Development of a coarse-grained lipid model, LIME 2.0, for DSPE using multistate iterative Boltzmann inversion and discontinuous molecular dynamics simulations

Kye Won Wang, Yiming Wang, and Carol K. Hall*

Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, North Carolina, 27695

Corresponding author: hall@ncsu.edu

Abstract

We suggest an improved version of the intermediate resolution implicit solvent model for lipids, LIME, that was previously developed for use with discontinuous molecular dynamics (DMD) simulations. LIME gets its geometrical and the energy parameters between bonded and nonbonded pairs of coarse-grained (CG) sites from atomistic simulations. The improved model, LIME 2.0, uses multiple square wells rather than the single square well used in original LIME to obtain intermolecular interactions that more faithfully mimic those from atomistic simulations. The multi-state iterative Boltzmann inversion (MS-IBI) scheme is used to determine the interaction parameters. This means that a single set of interaction parameters between coarse-grained sites can be used to represent the lipid bilayers at different temperatures. The physical properties of CG DSPE lipid bilayer are calculated using CG simulations and compared to atomistic simulations results to verify the improved model. The phase transition temperature of the lipid bilayer is measured accurately and the lipid translocation phenomenon, " flip-flop" is observed through CG simulation. These results suggest that CG parameterization using multiple square-well and the MS-IBI scheme is well suited to the study of lipid bilayers cross a range of temperatures with DMD simulations.

Keyword

Coarse-graining, Multi-state iterative Boltzmann inversion, multiple square well, discontinuous molecular dynamics simulation, lipid bilayer

Introduction

Lipid bilayers result from the self-assembly of two phospholipid leaflets[1]. When dispersed in water, lipid molecules spontaneously join together to form a lipid bilayer, hiding their hydrophobic tails from the surrounding water and exposing their hydrophilic heads[1-2] in structures stabilized by intermolecular van der Waals interactions[3-4]. Lipid bilayers are a common component of cell membranes and are essential in the lives of living organisms[5-6]. They play important roles in many biological systems such as cell interior protectors and cell signaling [7-11]. Lipid bilayers can form closed spherical membrane structures called liposomes[12-13], that are very similar to animal and plant cells[14-15]. For this reason they are often used in studies of cell membrane function[16-17]. The ease of adding functional groups to the constituent lipid molecules makes liposomes attractive for applications such as targeted drug delivery, long-circulating *in-vivo* particles, and cellular-uptake enhancers[18-20].

Liposomes have recently attracted attention as drug delivery vehicles for use in treating a variety of diseases, especially cancer[21-23]. Many drugs, including DNA-based drugs, are not able to enter their target cell when administered directly into the bloodstream. An alternative is to internally store or electrostatically bind drugs to liposomes and then transport them directly to cells[24-26]. Hydrophilic drugs can be encapsulated in the aqueous phase of the liposome core and hydrophobic materials can be entrapped within the hydrophobic bilayer because of the lipids' amphiphilic nature[27]. Fetterly et al. recently developed liposomes that incorporate paclitaxel, a well-known drug for cancer treatment[28]. Similarly, Felgner and colleagues first reported that cationic liposomes could be useful for gene transfer; they have been studying liposome-based gene therapy since 1987[29].

Various simulation methods have been developed to model the behavior of lipid bilayers or liposomes. These range from atomistic (high-resolution) simulations to coarse-grained (low-resolution) simulations. Atomistic simulations of lipid bilayers have been conducted to study the structure of lipid bilayers, the permeation of ions or molecules through a lipid membrane, and the interaction between lipid bilayers and various materials[30-31]. A limitation of atomistic simulations is that their detailed description of a molecular geometry and energetics makes them computationally intensive. Simulation efficiency can be improved by coarse-grained (CG) modeling in which groups of atoms are combined together into a single interaction unit[32-33]. Coarse-grained models are attractive for studying lipid bilayers because they can be used to

examine large systems of molecules at very long time scales. The most widely used coarse-grained model for lipids is the Martini model developed by Marrink[34-35]. The Martini model is able to reproduce lipid bilayer structure, dynamic properties, thermodynamic properties, and so on in a short time frame.

Another CG model for lipids is "LIME (Lipid Intermediate Resolution Model)" designed for use with discontinuous molecular dynamics (DMD) simulations[36-37]. LIME is an implicit solvent lipid model developed by Curtis and Hall to predict the assembly of 1,2-dipalmitoyl-snglycero-3-phosphocholine (DPPC). The site-site potentials in LIME are all discontinuous functions of the separation between the CG sites. The model parameters in LIME were obtained by using a multiscale modeling procedure: the radial distribution functions (RDF) between bonded and nonbonded pairs of CG sites along the phospholipid chain were taken from atomistic simulations and then used to extract the LIME geometrical and the energetic simulation parameters. The extracted energy parameters are "potentials of mean force", effective potentials (van der Waals interactions plus electrostatic interactions) between the groups of sites in water. DMD simulations of a system of 256 DPPC molecules using the LIME model led to the formation of lipid bilayers in short computation times. LIME predicted physical properties of lipid bilayers that were comparable to those predicted by the atomistic simulations. However, LIME has a few limitations: 1) the RDFs of the lipid bilayer calculated from the atomistic simulation do not match the RDFs from the coarse-grained simulation in detail, 2) the intermolecular attractions between coarse-grained sites is limited to a single square well potential, and 3) the set of coarse-grained parameters extracted from atomistic simulations at one temperature were not necessarily accurate at other temperatures.

In order to overcome these limitations and obtain a more accurate CG lipid model for LIME, we use multiple square-well intermolecular potentials with iterative Boltzmann inversion (IBI) to find the interaction parameters. Models in which multiple square wells have been used in conjunction with DMD simulations are the following. Rutkowski et al. analyzed the phase behavior of dipolar colloidal rods by expressing the potential between the colloidal particles in terms of three square wells[38]. Benner et al. developed a CG model for chitosan using multiple square wells; the reproduced CG RDFs matched the RDFs from atomistic simulations almost exactly[39]. IBI derives coarse-grained potentials by optimizing a potential to match target RDFs from an atomistic simulation[40]. Recently, Moore et al. developed an extension to the IBI method

to include target RDFs from multiple states, adding constraints to the potential optimization process[41]. By adding these constraints, they succeeded in extracting a single set of universal interaction parameters applicable to multiple temperatures in their study.

In this paper, we develop a more accurate set of LIME parameters by applying multiple square well intermolecular potentials and using multi-state iterative Boltzmann inversion (MS-IBI) method to calculate the CG potentials. The improved LIME interaction parameters allow us to conduct CG/DMD simulations of lipid bilayers for drug delivery applications at a variety of temperatures. The model lipid for this study is 1,2-distearoyl-sn-glycero-3-phosphoethanolamine (DSPE). Unlike the previously-derived LIME force field, the improved LIME uses multiple square wells to more faithfully represent the shape of RDFs in atomistic simulations. In addition, MS-IBI is used to determine the interaction energies between CG sites, thereby expanding the ability of LIME to model lipid bilayers at any temperature. We show that the discontinuous molecular dynamics simulations using the improved LIME force field accurately represent the structure of a DSPE lipid bilayer over a given temperature range.

Highlights of our results are the following: Different numbers of square wells are needed for each of the 6 distinct CG types on DSPE. Each intermolecular potential is expressed using 3 or 4 square wells depending on the shape of the atomistic RDFs. Then the IBI method is extended to conduct multi-state optimization. The intermolecular energies are adjusted to achieve simultaneous convergence of CG RDFs from multiple states to RDFs from atomistic simulations. The values of the energies converged after 66 iterations. CG simulations of DSPE lipid bilayer using the potentials from MS-IBI are conducted at 310, 340, and 360 K. The characteristics of the lipid bilayer are well predicted at each temperature. The physical properties of the CG lipid bilayer are measured at various temperatures, and they are comparable to that of the atomistic lipid bilayer. Finally, the gel-liquid crystal phase transition temperature is predicted from CG/DMD simulations.

Model and method

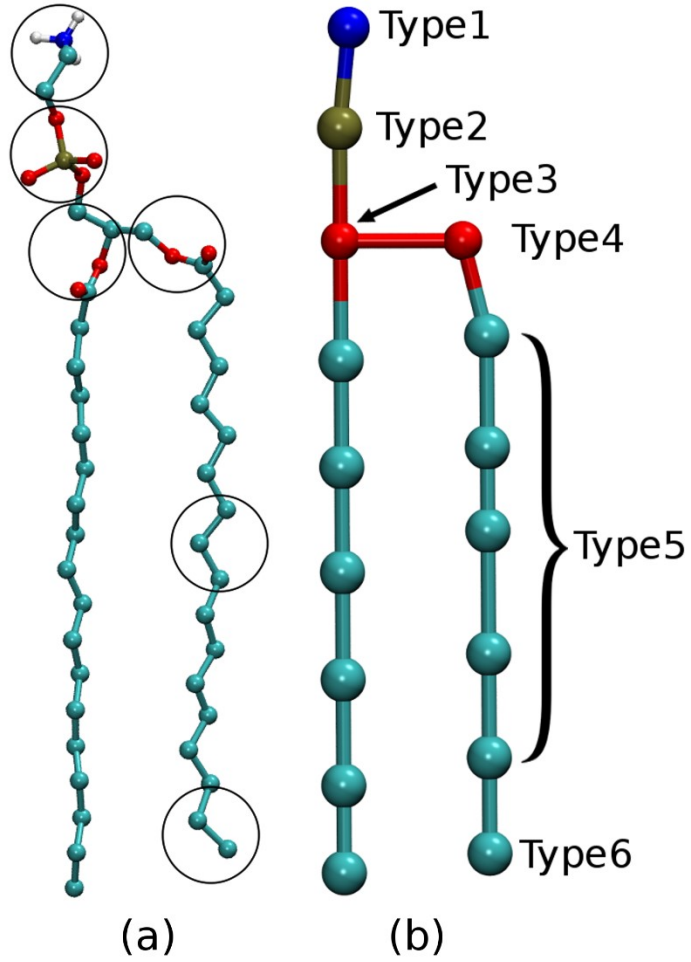

Figure 1. (a) United-atom and (b) coarse-grained representation of DSPE. The color scheme is: blue (ethanolamine); brown (phosphate group); red (ester linkage); cyan (alkyl tail group).

The representation of a lipid molecule in the improved LIME model is the same as that used in the original LIME model. Six coarse grained types are used to represent the 190 atoms that compose a DSPE molecule. Figure 1 shows the united-atom and coarse grained representations of a DSPE molecule. The coarse graining scheme massively reduces the number of sites, so that the DSPE molecule finally has 16 coarse-grained types. The coarse-grained DSPE is composed of a polar head group which includes an ethanolamine (blue), phosphate group (brown), two ester linkages (red), nonpolar hydrophobic alkyl group and a terminal tail

group (cyan). The mass of each coarse-grained site is the sum of the atoms' molar mass within each coarse grained site. Each CG site has a unique interaction potential that is different than those of the other sites.

