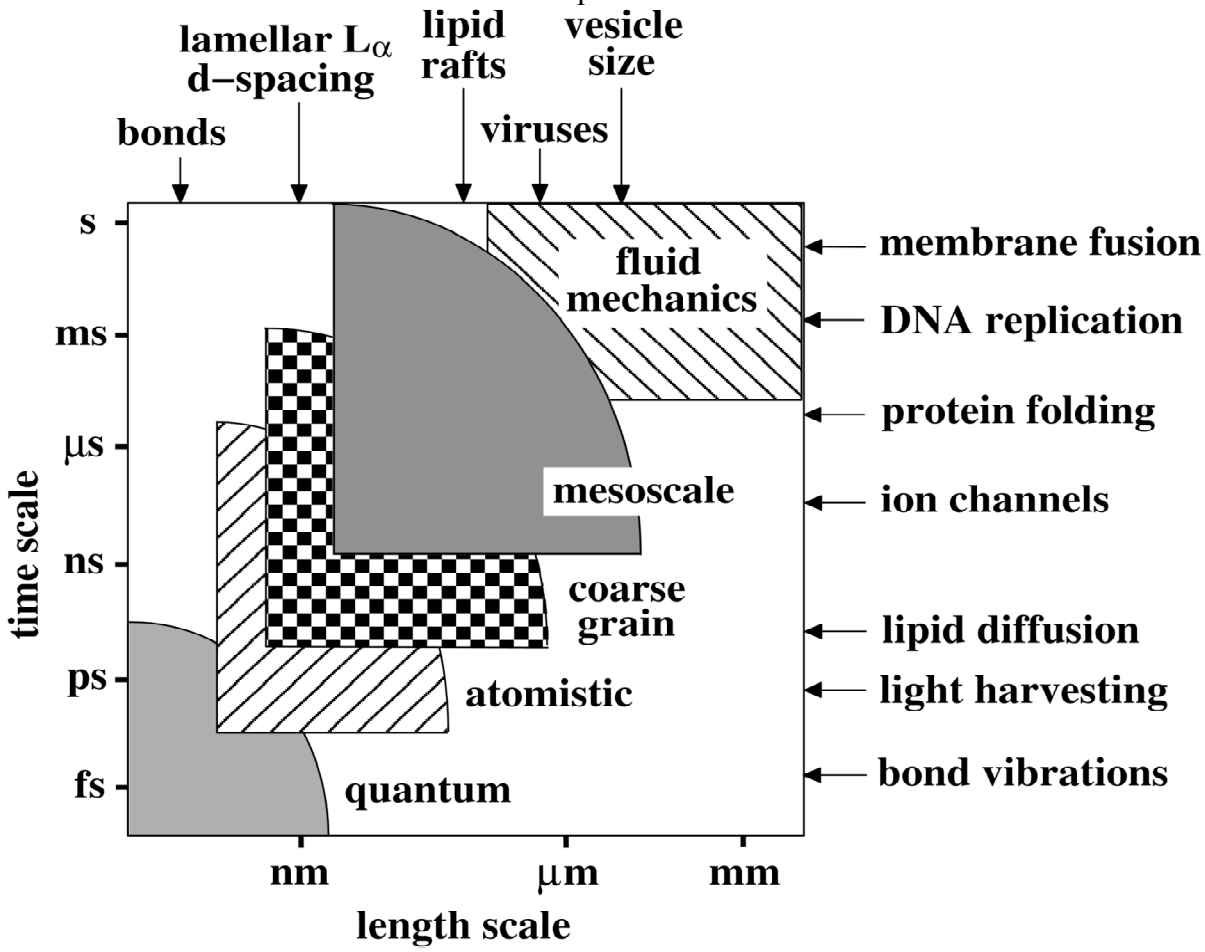

Fig. 1: Schematic of temporal and spatial scales accessible by simulation techniques. Also indicated are some characteristic membrane structures and events.

Possibly the least developed of these techniques are the ones aimed at studying events which are intermediate between the fully atomistic scale and mesoscale. Namely, events which occur on time scales of hundreds of nanoseconds to milliseconds and spatial scales that

approach a micron. Among the most powerful probes of biological systems are fluorescence based optical techniques and confocal microscopy , which routinely access precisely these time and spatial scales.

