

# A Systematically Coarse-Grained Solvent-Free Model for Quantitative Phospholipid Bilayer Simulations

Zun-Jing Wang\* and Markus Deserno

Department of Physics, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, Pennsylvania 15213

Received: March 21, 2010; Revised Manuscript Received: June 10, 2010

We present an implicit solvent coarse-grained (CG) model for quantitative simulations of 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) bilayers. The absence of explicit solvent enables membrane simulations on large length and time scales at moderate computational expense. Despite improved computational efficiency, the model preserves chemical specificity and quantitative accuracy. The bonded and nonbonded interactions together with the effective cohesion mimicking the hydrophobic effect were systematically tuned by matching structural and mechanical properties from experiments and all-atom bilayer simulations, such as saturated area per lipid, radial distribution functions, density and pressure profiles across the bilayer,  $P_2$  order, etc. The CG lipid model is shown to self-assemble into a bilayer starting from a random dispersion. Its line tension and elastic properties, such as bending and stretching modulus, are semiquantitatively consistent with experiments. The effects of (i) reduced molecular friction and (ii) more efficient integration combine to an overall speed-up of 3–4 orders of magnitude compared to all-atom bilayer simulations. Our CG lipid model is especially useful for studies of large-scale phenomena in membranes that nevertheless require a fair description of chemical specificity, e.g., membrane patches interacting with movable and transformable membrane proteins and peptides.

## 1. Introduction

Phospholipid bilayers are among the most versatile and significant biomaterials in nature. They enclose and protect all living cells and compartmentalize all eukaryotes into many specific organelles (such as nucleus, endoplasmic reticulum, Golgi apparatus, etc.) between which they enable highly regulated membrane-mediated traffic.<sup>1,2</sup> The biological functions of phospholipid bilayers depend strongly on their structural and mechanical properties,<sup>3,4</sup> and a great deal has been learned from experimental techniques such as NMR, scattering techniques, micropipet aspiration, etc.<sup>4–11</sup> However, important examples remain in which complex biomembrane behavior is key to biological function but hard to probe directly in experiment. In this situation (and in fact through all branches of physics), computational physics has long been valued as an attractive option to complement the laboratory.<sup>12–16</sup>

A strategy which has offered much insight in the past is to perform all-atom (AA) simulations.<sup>17–20</sup> However, experimental limitations are often felt particularly dearly on the mesoscale, the range smaller than what can be optically accessed but larger than localized chemistry—loosely speaking, tens to hundreds of nanometers and micro- to milliseconds in time scale. Unfortunately, with current (and foreseeable) computational resources, fully atomistic simulations are generally restricted to a few hundred to at most a few thousand lipids (10–50 nm) for up to a few hundred nanoseconds (smaller system in a few cases can be simulated for up to the order of a microsecond).<sup>21</sup> In this situation, simplified models promise to bridge the gap between experiment and simulation. Of particular interest to us are particle-based coarse-grained (CG) models, in which lipids or small lipid patches are replaced by objects with much fewer degrees of freedom.<sup>22–25</sup> The majority of these models follows

a *top-down* scheme,<sup>26–36</sup> in which not only the resolution is lowered but also no explicit effort is made to rescue local chemical specificity. Rather, these models are tuned to reproduce global observables, such as phase behavior or elastic properties, and are thus most useful when lipid- or chemistry-specific detail can safely be neglected.

However, the chemical composition of the fatty acid chains or the nature of the hydrophilic headgroup does not just determine macroscopic membrane properties; they matter if one wants to understand, say, how chain unsaturation affects fluidity or rigidity, or how inserted proteins interact with the bilayer and each other. Since these questions can still involve large scales (think for examples of antimicrobial peptide aggregation<sup>37,38</sup>), coarse-graining often remains the only viable way, but local detail must now be rescued in a systematic effort aimed to quantitatively bridge the finer and coarser scales. Along these lines, several *bottom-up* methodologies have been developed for deriving more reliable and computationally efficient biomembrane models. For instance, by matching structural properties, Shelley et al.<sup>39</sup> and their followers<sup>40</sup> developed explicit-solvent lipid CG models which quantitatively reproduce the radial distribution functions (RDFs) and cross-bilayer density profiles from fully atomistic simulations. Another approach is to match lipid partitioning coefficients between oil and water;<sup>28,41</sup> a recent popular incarnation of this strategy is the MARTINI force field<sup>42,43</sup> and its extension to peptides.<sup>44</sup> Simulations show that these models then also reproduce the cross-bilayer density profile,  $P_2$  orientational order parameter, and elastic properties.<sup>42,43</sup> Finally, Izvekov and Voth derived effective CG force fields by matching the total force on each CG particle to the one observed in the corresponding all-atom system.<sup>45</sup> This MS-CG methodology optimizes a different merit function while approximating the full  $N$ -body potential of mean force and thus generally differs from force fields derived through structure matching protocols. However, since both are approximative (structure matching

\* E-mail: zwang@cmu.edu or zunjingwang@gmail.com.

neglects higher than two-body correlations and force-matching projects the  $N$ -body potential onto the smaller subspace of realized interactions), no generally valid quality statement can be made (see ref 46 for more details.)