The CG parameters for the improved LIME model are extracted from the results of explicit-solvent NPT ensemble united-atom simulation. The simulation contains 128 DSPE lipids and 5888 water molecules. We used Gromacs simulation package (version 4.5.4) with the GROMOS96 53a6 force field[42]. The Lennard-Jones interaction parameters for DSPE lipid are taken from the simulation results of Qin et al.[43]. The restrained electrostatic potential (RESP) approach with Gaussian09 was used to derive the partial charges for DSPE. The initial configuration of the system was a pre-formed bilayer. The Berendsen thermostat was used to maintain the system temperature throughout the simulation with a time constant of 0.1 ps. Atomistic simulations were performed at 310, 340, and 360 K as reference systems for the MS-IBI process. At 310 and 360 K, the lipid bilayer is considered to be in the gel (or solid) phase and in the liquid phase, respectively; 340 K is the temperature at which the transition between the two phases occurs. The simulations were run for 100 ns with a time step of 0.001 ps. The pressure was kept at 1.0 bar. The trajectories of all atoms were collected for the final 40 ns with a time step of 0.001 ps and used to calculate the centers of mass for each CG site to convert the atomistic coordinates to a CG representation. The trajectory was saved every 4000 steps, thereby a total of 10000 frames was used for the calculation. The physical properties of the atomistic lipid bilayer model were measured using the converted trajectories as well.

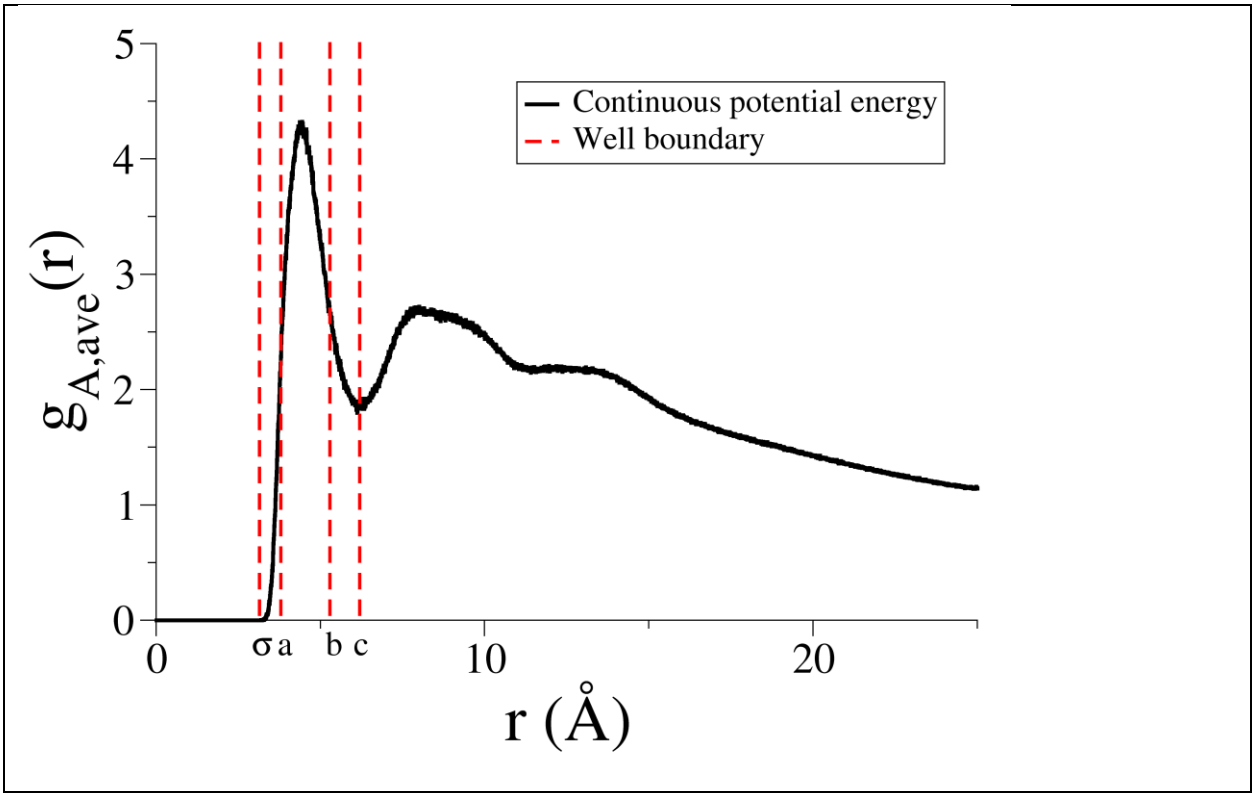

Figure 2. Radial distribution function from atomistic simulation (black) and square well boundaries marked with red dashed lines (a, b, and c indicate the first, second, and third well width in order) describes the intermolecular interaction between CG types 1 and 4.

The intermolecular interactions between the CG sites on DSPE are represented using hard spheres and multiple square wells (or shoulders). The hard sphere diameter, σ, and square-well widths, λ, are determined from averages of the RDFs between pairs of nonbonded coarse-grained sites obtained at three different temperatures (310, 340, and 360 K) in the atomistic simulations, $g_{A,ave}(r)$, where A and ave stand for atomistic simulation and average value, respectively. Figure 2 shows the average intermolecular radial distribution function of the three different temperatures from atomistic simulation (black) for the pairwise interaction between CG types 1 and 4. In general, hard sphere diameters for each pair of interaction sites were chosen by locating the minimum nonzero separation distance between the two sites. Since our goal was to obtain one parameter set that is applicable to various temperatures, we calculated nonzero distances at three different temperatures and selected the smallest value as the hard sphere diameter for that pair of interacting sites (σ in Figure 2).

The number of square wells or shoulders for each CG pair is set so that the RDFs of the CG simulations can reproduce the RDFs of the atomistic simulations and this number depends on the shape of the RDFs. Although there may be multiple peaks in the RDF graph, we focused only on the first peak as this approach appeared to be sufficient to describe the intermolecular interaction empirically. Figure 2 illustrates the well boundaries marked with red dashed lines for the first peak of the RDF between types 1 and 4. The first well width is selected to cover the range from the hard-sphere diameter to the smallest value of r at which $g_{A,ave}(r)$ is approximately 30 - 70% of the maximum height of the first peak ("a" in Figure 2). The second well width is selected to cover the range from the end of the first well to the smallest distance (larger than the maximum peak distance) at which $g_{A,ave}(r)$ is approximately 30 - 70% of the maximum value of the first peak ("b" in Figure 2). The third well width is selected to be the location at which $g_{A,ave}(r)$ is the first local minimum ("c" in Figure 2). Although the criterion for deciding the first and second well widths is somewhat ambiguous (the locations are at 30 -70% of the maximum peak height), this is because the shapes of peaks are not perfectly symmetrical. We found that abnormally narrow or wide well widths could be set if a single criteria were applied. Thus, slightly different criteria are applied depending on the shape of the peak. All intermolecular pairs of CG site have 3 separate square well potentials except for the interaction between types 1 and 3. Because the average atomistic RDF between types 1 and 3 has a second peak very near to the first peak, one more square well was used to obtain a more detailed representation. All the shapes of atomistic RDFs are shown in Supplementary Material. If the atomistic RDFs do not have noticeable local maximum, the interaction between the CG pair is represented using the hard-sphere potential. The determined hard sphere diameter and square well widths for all pairs of CG types are in Table S.1.

The square well depth (or shoulder height) , ε, between CG sites are determined by an IBI scheme modified for discontinuous potentials developed by Benner and Hall[39]. In the IBI method for a single temperature, the potential between the CG sites is chosen such that the RDFs of the CG simulations match the RDFs of the atomistic simulation. Those discontinuous potentials (square well depth) are updated according to

$$U_{CG}^{(i+1)}(r) = U_{CG}^{(i)}(r) - k_B T \ln \frac{g_{CG}^{(i)}(r)}{g_A(r)}, \qquad (1)$$

where U_{CG}(r) is the coarse-grained potential, i represents the current iteration, k_B is the Boltzmann constant, T is the absolute temperature, r is the separation between particles, $g_{CG}(r)$ is the CG RDF, and $g_A(r)$ is the RDF between pairs of nonbonded coarse-grained sites in the atomistic simulations. The IBI method for multiple temperatures is similar to that for a single temperature, but it has to be modified to cover N states (i.e. MS-IBI). The potentials are updated according to

$$U_{CG}^{(i+1)}(r) = U_{CG}^{(i)}(r) - \frac{1}{N}\sum_{j=1}^{N} k_B T_j \ln \frac{g_{CG,j}^{(i)}(r)}{g_{A,j}(r)}, \qquad (2)$$

where j represents the state, $g_{CG,j}(r)$ is the CG RDF for the jth state, and $g_{A,j}(r)$ is the atomistic RDF in the jth state[41].

The MS-IBI procedure has been modified to find square well (or shoulder) depths for CG/DMD simulations. An in-house DMD simulation code is used to perform all simulations. Each CG simulation has the same box size as the atomistic simulations. A pre-formed lipid bilayer consisting of 128 DSPEs is used as the initial configuration in the CG simulations. The Anderson thermostat was used to maintain the temperature constant; in this method, the velocity of a CG bead is adjusted to keep the system's Maxwell–Boltzmann velocity distribution consistent with the set temperature. The initial approximations for all CG square-well depths are set to be random numbers between -1.0 and 1.0. The square-well depths are updated every 200 million collisions. When the ratio of the average CG RDF of three states to the corresponding atomistic RDF for the same range of intermolecular distances is less than 1, the well depth is updated to have a stronger interaction. When the ratio is greater than 1, the well depth is updated to weaken the interaction. The details of the updating criteria are as follows. If the ratio of the average value of the CG RDF to the average value of atomistic RDF for a particular square well is smaller than 0.25 or greater than 1.75, the square well depth is updated by -0.10 or +0.10. If the ratio is between 0.25 and 0.75 or between 1.25 and 1.75, the well depth is adjusted by -0.05 or +0.05. If it is between 0.75 and 0.90 or between 1.10 and 1.25, the well depth is updated by -0.005 or +0.005 for careful adjustment. When the ratio of the average value of CG RDF to the average value of atomistic RDF in the corresponding region is between 0.90 and 1.10, the CG RDF is considered as "converged" and the value of the square well depth remains the same without updating. The simulations continue until all the CG RDFs converge to atomistic RDFs.