Challenges

To understand the biological function of lipids, their physical properties must be studied in the context of membranes composed of lipid/protein mixtures. Membrane lipid composition varies widely over different organelle membranes, within a single membrane, and even across leaflets of the same bilayer membrane. These differences in membrane composition range throughout the whole spectrum of living organisms from protozoans to higher organisms such as mammals. For example, the transbilayer lipid distribution is symmetrical in the endoplasmic reticulum of mammalian cells, while it is markedly asymmetrical in the plasma membrane. In the plasma membrane the majority of sphingolipids are found in the outer leaflet while most of the phosphatidylserine and phosphatidylethanolamine lipids are found in the cytosolic leaflet. Local variations in the physical properties of bilayers allow for membrane deformation and facilitate vesicle budding and fusion. Proteins can also stimulate lipid exchange between membranes by bringing them into contact. Leaflet flips can occur by the action of protein flippases, which are thought to drive vesicle budding from the plasma membrane by transferring lipids from the cytosolic to the outer leaflet. An understanding of these processes at a mesoscopic or atomic level is currently lacking[5].

It has been proposed that lipid domains of different hydrophobic thickness and composition can aid membrane protein localization, and can influence membrane protein function. In higher organisms, membrane proteins that are destined for the plasma membrane are separated from Golgi proteins based on the length of their transmembane domains. These proteins can also be chemically sorted using address labels on their cytosolic tails to interact with a protein coat. It is clear that hydrophobic matching between the protein and its matrix is essential for protein function.

CG model

It is important not to overly distort the geometrical shape of a macromolecule when doing the CG grouping of molecular fragments. As an illustrative example, for amphiphilic molecules the head group size compared to the average tail size normal to its length determines whether micelles or inverse micelles are preferred. Cone shaped molecules like phosphatidylethanolamine have small head groups and tend to form inverse micelles. Inverted cone shaped molecules like lysophosphatidylcholine tend to form micelles. Some CG models use anisotropic interaction sites to capture the underlying shape.

After selecting the interaction sites the bonded and non-bonded potentials which couple them must be determined. Short atomistic simulations can be used to attempt to include fine scale detail in a statistical manner. This is done by appealing to reverse MC simulation techniques to implicitly capture fine structure.

For aqueous amphiphilic systems we have adopted a hierarchical approach which we will now describe. This approach is quite general in nature and could easily be adapted to other systems. We begin with water. Actually, the treatment of water in the present model is somewhat complicated because water molecules are accounted for both implicitly and explicitly as will be described below. We only discuss the explicit representation here. With a Langmuir monolayer in mind we desire the CG water to be able to maintain a subcritical interface. Since we will, on average, be grouping about 10 atoms together to form CG units,

we choose to represent a loose grouping of three water molecules as a single CG site. With no internal structure and no electrostatic partial charges, we need only specify the intermolecular potential to complete the CG water model. CG interaction sites always have softer potentials than their all-atom counterparts because the constituent atomic sites become smeared out under the spherical (isotropic) averaging. We choose a Lennard-Jones 6-4 potential[7] for this reason. With two (Lennard-Jones) intermolecular parameters we can at best hope to match two target observables. Grand canonical MC is used to optimize the two parameters to match the experimental bulk density and vapor pressure of water at room temperature.

We next consider hydrocarbons. These may be branched and/or unsaturated. We usually represent three consecutive carbon atoms and their respective hydrogens as a single site. Sites are connected by harmonic bond and bend potentials. The bond and bend force constants and equilibrium values are chosen so that the bond and bend distributions best match those of the corresponding all-atom simulation. The comparison can be made by grouping the all-atom data to correspond to the CG sites.

No torsional potentials are implemented for the hydrocarbon model. Saturated straight chain hydrocarbons are quite floppy; the torsional angle distribution over four CG sites would be featureless. If the branching and unsaturation in particular cases warrants including a torsional potential this could be done. The non-bonded potential is expected to be soft and roughly the same for different hydrocarbons. We take it to be a Lennard-Jones 9-6 potential. The well depth and contact distance are adjusted so as to reproduce the experimentally determined vapor pressure and bulk density at room temperature for a few typical bulk alkane systems.

Once the pure water and pure hydrocarbon systems have been parameterized, these parameters can be used in more complicated systems without being subject to further modification. This reduces the number of free parameters in complex systems. In this spirit, we obtain one more set of parameters: the hydrocarbon-water interaction potential. Towards this end hydrocarbon-water CG systems are simulated. The interaction potentials are chosen to be of Lennard-Jones 9-6 type and the parameters are selected to obtain phase separation and reasonable width for the hydrocarbon-water interface.