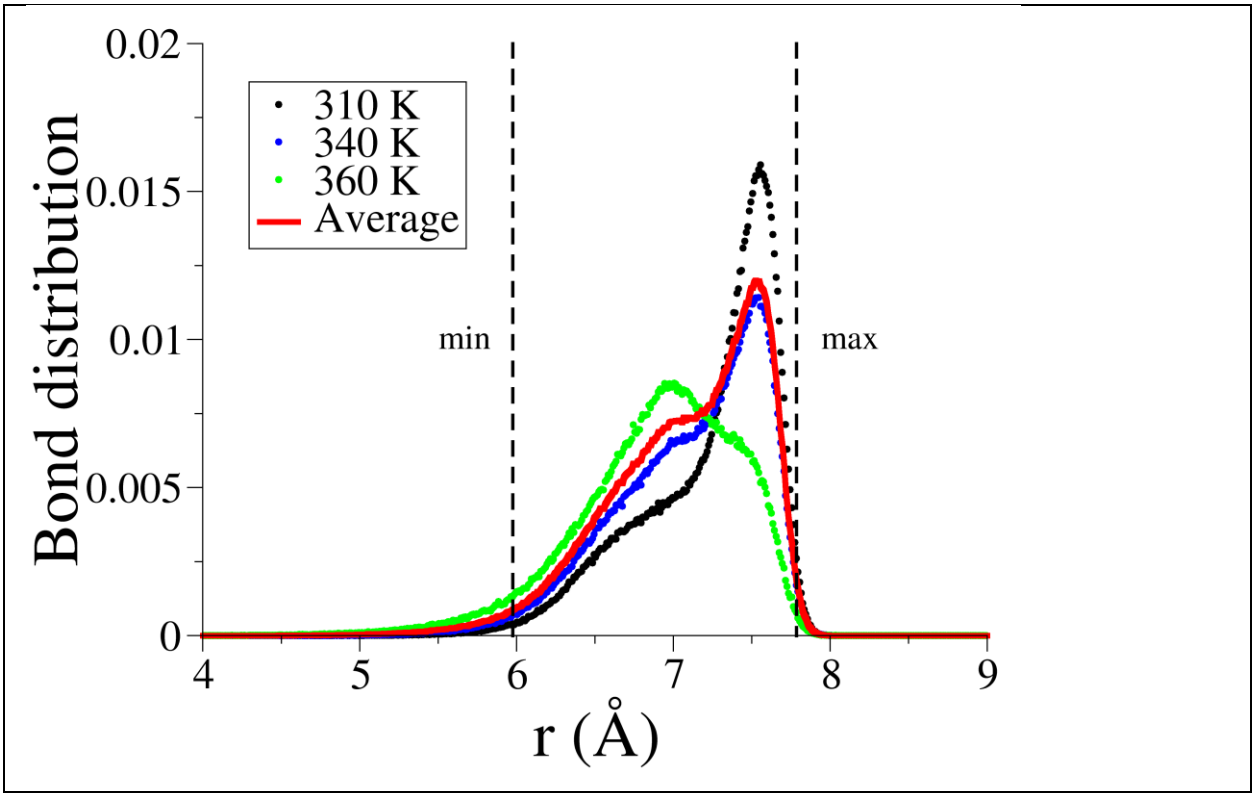

Figure 3. Intramolecular bond distribution functions for coarse-grained sites between 3 (ester group, type 3) and 5 (alkyl tail group, type 5).

To maintain the connectivity and stiffness of CG DSPE, three types of bonds are used; covalent bonds, pseudobonds to maintain bond angles, and pseudobonds to maintain torsional angles. The covalent bond indicates a real bond between adjacent DSPE CG sites. The bond angle is maintained via a pseudobond, an invisible bond between a CG site and its next nearest neighbor CG site along the DSPE chain. The torsional angle is maintained via a pseudobond between a CG and its second nearest neighbor CG site. The lengths of all real bonds and pseudobonds fluctuate between maximum and minimum values. An infinite repulsion force is exerted so that the distance between the CG sites does not fall outside of the minimum or maximum distances associated with each bond length.

A total of 45 covalent bonds and pseudobonds between CG sites along DSPE are defined. The minimum and maximum bond lengths for those bonds are calculated from atomistic simulation results for the center of mass distance between bonded CG sites. In CG/DMD simulations, distributions for covalent bonds and pseudobonds similar to those in atomistic

simulations can be achieved by limiting the minimum and maximum distance to an appropriate range. The average bond distributions at three different temperatures for bonded pairs are used to find the CG bond length as in Figure 3. The minimum bond length (label “min” in Figure 3) was selected by finding the smallest distance at which the average bond distribution function reaches 5-10% of its maximum peak value. The maximum bond length (label “max” of Figure 3) was determined by finding the smallest distance (larger than the maximum peak distance) where the average bond distribution function reaches 5-10% of its maximum peak value. The determined minimum and maximum bond lengths for all bonded CG sites are in Table S.2.

Results and discussion

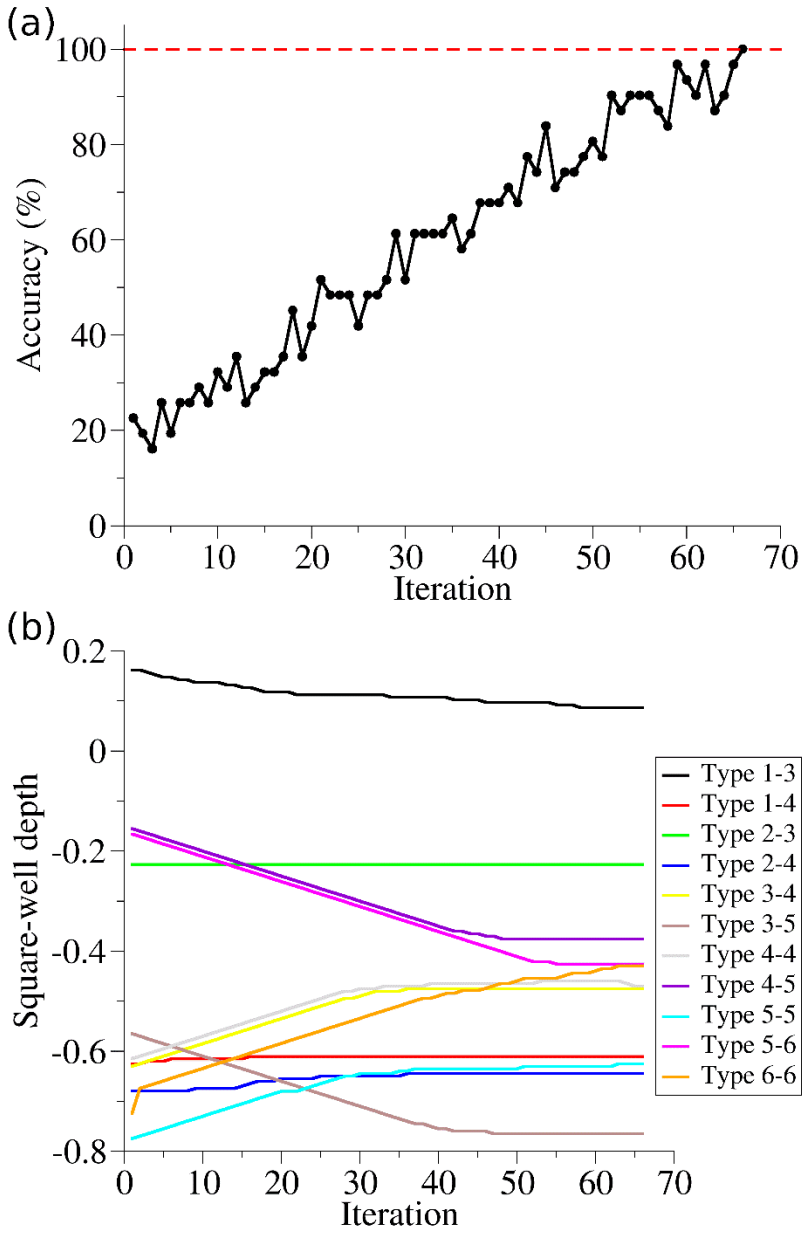

Figure 4. (a) The accuracy vs. the number of iterations and (b) the values of the second square well (or shoulder) depth as the iterations progress during MS-IBI.

The CG interaction energies were determined using the new MS-IBI approach. MS-IBI is executed to find a single set of potentials that can well represent the target structure data at three different temperatures. Figure 4a shows the accuracy of MS-IBI over the number of iterations.

The accuracy is defined as the number of CG RDFs that converge to their atomistic RDF counterparts divided by the total number of square wells considered. Note that the first square wells for all the paired interactions are not included in the accuracy calculation because it takes a very long time to converge and the CG RDFs in the first well are always comparable to the atomistic RDFs in the corresponding region regardless of the change in the square well depth. As described in the methods section, the initial values for the square well depths are assigned to be random numbers between -1.0 to 1.0, so the accuracy at the beginning is only 22.6% (Figure 4b). As the iteration progresses, the accuracy gradually increases. After 66 iterations, the percentage reaches 100%. Note that the number of iterations could be varied depending on system conditions, such as the total number of potentials, the total number of atoms, and the temperature range. Figure 4b shows the change in the values of the second square well depths for several coarse-grained types over the course of the iterations. Almost all of the values of the square well depths have reached a plateau at the end of the iteration procedure, which means that the values have evolved to their optimum values under given conditions. The final set of intermolecular CG potentials is listed in Table S.3.

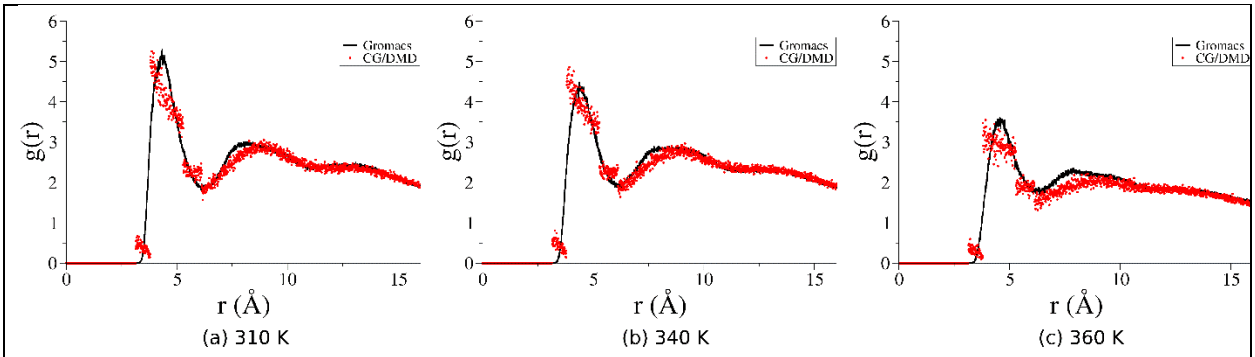

Figure 5. Comparison of atomistic (black) and CG (red) RDFs between CG types 1 and 4 at (b) 310 K, (b) 340 K, and (c) 360 K.