Amphiphilic systems

We are now ready to parametrize an aqueous amphiphile system. We will focus on the phospholipid dimyristoylphosphatidylcholine (DMPC), but the strategy used can be applied to other systems. The DMPC molecule is coarse grained using 13 sites to represent the 118 atoms as shown in Fig. 2. An all-atom simulation of an equilibrated DMPC bilayer in the L_a phase is used to parametrize the CG model. The CG system that we will calibrate is also prepared as a bilayer so that we are treating the same thermodynamic phase. The intra-molecular force field is parametrized with harmonic functions by matching bond and bend distributions with the corresponding distributions from the all-atom simulation. The acyl tails of the DMPC lipid are straight alkane chains and the non-bonded parameters for the tail units interacting with other tail units or water are taken from the already determined hydrocarbon-water parameters.

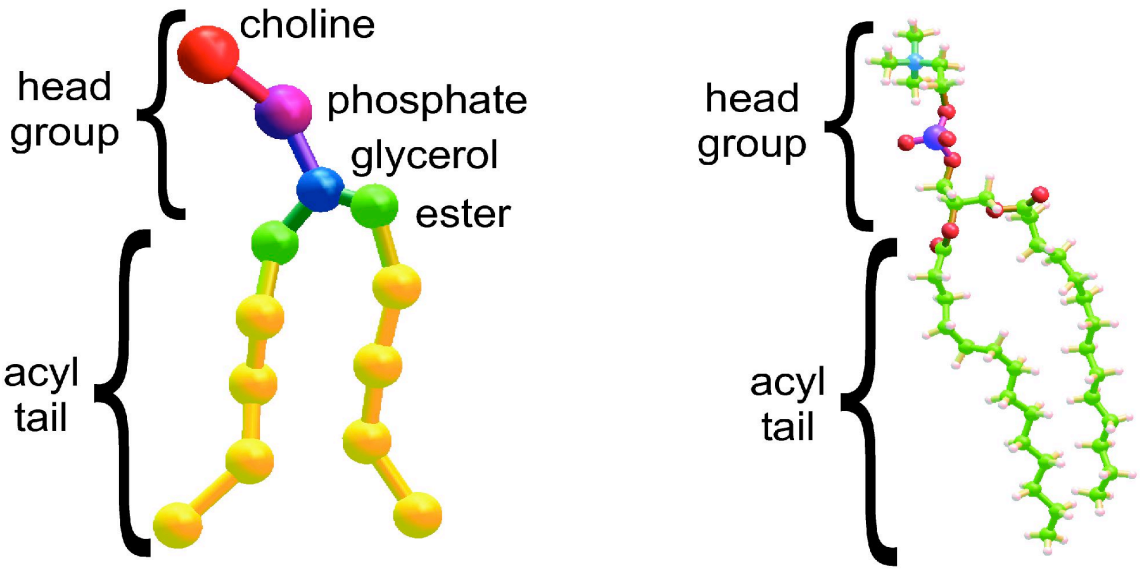

Fig. 2: Thirteen site model and all-atom version of DMPC. The choline and phosphate sites carry positive and negative electrostatic charges, respectively, of equal magnitude.

The lipid head groups are coarse grained into a positively charged choline site, a negatively charged phosphate site, a glycerol site, and two ester sites which have the two acyl tails attached to them (see Fig. 2). All combinations of non-bonded pairwise interactions among these head groups are modeled with tabulated effective potentials which aim to reproduce the radial distribution functions from the appropriately grouped all-atom simulation data. These head group -- head group interactions implicitly include water solvation shell structure. However, matters are not as straightforward as in the implicit solvent model of Lyubartsev and Laaksonen[6] because our solvent is also present explicitly. The explicit water serves as a momentum carrier (in MD simulations) and is desirable for dynamical studies. Furthermore, the Lennard-Jones nature of the potential for the explicit water site has an attractive well region which allows it to maintain an interface. Nonetheless, the parameterization strategy is unchanged: the tabulated potential attempts to mimic the all-atom reference radial distribution function in the fully interacting CG system. This means that the magnitude of the tabulated potentials are reduced compared to the case when no explicit solvent is present.

The lipid head group non-bonded interaction with both hydrocarbon and water is taken to be of Lennard-Jones form and parametrized so as to roughly match the integrated radial distribution function (in the bilayer configuration) of the two species in question out to a few selected distances. This simple functional form is chosen because these interactions are of lesser importance and we only attempt to roughly capture them.

Some of the parameterization strategy just described builds specific L_a bilayer structure directly into the force field. Furthermore the derived force field is only valid for a small temperature range. This is part of the tradeoff in moving to a more efficient simulation method; generality is sacrificed. However, the situation is not as limiting as it first appears. Only the lipid head group -- head group interaction potentials contain explicit bilayer information since these were the only potentials that were tabulated to reproduce the

thermodynamic phase-specific atomistic radial distribution function data. The rest of the interaction potentials are quite general. Furthermore, the enthalpic lipid tail and entropic changes which occur when the lipid/water system is in a different phase can partially override the structure inherent in the non-bonded potentials. Ongoing studies using this L_a bilayer derived force field include Langmuir monolayers and inverse hexagonal phases; the results (e.g., the surface tension of Langmuir monolayers) are encouraging and agree semi-quantitatively with experiments.

Efficiency: CG versus all-atom MD

The softer interaction potentials allow the use of a one order of magnitude larger propagation time-step. The reduced number of interaction sites and potentials between them yield another two orders of magnitude speedup. In the case of DMPC, the CG model consists of 13 sites and 24 internal potentials (12 bonds and 12 bends). The all-atom CHARMM DMPC encompasses 118 atoms and 971 internal potentials (117 bonds, 226 bends, 315 torsions, and 313 one-fours). A further two orders of magnitude efficiency gain comes from enhanced diffusion of the lipid species, for example in the plane of a bilayer or Langmuir monolayer. This is a result of the soft interaction potentials and the lack of an explicit hydrogen bonding network at the interface between lipid head groups and water. We have quantified this diffusional speedup as follows. The two dimensional diffusion constant for L_a phase DMPC in the plane of the bilayer is 6.5 x 10^{-8} cm^2/s for an all-atom simulation and 6.3 x10^{-6} cm^2/s for the CG model.

Membrane self-assembly

Amphiphilic self-assembly is well established for generic model systems using coarse grain simulation techniques and is the subject of recent atomistic studies[8]. We have studied many self-assembly processes from uniformly random initial conditions using the current CG model, and will present a representative selection of them to demonstrate the potential and limitations of the model.

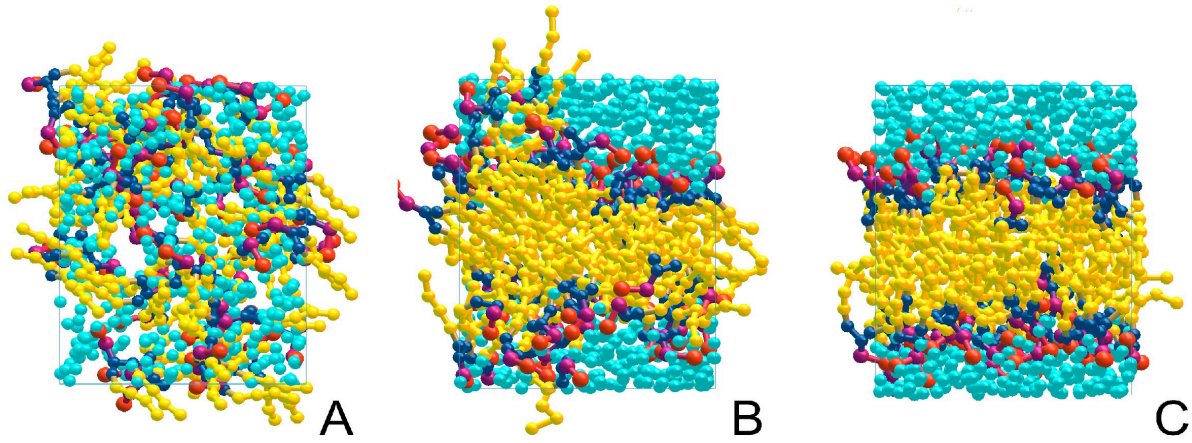

Fig. 3: Bilayer self assembly. A random initial condition (panel A) consisting of 64 DMPC lipids and 548 water sites (representing three water molecules each) assembles to its known thermodynamic state (panel C, the L_a phase) after passing through an intermediate state

(panel B) with some defects. Water is colored light blue; lipid acyl tails yellow, lipid head groups red, blue, and purple.