Figure 5 compares the atomistic and CG RDFs between CG types 1 and 4 at (a) 310 K, (b) 340 K, and (c) 360 K. The atomistic RDFs are obtained from Gromacs simulations and the CG RDFs are obtained at the end of the MS-IBI. The black line and red dots represent the RDFs of atomistic simulations and those of CG/DMD simulations, respectively. The comparison clearly reveals the characteristics of the two different simulation methods; it can be seen that the

atomistic RDFs are smoothly connected, while the CG RDFs are disconnected at the well boundaries, as expected. The change in the shape of the CG RDFs with temperature is demonstrated as is that of the atomistic RDFs. The amplitude of the first peak in the atomistic model gradually decreases as the temperature increases while the location of the peak remains nearly constant. Similarly, the height of the peak in the CG RDFs decreases as the temperature increases. Interestingly, even though our CG intermolecular potentials are set to match only the first peak of the atomistic RDFs, the CG RDFs of the remaining parts were nearly identical. The other CG RDFs are also in good agreement with the atomistic RDFs, and data on this comparison can be found in the Supplementary Material. It is apparent that the CG intermolecular potential parameters from the MS-IMI method well represent the distinctive characteristics from atomistic simulations in the various states.

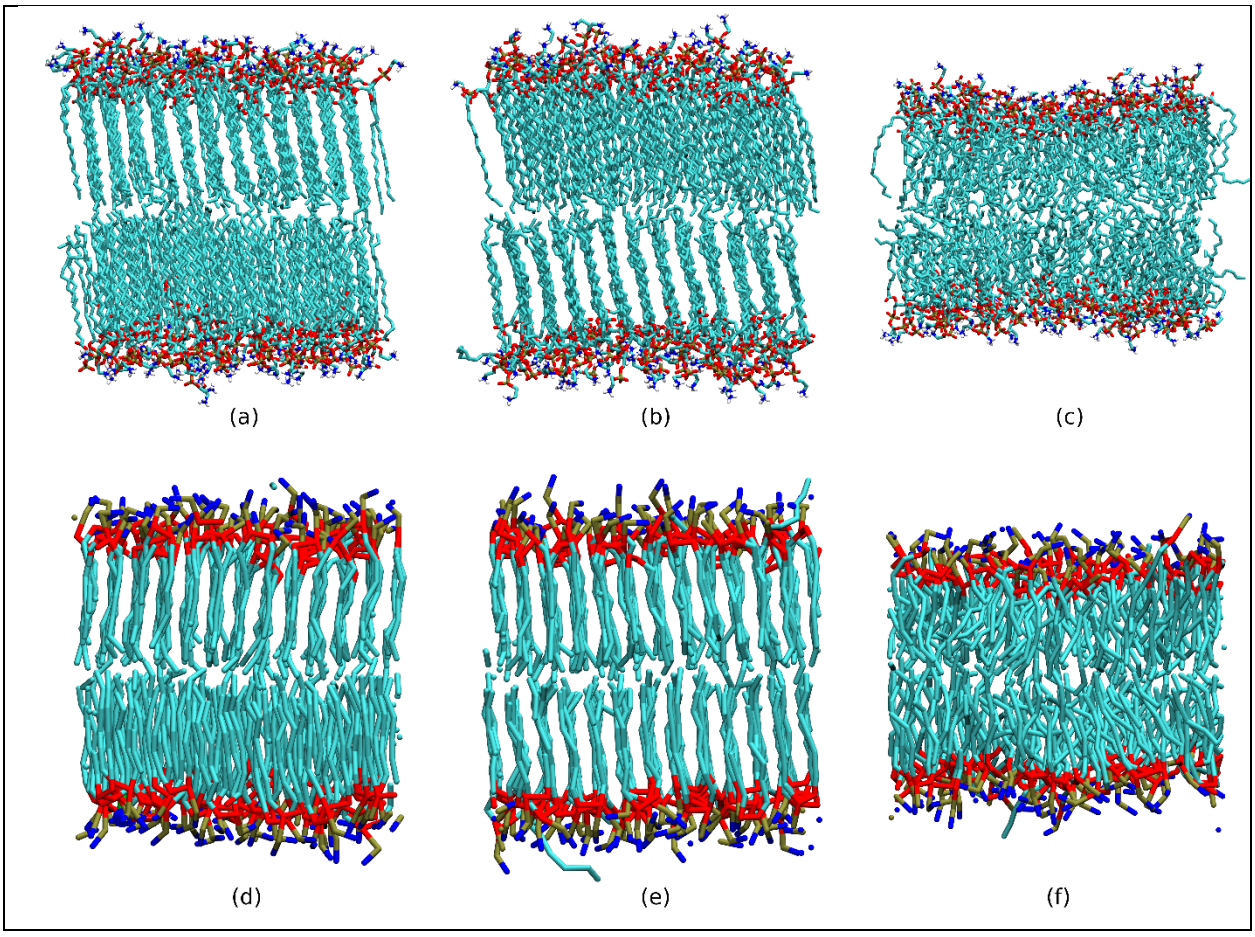

Figure 6. Snapshots of a DSPE lipid bilayer in atomistic representation at (a) T = 310 K, (b) T = 340 K, and (c) T = 360 K. Snapshots of DSPE lipid bilayer in CG representation at (d) T=310 K, (e) T=340 K, and (f) T=360 K.

The conformations of the CG lipid bilayers in three different states (310, 340, and 360 K) obtained from the MS-IBI were compared to those in the corresponding atomistic simulations. The lipid bilayers at 310, 340, and 360 K represent the lipid assembled in the gel state, the transition between the gel and liquid crystalline states, and the liquid crystalline state of DSPE lipid, respectively. Figure 6 shows snapshots of the lipid bilayer formed in atomistic and coarse-grained resolutions at the three temperatures. Blue and brown spheres indicate CG sites 1 and 2, which are ethanolamine and phosphate, respectively. Red and cyan colors represent ester linkages and hydrophobic tails. At 310K, the DSPE lipid chains in both atomistic and CG representations are well organized, this is one of the gel state configurations of lipid bilayer. In addition, a clear boundary between the upper and lower leaflets is identifiable. The snapshot at 340K shows a bilayer configuration that is very similar to that of the gel state. The tail arrangement of lipids did not change much from the arrangement at 310 K. The boundary between the upper and the lower layers is still observed, but the cohesion of the head group is weakened and the arrangement is disturbed. Finally, at 360K, the hydrocarbon tails are disordered and lose their regular arrangement. In addition, the interface between the upper and lower layers has vanished. As a result of visual observation of the model, it appears that the CG lipid bilayer models from MS-IBI mimic the atomistic models well at each temperature.

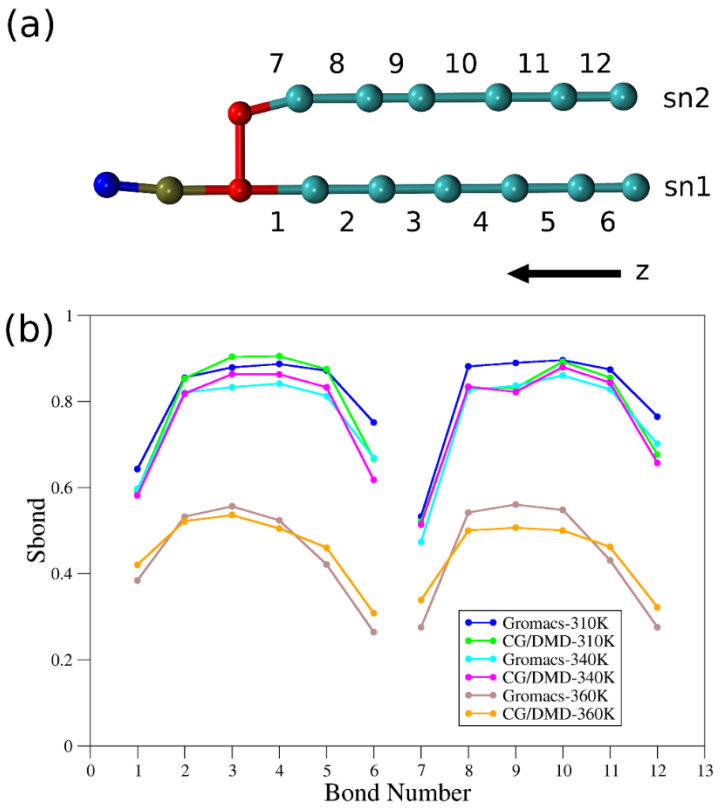

Figure 7. (a) The bond numbers and their positions. (b) Orientation order parameter of lipid tails for the different bonds at three different temperatures.

The orientation order parameter for the lipid tail is used to further investigate how closely the structure properties of the coarse-grained lipid bilayers match those of the atomistic lipid bilayer. The orientation order parameter is defined to be

$$S_{bond} = \frac{1}{2} < 3cos^2\theta - 1 >$$

where θ is the angle between the vector along a coarse-grained bond and the z-axis (bilayer normal). The value of S_{bond} is -0.5 for a perpendicular alignment and 1.0 for a parallel alignment to the bilayer normal. When bonds are in a completely disordered state, the value will be close to zero. There are 12 bonds in a CG DSPE lipid that correspond to the tails, and the bond numbers and their positions are shown in Figure 7a. The alkyl chain containing bonds 1 to 6 is called sn1 and the alkyl chain containing bonds 7 to 12 is called sn2. We compare the order parameter of atomistic and CG models at three different states (310, 340, and 360 K) and present the results in

Figure 7b. The order parameter value varied from 0.2 to 0.9 depending on the temperature: it is close to 1 at 310 or 340 K and to zero at 360 K. This result confirms that our CG lipid model can represent the vertical alignment of the lipid bilayer at lower temperatures. The orientation order parameter values from CG models and the atomistic models are fairly similar. The percent error is the difference between the CG and the atomistic order parameter values divided by the atomistic order parameter value, multiplied by 100. The average value of each bond's percent error is 5.9 %, but it varies by temperature and location. The average errors are different at different temperatures: 4.45 % at 310 K, 3.42 % at 340 K, and 9.79 % at 360 K. The error increases as the temperature increases. The average error for the sn1 chain is 4.90%, but it increases to 6.88% for the sn2 chain. The CG model replicates the atomistic model quite well, but the discrepancies at high temperature and for the sn2 chain cause the error to increase.