At 303.15 K and with a water to lipid ratio of about 25:1, the thermodynamic state of DMPC is the L_a bilayer phase. Such a system with 64 lipids in the orthorhombic simulation cell is prepared randomly as shown in Fig. 3A. From this setup, a MC simulation with a simple move set assembles into a bilayer structure with some defects as shown in Fig. 3B. Changing to MD causes the defects to heal quickly (within 6 ns) to the final state as shown in Fig. 3C.

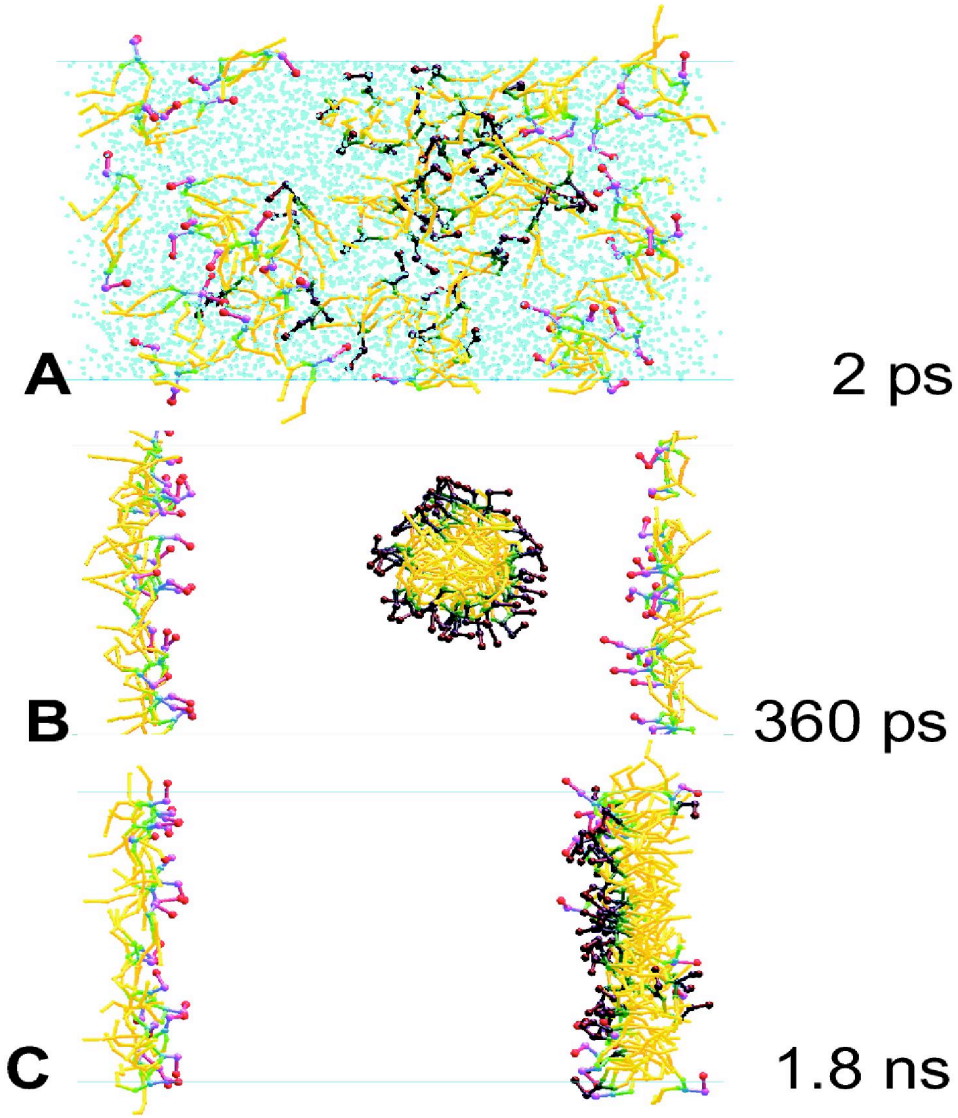

Fig. 4: Langmuir monolayer self assembly. A random slab initial condition (panel A) consisting of 80 DMPC lipids and 5000 water sites organizes into two Langmuir monolayers and a cylindrical micelle (panel B). The micelle then fuses with one of the monolayers (panel C). Water is shown in light blue and is omitted from panels B and C for clarity. The lipid acyl tails are colored yellow. The micellar lipids of panel B are shown with dark head groups in all three panels. The remaining head groups are colored red, purple, and green.

A monolayer self assembly process is studied by randomly placing lipid and water molecules in a slab geometry with two air/liquid interfaces (Fig. 4A). The system self-assembles within 300 ps into two Langmuir monolayers and a cylindrical micelle in the bulk water region (Fig.

4B). The micelle drifts towards and fuses with one of the monolayers within 1.5 ns, giving a final configuration of two unequally populated monolayers (Fig. 4C).

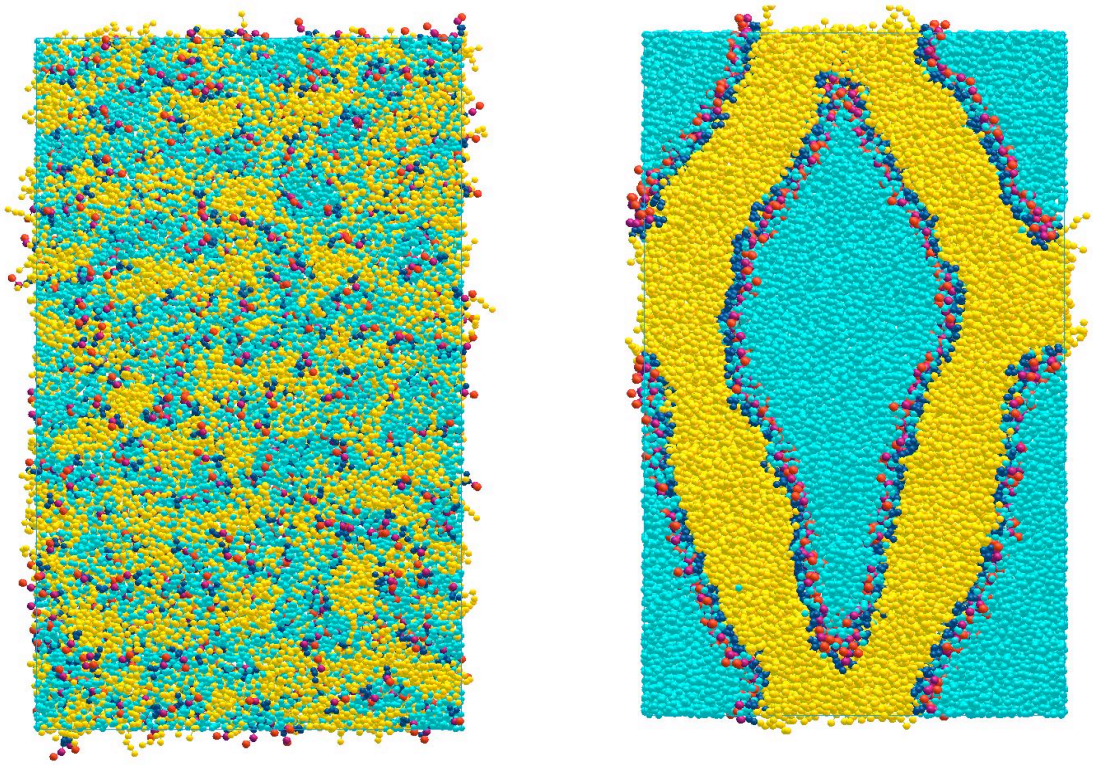

Fig. 5: Inverse hexagonal self assembly. A random initial condition (first panel) consisting of 738 DHPC lipids, 1968 nonane molecules, and 15060 water sites assembles into an inverted hexagonal phase (second panel). Water is colored light blue; lipid acyl tail and nonane yellow; lipid head groups red, blue, and purple.

An inverse hexagonal phase was self-assembled from a random initial condition consisting of 50.5 weight percent water, 33.8 weight percent diheptadecanoylphosphatidylcholine, and 15.7 weight percent nonane, as depicted in Fig. 5. This composition is experimentally known to form the inverse hexagonal phase.

However, the periodic boundary conditions and small simulation cell size may result in the stabilization of metastable structures. Care should be taken to assert whether or not the final assembled structure is the thermodynamic state.