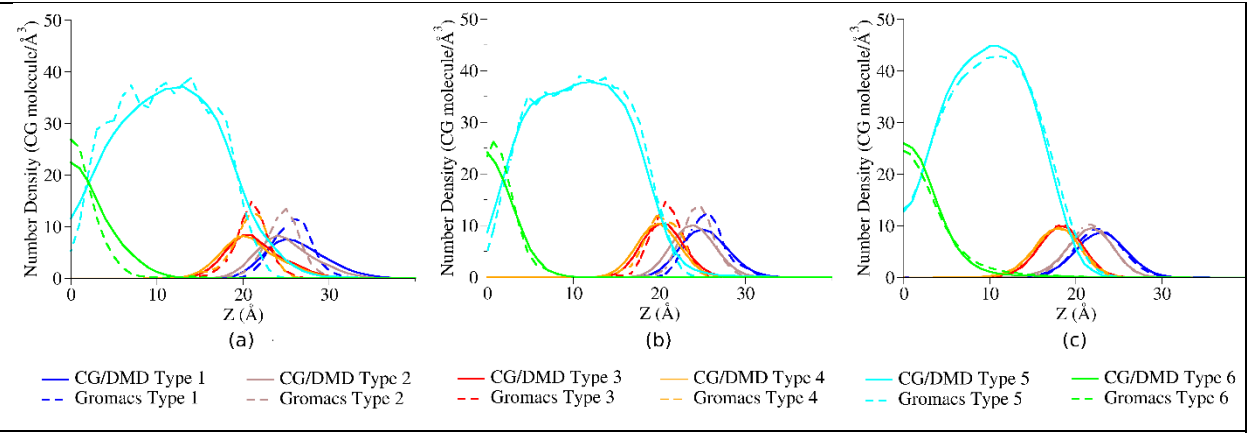

Figure 8. The number density profiles for the CG types between the CG and Gromacs simulations at (a) 310, (b) 340, and (c) 360 K.

A comparison of the number density profiles for the CG types between the CG and Gromacs simulations is shown in Figure 8. The number density is defined as the number of average atoms per unit volume at a distance (z) from the bilayer normal. The number density profile describes the probability that an atom (CG bead) of a specific type will be present at a specific location. It is useful to map out the structure of lipid bilayers by mapping out the probability that each atom will be at a given position. The density profiles for the atomistic simulations are obtained by coarse-graining the Gromacs simulation trajectories and that for the CG simulations are obtained by running DMD simulations for 100 million collisions. The

density profiles of all coarse-grained types are compared at three different temperatures. At 360 K, the height of the peaks of CG density profiles are almost perfectly matched to that of the atomistic profiles, but at the other temperatures, the heights are slightly underestimated. However, it is important to note that the peak locations of the density profiles differ by no more than 2 Å from those of atomistic density profiles at all temperatures. In general, the CG density profiles align well with atomistic density profiles, which tells us that our CG parameters are suitable to represent the position of the atoms in the lipid bilayers.

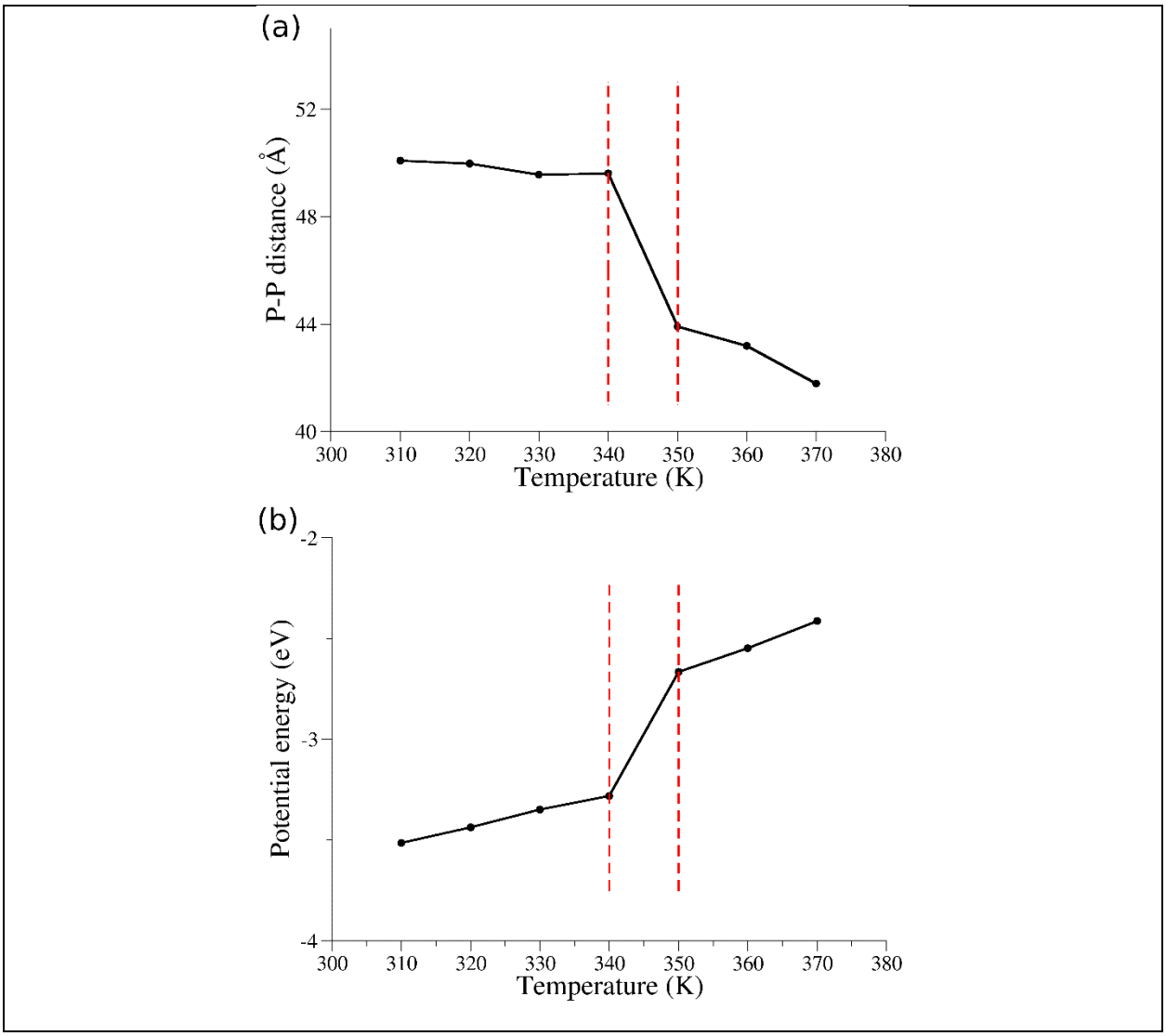

Figure 9. (a) Change in the bilayer thickness of CG DSPE with temperature. (b) Average potential energy per bead for CG type 5 at different temperatures. Red dashed lines represent the region at which the phase transition occurs.

We simulated the DSPE lipid bilayers at various temperatures other than 310, 340, and 360 K using the parameters obtained by MD-IBI and used the simulation data to estimate the phase transition temperature. In order to predict the phase transition temperature, the bilayer thickness and the average intermolecular potential energy per bead for CG type 5 are computed. The simulations are conducted at 310, 320, 330, 340, 350, 360, and 370 K for 1 billion collisions, and the trajectory of the last 100 million collisions is used to generate the necessary data. The bilayer thickness is measured at each temperature. The bilayer thickness is defined to be the mean distance between phosphate atoms in the upper and lower layers. In the case of atomistic simulations, the thickness is directly determined according to the definition. However, in CG simulation, the thickness is measured using the locations of the CG sites containing phosphate. Figure 9a shows the DSPE lipid bilayer thickness in the CG simulations using parameters obtained from MS-IBI. As shown in the figure, the lipid bilayer thickness is high at temperatures between 310 to 340 K, reflecting the characteristic thickness of the lipid bilayer in the gel state. The thickness decreases rapidly at 340K, and then decreases a little more slowly.

To help us quantify the temperature of the phase transition, the intermolecular potential energy per bead at various temperatures (Figure 9b) associated with the CG beads of type 5, which is the alkyl chain CG bead, was computed. Type 5 (the tail of lipid) was chosen for this as it plays a significant role in the phase transition. (Even though the RDF is a way to see the structural change of the lipid bilayer, it is hard to understand the phase transition simply by examining the RDFs.) The average intermolecular potential energy per bead between the beads of CG type 5 is lowest at 310 K, which means that a large number of beads are located around a central bead. As the temperature increases, the energy increases until 340 K, at which point the slope of the curve significantly increases. After 350 K, the slope is not as large. The bilayer thickness and potential energy per bead change drastically in the region of 340-350 K, allowing us to predict that the phase transition occurs in this temperature range. Actually, the experimentally measured phase transition temperature for DSPE lipids is 347 K[44]. It is apparent

that the set of intermolecular potential parameters obtained through MS-IBI is well suited for CG simulations at various temperatures, and that it is able to predict the known phase transition temperature.

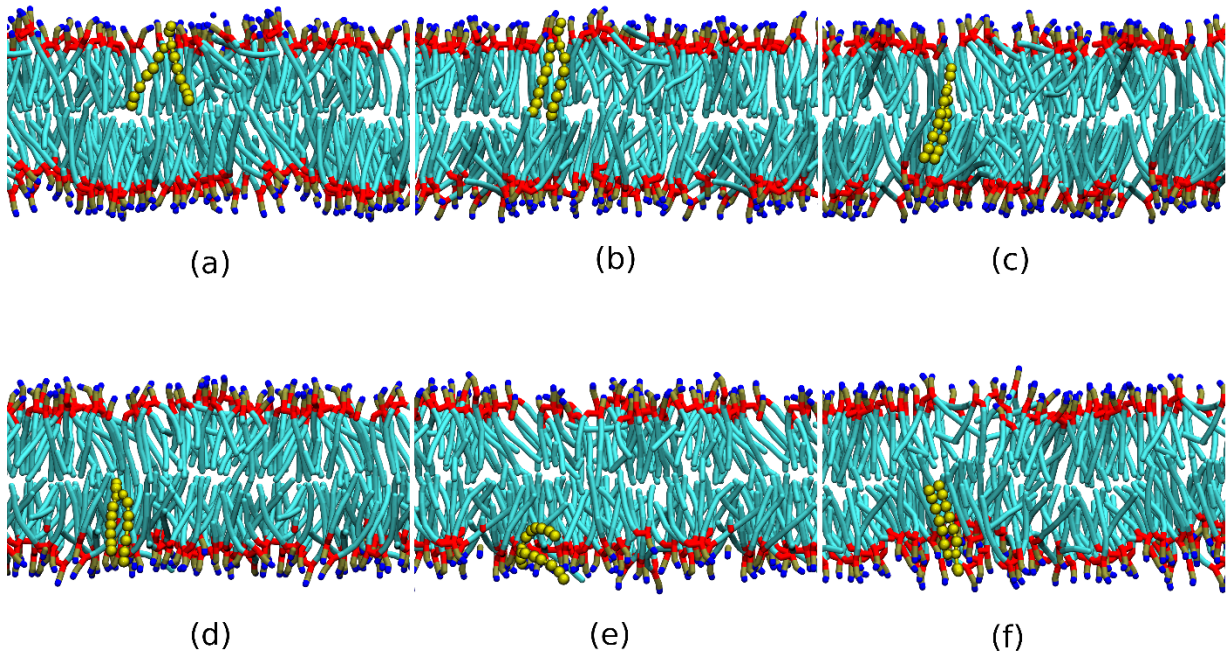

Figure 10. A series of snapshots from CG/DMD simulation of a DSPE bilayer at 370 K showing the translocation of a phospholipid (yellow) from upper to lower leaflet.

Our CG simulations also predicted the translocation of phospholipids (known as "flip-flop"), which is defined to be the trans-bilayer movement of a phospholipid from one leaflet to another. "Flip flop" plays an important role in many cellular activities and in the drug delivery process. We examined the trajectory files obtained at each simulation to see if this phenomenon occurred. Translocation of a lipid was only observed in the simulation at 370 K, which is the liquid phase. Figure 10 displays a series of snapshots in which a lipid (yellow) in the bilayer moves from the upper layer to the bottom layer. Two interesting features are noticeable in the snapshots. The first is that the lipid moves from one leaflet to another when the two tails of a lipid are in tight contact. Figure 10 a, b and c are snapshots of the morphological change of a lipid when the positional interchange occurs. Two alkyl chains are far away from each other

initially, but the tails become closer just before the translocation occurs. The second is that a lipid coming from the opposite leaflet changes its orientation at the bilayer surface (Figure 10 d, e and f). This phenomenon is found at very high temperatures because the average distance between the lipid must necessarily be sufficient for the lipid to move freely.

Conclusion

We have developed an improved version of the LIME model for DSPE lipids that can be used in coarse grained (CG) simulation with discontinuous molecular dynamics (DMD) simulation. This was accomplished by using more than one square well to describe each CG interaction and by using the multi-state iterative Boltzmann inversion (MS-IBI) to calculate the CG parameters. Like the original LIME, the improved LIME extracts the parameters for the CG simulation from radial distribution functions (RDFs) based on the atomistic simulation results. In the original LIME, the intermolecular interactions between CG sites are represented using a single square well potential, and a simple single-step Boltzmann inversion technique to extract the intermolecular potentials between CG sites. Those simplifications led to some discrepancies between the RDFs formed in CG simulations and those in atomistic simulations. However, the improved model has multiple square wells which gives us a CG RDF that better reflects the atomistic RDF. In addition, in the original LIME the CG parameters were extracted from atomistic simulation results at a single temperature, which meant that the parameters were likely only valid at that temperature. CG-DMD simulations at a different temperature should in principle have different parameter sets. To overcome this drawback, we introduced a MS-IBI process to calculate a single set of intermolecular parameters that can be applied at various temperatures.

Using the parameters obtained through MS-IBI, the CG DSPE lipid bilayer was simulated at various temperatures from 310 to 370 K. The accuracy of the improved model is confirmed by comparing physical properties of the CG and atomistic models. The phase transition temperature is measured accurately and translocation of a single lipid molecule is found in high temperature simulations. Thus, the obtained parameters are well suited for use in systems where CG simulations need to be performed at various temperatures, such as lipid bilayer phase behavior. This study indicates that CG parameterization using the multi square-well and MS-IBI is highly appropriate for study of lipid bilayers with DMD simulations.

Even though our method is well suited for representing the multi-phase behaviors of lipid bilayers, future research could be directed towards improving the approach so as to obtain a coarse grained potential that more faithfully represents the atomistic potential that it is designed to mimic. This could be done by adding new square wells to represent the second and higher peaks that are observed in the RDFs. Note however that we did observe the second and the third peaks in the RDFs of the CG simulation even though those peaks were not included in defining the force field parameters. This result seems to suggest that the peaks other than the first peak is not that important in determining the intermolecular potential functions. It may be that the appearance of those peaks in RDFs is mainly a reflection of the structural characteristics of the lipid bilayer in which molecules are densely packed. Nevertheless, using square wells that accounting for more than the first peak could lead to accurate representation of pairwise interactions.

The second suggestion is to refine the force field parameters by conducting MS-IBI several times. Theoretically, MS-IBI for our CG/DMD system could produce many force field parameter sets satisfying the convergence tolerance int the iteration process. Because we performed MS-IBI process once, it was possible that our force field parameters were determined to have a better representation at a particular temperature. This could be the reason that our CG simulations did not accurately reproduce the structures of lipid bilayers at some temperatures. We believe that more optimized force field parameter set producing accurate simulation results at all temperatures could be obtained by averaging force field parameters sets from many MS-IBI.

Another possible future direction is to expand the coarse-graining approach to other biomolecular systems. We have already developed CG force fields for DNA,[45] peptides[46], and biocompatible polymers[39, 47], so we could expand our approach to multicomponent systems such as DNA – lipid bilayer, peptide – lipid bilayer, hydrogel – lipid bilayer interactions. It might, for example, be of interest to apply such an approach to the case of amyloidogenic peptides, immersed in a lipid membrane as this could help explain the origin of cell toxicity in amyloid diseases[48] or the effect of Leucine zipper lipopeptides modulating on lipid membrane phase transition temperature for drug delivery[49].

Acknowledgements

This work was supported by the National Institutes of Health, EB006006, the NSF's Research Triangle MRSEC on Programmable Soft Matter, DMR-1121107, CBET-1512059, CBET-1743432, and the NSF Partnership for Research and Education in Materials (PREM), DMR-1205670.

Supplementary material

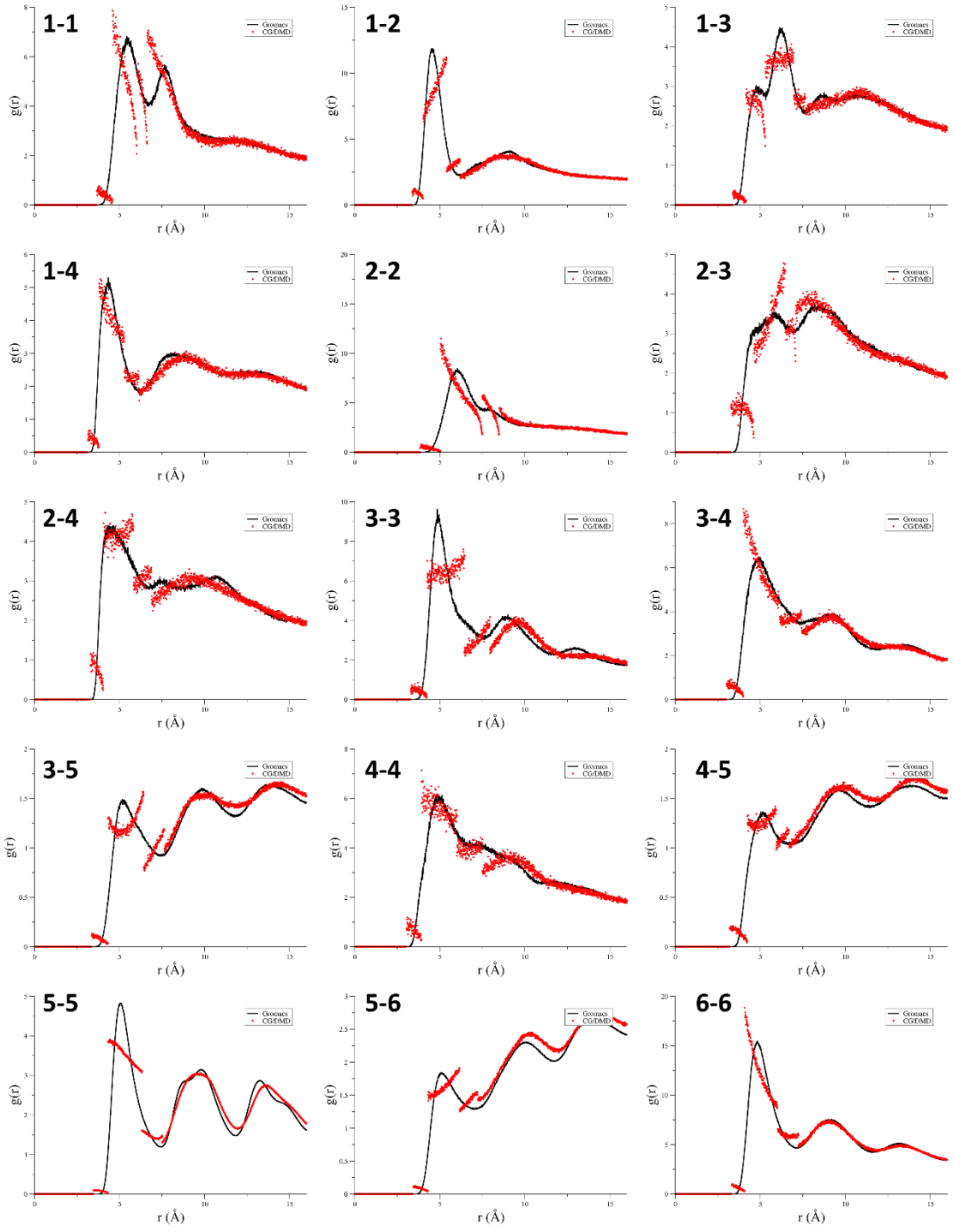

Figure S1. Comparison of atomistic (black) and CG (red) RDFs for each pair of CG sites at 310 K.