Future perspectives

There are many avenues along which the current CG studies can be continued. Firstly, systems can be studied whose constituent components have already been parameterized. An example of this is the study of Langmuir monolayers using existing water and phospholipid parameters. Inaccuracies which come to light in such studies can point the way to improving the model, which has been parametrized only for lamellar phases. Secondly, existing parameters can be used as building blocks for new species. Thirdly, new species can be fully parameterized from a combination of all-atom MD and experimental observables.

The appeal of the second direction is that a new species can be constructed very quickly. Moreover, the construction can be artificial in the sense that some interactions can be

deliberately excluded or modified in order to assess their impact on the system under study. The drawbacks to this partial parameterization are that the model loses its predictive power for specific molecular systems, and the interaction parameters may not even be amenable to a crude guess based on the existing force field. As an example of this last point, the effect of unsaturation in the lipid acyl tail cannot be mimicked based on the DMPC parameter set.

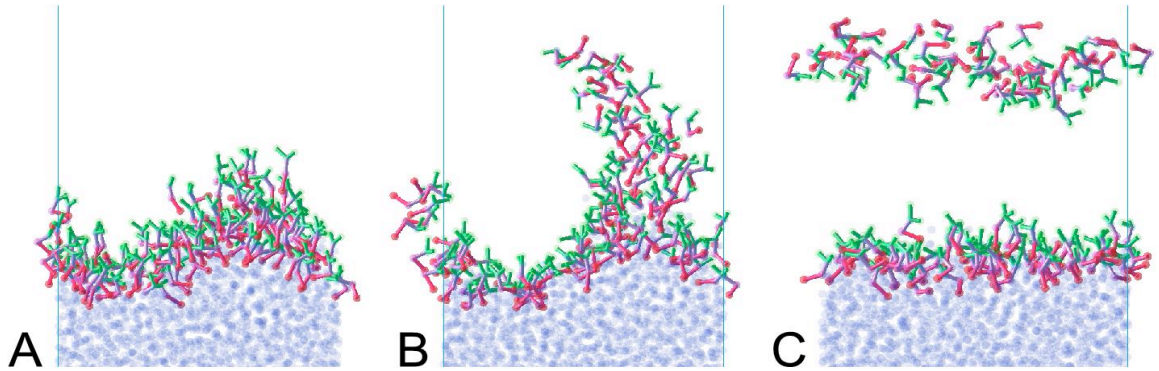

Fig. 6: Monolayer instability and collapse shown at one interface of the 120 long lipid per interface system. The initially flat interface (not shown) develops some curvature (panel A) and then opens a bridge to the exterior of the leaflet (panel B). This bridge transports enough material to eventually bring the system back into equilibrium with a flat monolayer interface (panel C). See Fig. 7for the time and the surface tension of these snapshots. Acyl tails not shown. Coloring is as follows: Water blue, choline red, phosphate purple, glycerol blue, and ester green.