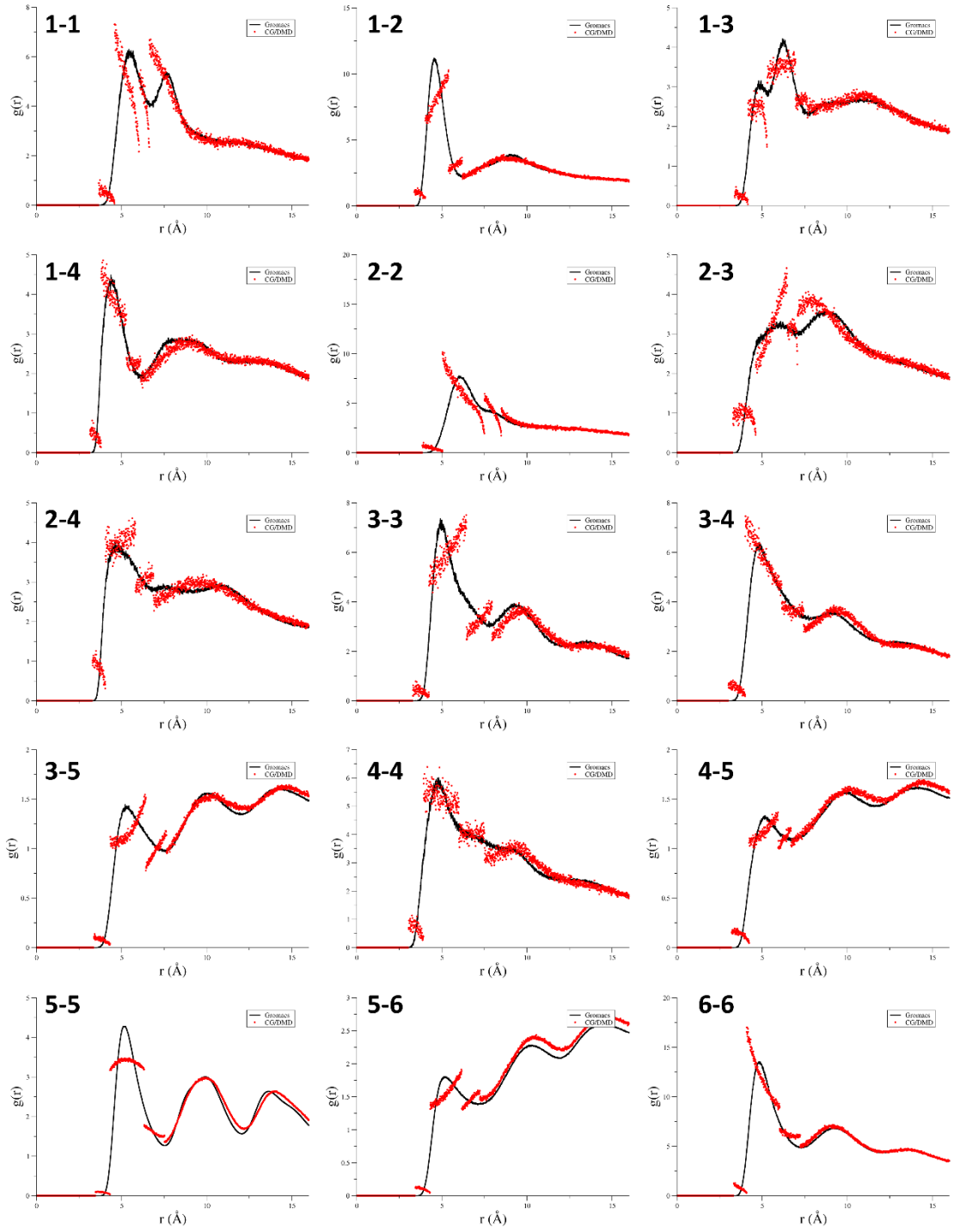

Figure S2. Comparison of atomistic (black) and CG (red) RDFs for each pair of CG sites at 340 K.

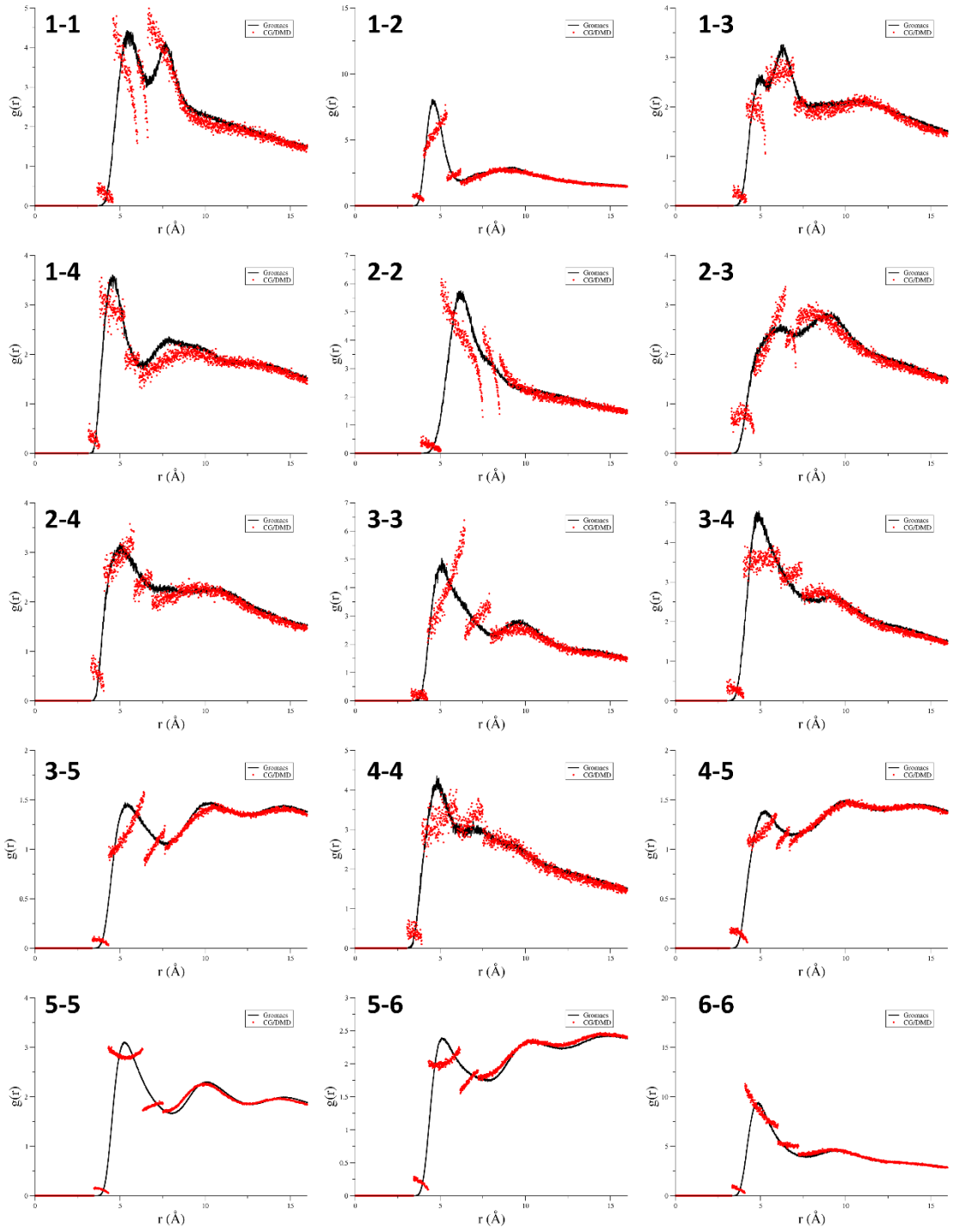

Figure S3. Comparison of atomistic (black) and CG (red) RDFs for each pair of CG sites at 360 K.

Table S.1. Hard sphere diameter and square well widths for each pair of CG type.

CG type i	CG type j	σ (Å)	λ_1 (Å)	λ_2 (Å)	λ_3 (Å)	λ_4 (Å)
1	1	3.655	4.585	6.025	6.625	
1	2	3.375	4.025	5.395	6.195	
1	3	3.385	4.165	5.295	6.965	7.655
1	4	3.135	3.795	5.285	6.125	
1	5	3.545				
1	6	3.625				
2	2	3.855	5.035	7.495	8.495	
2	3	3.275	4.625	6.475	7.075	
2	4	3.285	4.045	5.815	6.875	
2	5	3.575				
2	6	3.415				
3	3	3.285	4.245	6.445	7.945	
3	4	3.015	4.005	6.145	7.435	
3	5	3.365	4.325	6.425	7.635	
3	6	3.395				
4	4	3.035	3.915	6.005	7.505	
4	5	3.205	4.235	5.935	6.695	
4	6	3.245				
5	5	3.465	4.315	6.325	7.515	
5	6	3.435	4.305	6.165	7.225	
6	6	3.325	4.085	6.025	7.235	

Table S.2. Minimum and maximum bond lengths for all bonded CG sites.

CG site i	CG site j	Minimum (Å)	Maximum (Å)
1	2	3.155	3.755
1	3	5.005	7.855
1	4	6.555	11.555
1	10	5.535	9.725
2	3	3.305	4.975
2	4	6.405	8.735
2	5	9.495	12.295
2	10	4.845	6.725
2	11	7.555	9.825
3	4	3.325	4.015
3	5	5.975	7.785
3	6	9.065	11.535
3	10	2.605	4.105
3	11	4.695	6.795
3	12	6.705	10.345
4	5	3.275	4.075
4	6	6.735	7.765
4	7	9.265	11.815
4	10	4.355	6.985
4	11	4.595	7.945
5	6	3.245	3.975
5	7	6.765	7.745
5	8	9.555	11.505
5	10	5.585	10.365
6	7	3.345	3.995
6	8	6.665	7.715
6	9	8.575	10.865
7	8	3.285	3.905
7	9	5.445	7.145
8	9	2.715	3.365
10	11	2.805	3.545
10	12	5.765	7.825
10	13	8.295	10.735
11	12	3.655	4.305
11	13	6.015	7.315
11	14	9.465	11.535
12	13	2.715	3.075
12	14	6.215	7.305
12	15	9.425	11.025
13	14	3.665	4.385
13	15	6.795	8.105
13	16	8.745	11.325
14	15	3.175	3.965
14	16	5.505	7.145
15	16	2.735	3.255

Table S.3. Square well depths for each pair of CG type.