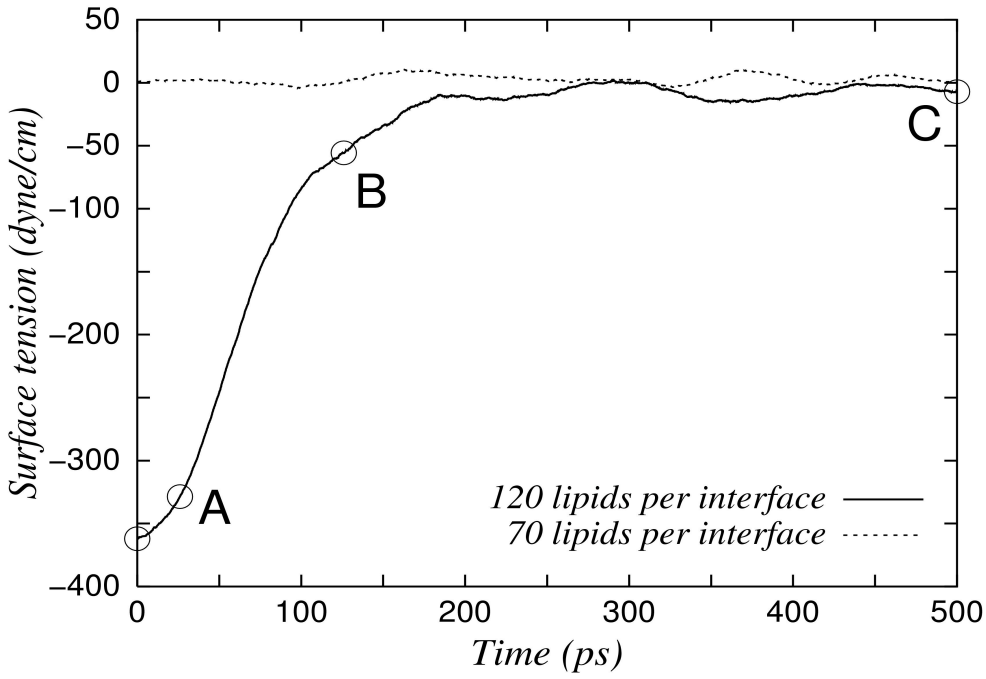

Fig. 7: Instantaneous surface tension (dyne/cm) versus raw simulation time (ps) for the 120 long-lipid per interface unstable system shown in Fig. 6. The data is smoothed with a 100 ps wide symmetrical second order Savitzky-Golay filter. Shown for comparison is the corresponding 70 long-lipid per interface curve, which has a surface tension of roughly zero (corresponding to a surface pressure of roughly 72 dyne/cm). The four circles are (from left to right): The initially flat 120 lipid per interface monolayer; Fig. 6A when the interface has

developed curvature; Fig. 6B when the monolayer is expelling lipids and Fig. 6C when the monolayer has come to equilibrium.

There are numerous topics which are amenable to study with the CG method, some of which we now mention. Bilayers and monolayers involving a few lipid species and cholesterol are suitable to study raft formation. Mixed lipid Langmuir monolayers display a rich variety of microdomains of different composition and phase. It would be interesting to see how many of these are accessible with CG models. The compression/expansion cycle of the lung surfactant DPPC could be studied in the presence of the surfactant proteins (SPs) SP-A, SP-B, and SP-C, which are known to alter monolayer collapse. Lipid mediated protein-protein interactions can be used to explore membrane protein crystallization. Oligomeric channel protein insertion into membranes and their assembly and mutual orientation across the bilayer are of interest both for antimicrobial and purely structural studies. Related to this is cyclic D,L-beta-peptide self-assembly and membrane insertion and disruption. Monolayer structure at solid/water interfaces displays novel geometry such as a hemicylindrical micelles which is being elucidated with atomic force microscopy. Entropic and enthalpic interactions between amphiphile surfaces such as micelles, lipid bilayers, microemulsion droplets, and combinations thereof can be computed as potentials of mean force. Protein alignment can be studied as a function of surface pressure in Langmuir monolayers. Self-assembled vesicles from non-lipid species such as diblock copolymers and surfactant-like peptides offer alternatives for many applications including targeted drug delivery. In conclusion, there are many fruitful future applications of CG models of the type described herein..

REFERENCES

1 Mouritsen O.G., Jorgensen K.: Small-scale lipid-membrane structure: Simulation versus experiment. Curr. Opinion Struct. Biol. 7 pp 518-527 (1997)

2 Blume A. Lipid Dynamics. In Phospholipids Handbook editor Gregor Cevc.(1993)

3 Schutz, G. J., Kada, G., Pastushenko, V. Ph., Schindler, H. Properties of lipid microdomains in a muscle cell membrane visualized by single molecule microscopy. EMBO 19, pp 892 (2000)

4 See for example: Douliez, J. P., Leonard, A., Dufourc E. J., Conformational order of DMPC sn-1 versus sn-2 chains and membrane thickness: An approach to molecular protrusion by solid state H^2-NMR and neutron diffraction. J. Phys. Chem. 100 pp 18450-18457 (1996) and references therein.

5 H. Sprong, P. van der Sluijs, and G. van Meer, How proteins move lipids and lipids move proteins, Nature Rev. Mol. Cell Biol. **2** pp 504-513 (2001).

6 A. P. Lyubartsev and A. Laaksonen, Calculation of effective interaction potentials from radial distribution functions: a reverse Monte Carlo approach, Phys. Rev. E. **52** pp 3730-3737 (1995).

7 J. C. Shelley, M. Y. Shelley, R. C. Reeder, S. Bandyopadhyay, and M. L. Klein, A coarse grain model for phospholipid simulations, J. Phys. Chem. B **105** pp 4464-4470 (2001).

8 S. J. Marrink, D. P. Tieleman, and A. E. Mark, J. Phys. Chem. B **104** pp 12165 (2000).