CG type i	CG type j	ε_1 (eV)	ε_2 (eV)	ε_3 (eV)	ε_4 (eV)
1	1	4.5850	1.2365	0.6250	
1	2	0.2050	-1.8797	-0.4750	
1	3	2.7050	0.0872	-0.4050	-0.0800
1	4	1.8600	-0.6104	-0.2100	
1	5	0.0000			
1	6	0.0000			
2	2	5.2050	1.2828	0.5850	
2	3	0.7950	-0.2266	0.1750	
2	4	1.1350	-0.6448	-0.2600	
2	5	0.0000			
2	6	0.0000			
3	3	1.2450	-1.4037	-0.4000	
3	4	2.2900	-0.4751	-0.2500	
3	5	2.0700	-0.7651	-0.1700	
3	6	0.0000			
4	4	1.6400	-0.4701	-0.2250	
4	5	1.9450	-0.3751	-0.1150	
4	6	0.0000			
5	5	2.8250	-0.6250	-0.0900	
5	6	2.2600	-0.4259	-0.0750	
6	6	2.6900	-0.4294	-0.1500	

References

1. Toyoki, K., Synthetic Bilayer Membranes: Molecular Design, Self-Organization, and Application. *Angewandte Chemie International Edition in English* **1992,** *31* (6), 709-726.
2. Simons, K.; Vaz, W. L. C., Model Systems, Lipid Rafts, and Cell Membranes. *Annual Review of Biophysics and Biomolecular Structure* **2004,** *33* (1), 269-295.
3. Israelachvili, J. N., Strength of van der Waals Attraction between Lipid Bilayers. *Langmuir* **1994,** *10* (9), 3369-3370.
4. Parsegian, V. A., Reconciliation of van der Waals force measurements between phosphatidylcholine bilayers in water and between bilayer-coated mica surfaces. *Langmuir* **1993,** *9* (12), 3625-3628.
5. Tero, R.; Fukumoto, K.; Motegi, T.; Yoshida, M.; Niwano, M.; Hirano-Iwata, A., Formation of Cell Membrane Component Domains in Artificial Lipid Bilayer. *Scientific Reports* **2017,** *7* (1), 17905.
6. van Uitert, I.; Le Gac, S.; van den Berg, A., The influence of different membrane components on the electrical stability of bilayer lipid membranes. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2010,** *1798* (1), 21-31.
7. van Meer, G.; Voelker, D. R.; Feigenson, G. W., Membrane lipids: where they are and how they behave. *Nature reviews. Molecular cell biology* **2008,** *9* (2), 112-124.
8. Wu, W.; Shi, X.; Xu, C., Regulation of T cell signalling by membrane lipids. *Nature Reviews Immunology* **2016,** *16*, 690.
9. Lladó, V.; López, D. J.; Ibarguren, M.; Alonso, M.; Soriano, J. B.; Escribá, P. V.; Busquets, X., Regulation of the cancer cell membrane lipid composition by NaCHOleate: Effects on cell signaling and therapeutical relevance in glioma. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2014,** *1838* (6), 1619-1627.
10. Sunshine, H.; Iruela-Arispe, M. L., Membrane lipids and cell signaling. *Current Opinion in Lipidology* **2017,** *28* (5), 408-413.
11. Lundbæk, J. A.; Collingwood, S. A.; Ingólfsson, H. I.; Kapoor, R.; Andersen, O. S., Lipid bilayer regulation of membrane protein function: gramicidin channels as molecular force probes. *Journal of The Royal Society Interface* **2010,** *7* (44), 373-395.
12. Bangham, A. D.; Standish, M. M.; Watkins, J. C., Diffusion of univalent ions across the lamellae of swollen phospholipids. *Journal of Molecular Biology* **1965,** *13* (1), 238-IN27.
13. Akbarzadeh, A.; Rezaei-Sadabady, R.; Davaran, S.; Joo, S. W.; Zarghami, N.; Hanifehpour, Y.; Samiei, M.; Kouhi, M.; Nejati-Koshki, K., Liposome: classification, preparation, and applications. *Nanoscale Research Letters* **2013,** *8* (1), 102-102.
14. Li, J.; Wang, X.; Zhang, T.; Wang, C.; Huang, Z.; Luo, X.; Deng, Y., A review on phospholipids and their main applications in drug delivery systems. *Asian Journal of Pharmaceutical Sciences* **2015,** *10* (2), 81-98.
15. Monteiro, N.; Martins, A.; Reis, R. L.; Neves, N. M., Liposomes in tissue engineering and regenerative medicine. *Journal of the Royal Society Interface* **2014,** *11* (101), 20140459.
16. Lagny, T. J.; Bassereau, P., Bioinspired membrane-based systems for a physical approach of cell organization and dynamics: usefulness and limitations. *Interface Focus* **2015,** *5* (4).
17. Zepik, H. H.; Walde, P.; Kostoryz, E. L.; Code, J.; Yourtee, D. M., Lipid Vesicles as Membrane Models for Toxicological Assessment of Xenobiotics. *Critical Reviews in Toxicology* **2008,** *38* (1), 1-11.
18. Chatin, B.; Mével, M.; Devallière, J.; Dallet, L.; Haudebourg, T.; Peuziat, P.; Colombani, T.; Berchel, M.; Lambert, O.; Edelman, A.; Pitard, B., Liposome-based Formulation for Intracellular Delivery of Functional Proteins. *Molecular Therapy - Nucleic Acids* **2015,** *4*, e244.

19. Sercombe, L.; Veerati, T.; Moheimani, F.; Wu, S. Y.; Sood, A. K.; Hua, S., Advances and Challenges of Liposome Assisted Drug Delivery. *Frontiers in Pharmacology* **2015,** *6*, 286.
20. Pochini, L.; Oppedisano, F.; Indiveri, C., Reconstitution into liposomes and functional characterization of the carnitine transporter from renal cell plasma membrane. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2004,** *1661* (1), 78-86.
21. Peetla, C.; Stine, A.; Labhasetwar, V., Biophysical interactions with model lipid membranes: applications in drug discovery and drug delivery. *Molecular pharmaceutics* **2009,** *6* (5), 1264-1276.
22. Allen, T. M.; Cullis, P. R., Liposomal drug delivery systems: From concept to clinical applications. *Advanced Drug Delivery Reviews* **2013,** *65* (1), 36-48.
23. Adlakha-Hutcheon, G.; Bally, M. B.; Shew, C. R.; Madden, T. D., Controlled destabilization of a liposomal drug delivery system enhances mitoxantrone antitumor activity. *Nature Biotechnology* **1999,** *17*, 775.
24. Bozzuto, G.; Molinari, A., Liposomes as nanomedical devices. *International Journal of Nanomedicine* **2015,** *10*, 975-999.
25. Matos, C.; Lima, J. L. C.; Reis, S.; Lopes, A.; Bastos, M., Interaction of Antiinflammatory Drugs with EPC Liposomes: Calorimetric Study in a Broad Concentration Range. *Biophysical Journal* **2004,** *86* (2), 946-954.
26. Lombardo, D.; Calandra, P.; Barreca, D.; Magazù, S.; Kiselev, M. A., Soft Interaction in Liposome Nanocarriers for Therapeutic Drug Delivery. *Nanomaterials* **2016,** *6* (7), 125.
27. Sharma, A.; Sharma, U. S., Liposomes in drug delivery: progress and limitations. *International Journal of Pharmaceutics* **1997,** *154* (2), 123-140.
28. Fetterly, G. J.; Straubinger, R. M., Pharmacokinetics of paclitaxel-containing liposomes in rats. *AAPS PharmSci* **2003,** *5* (4), 90-100.
29. Felgner, P. L.; Gadek, T. R.; Holm, M.; Roman, R.; Chan, H. W.; Wenz, M.; Northrop, J. P.; Ringold, G. M.; Danielsen, M., Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. *Proceedings of the National Academy of Sciences* **1987,** *84* (21), 7413.
30. Ulander, J.; Haymet, A. D. J., Permeation Across Hydrated DPPC Lipid Bilayers: Simulation of the Titrable Amphiphilic Drug Valproic Acid. *Biophysical Journal* **2003,** *85* (6), 3475-3484.
31. Bendas, G.; Vogel, J.; Bakowski, U.; Krause, A.; Müller, J.; Rothe, U., A liposome-based model system for the simulation of lectin-induced cell adhesion. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1997,** *1325* (2), 297-308.
32. Levitt, M.; Warshel, A., Computer simulation of protein folding. *Nature* **1975,** *253*, 694.
33. Warshel, A.; Levitt, M., Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *Journal of Molecular Biology* **1976,** *103* (2), 227-249.
34. Marrink, S. J.; de Vries, A. H.; Mark, A. E., Coarse grained model for semiquantitative lipid simulations. *Journal of Physical Chemistry B* **2004,** *108* (2), 750-760.
35. Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H., The MARTINI force field: Coarse grained model for biomolecular simulations. *Journal of Physical Chemistry B* **2007,** *111* (27), 7812-7824.
36. Curtis, E. M.; Hall, C. K., Molecular Dynamics Simulations of DPPC Bilayers Using “LIME”, a New Coarse-Grained Model. *The Journal of Physical Chemistry B* **2013,** *117* (17), 5019-5030.
37. Curtis, E. M.; Xiao, X.; Sofou, S.; Hall, C. K., Phase Separation Behavior of Mixed Lipid Systems at Neutral and Low pH: Coarse-Grained Simulations with DMD/LIME. *Langmuir* **2015,** *31* (3), 1086-1094.
38. Rutkowski, D. M.; Velev, O. D.; Klapp, S. H. L.; Hall, C. K., The effect of charge separation on the phase behavior of dipolar colloidal rods. *Soft Matter* **2016,** *12* (22), 4932-4943.

39. Benner, S. W.; Hall, C. K., Development of a Coarse-Grained Model of Chitosan for Predicting Solution Behavior. *The Journal of Physical Chemistry B* **2016,** *120* (29), 7253-7264.
40. Agrawal, V.; Arya, G.; Oswald, J., Simultaneous Iterative Boltzmann Inversion for Coarse-Graining of Polyurea. *Macromolecules* **2014,** *47* (10), 3378-3389.
41. Moore, T. C.; Iacovella, C. R.; McCabe, C., Derivation of coarse-grained potentials via multistate iterative Boltzmann inversion. *The Journal of Chemical Physics* **2014,** *140* (22), 224104.
42. Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R., GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications* **1995,** *91* (1), 43-56.
43. Qin, S.-S.; Yu, Z.-W.; Yu, Y.-X., Structural Characterization on the Gel to Liquid-Crystal Phase Transition of Fully Hydrated DSPC and DSPE Bilayers. *The Journal of Physical Chemistry B* **2009,** *113* (23), 8114-8123.
44. Kastantin, M.; Ananthanarayanan, B.; Karmali, P.; Ruoslahti, E.; Tirrell, M., Effect of the Lipid Chain Melting Transition on the Stability of DSPE-PEG(2000) Micelles. *Langmuir* **2009,** *25* (13), 7279-7286.
45. Wang, K. W.; Barker, K.; Benner, S.; Betancourt, T.; Hall, C. K., Development of a simple coarse-grained DNA model for analysis of oligonucleotide complex formation. *Molecular Simulation* **2018,** *44* (12), 1004-1015.
46. Cheon, M.; Chang, I.; Hall, C. K., Extending the PRIME model for protein aggregation to all 20 amino acids. *Proteins: Structure, Function, and Bioinformatics* **2010,** *78* (14), 2950-2960.
47. Wang, K. W.; Betancourt, T.; Hall, C. K., Computational Study of DNA-Cross-Linked Hydrogel Formation for Drug Delivery Applications. *Macromolecules* **2018,** *51* (23), 9758-9768.
48. Dignon, G. L.; Zerze, G. H.; Mittal, J., Interplay Between Membrane Composition and Structural Stability of Membrane-Bound hIAPP. *The Journal of Physical Chemistry B* **2017,** *121* (37), 8661-8668.
49. Xu, X.; Xiao, X.; Wang, Y.; Xu, S.; Liu, H., Modulation of phase transition of thermosensitive liposomes with leucine zipper-structured lipopeptides. *Physical Chemistry Chemical Physics* **2018,** *20* (23), 15916-15925.