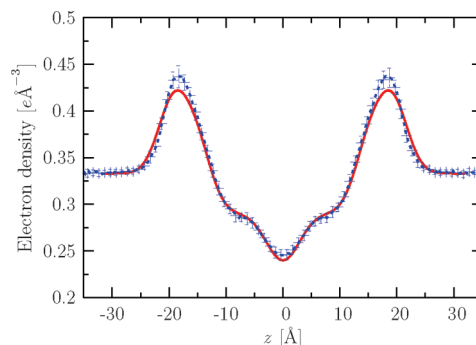
Given that membranes are two-dimensional structures in three-dimensional space, accounting for the embedding solvent is prone to becoming the computationally dominating task, especially once the membrane bends significantly. This makes any strategy that eliminates the solvent very attractive.<sup>47</sup> Successes in removing solvent particles from simulations have first been achieved in top-down CG models.<sup>29–36</sup> These models, however, never aimed at reliable structural information. Recently, Murtola et al. found that structure matching based on inverse Monte Carlo gave rise to some unphysical thermodynamic properties of their solvent-free CG lipid model.<sup>48</sup> Although these authors explicitly derived their CG force field from a bilayer state, the resulting CG lipids occasionally formed dense clusters separated by empty space. In order to stabilize the membrane state, Murtola et al. had to artificially couple a surface tension constrain to the CG model. This suggests that structure matching by itself might not suffice for arriving at a solvent-free CG lipid model which yields a stable stress-free membrane. Izvekov and Voth also recently presented a bottom-up solvent-free CG lipid model based on the MS-CG protocol.<sup>49</sup> Although their CG model was shown to reproduce most atomistic RDFs (a nontrivial statement if mapping is based on forces), their CG lipids self-assembled into “bicelle”-like aggregates in which boundary hydrocarbon chains remained exposed at the boundaries (see ref 49). At this point it is unclear whether this is a deficiency of the parametrization or merely an indication of a kinetically trapped state.

Our primary goal in this paper is to propose a novel implicit solvent CG model for lipid membranes which overcomes several of the difficulties mentioned above. We derive a structure-matched bottom-up solvent-free CG lipid force field which yields a stable zero-tension membrane at the correct area per lipid, while retaining computational efficiency, chemical specificity, structural accuracy, and, as a consequence, essentially correct elastic properties. Along the way, we also suggest a strategy for how to incorporate aspects of the equation of state (specifically, the question what the correct zero-tension state should be) into the iteration loop that refines the CG potentials. Finally, since our coarse-graining scheme differs both from the one employed by Murtola et al.<sup>48</sup> and Izvekov and Voth,<sup>49</sup> we thereby also provide an alternative model which will be valuable when assessing remaining questions about the still subtle step of eliminating the solvent in simulations aiming to preserve specificity.

The remainder of this article is organized as follows. In section 2, we describe the simulation details, including the reference system, CG model, the method used to derive the CG force field, and the resulting CG interaction potentials. In section 3, we evaluate the performance of our model in terms of RDFs, cross-bilayer density profile, pressure profile, the saturated area per lipid, self-assembly, elastic properties, line tension, and diffusion behavior.

## 2. Model and Method

**2.1. Reference System.** The reference data for coarse graining (e.g., RDFs,  $P_2$  order parameter, density profiles, pressure profiles, etc.) were obtained from all-atom (AA) simulations. Also, available experimental data were compared with AA simulations to examine their reliability. Beyond this, experiments



**Figure 1.** Comparison of electron density profiles between experiment and simulation. The red solid line represents experimental data; the blue dashed line and error bars were computed via block average of an equilibrated all-atom  $NP_{zz}AT$  simulation.

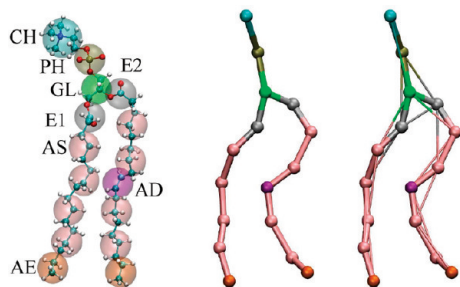
were also used as reference for saturated area per lipid (i.e., the zero-tension area per lipid) as well as stretching and bending modulus.

All AA simulations were performed with the parallel molecular dynamics (MD) program NAMD.<sup>50</sup> Initial coordinates of the lipids were taken from a previous simulation by Feller and MacKerell.<sup>51</sup> All systems were simulated using the fully atomistic CHARMM27<sup>52</sup> parameters to describe the lipid interactions. The visualizations were performed with VMD.<sup>53</sup>

Periodic boundary conditions were applied and a constant temperature (310 K) was maintained using a Langevin thermostat with a damping coefficient of  $0.5 \text{ ps}^{-1}$ . Constant pressure (1 atm) in the simulations was obtained using a Langevin-piston barostat with a piston period of 2 ps, a damping time of 2 ps, and a fully anisotropic pressure coupling.<sup>51</sup> The long-range electrostatic interactions were computed every time step with the particle mesh Ewald (PME) algorithm,<sup>54</sup> employing a real-space cutoff of 12 Å. The integration time step was 1 fs.

The smallest square membrane system consisted of 72 lipids fully hydrated with 3881 water molecules and equilibrated for about 30 ns in the  $NP_{zz}AT$  ensemble, with the pressure normal to the membrane  $P_{zz} = 1 \text{ atm}$  (box size in  $z$ -direction  $\langle L_{zz} \rangle = 83.8 \pm 0.3 \text{ Å}$ ) and the projected membrane area  $A = L_x \times L_y$  corresponding to an area per lipid of  $68.3 \text{ Å}^2$  (the saturated area measured in experiments<sup>9</sup>). A larger membrane systems with 288 lipids was then created from this base system by replicating it along both the  $x$ - and  $y$ -directions. A comparison of electron density profiles between AA simulations and experiment<sup>9</sup> is shown in Figure 1, which validates the structure from an area-fixed AA simulation. Note that Debye–Waller broadening due to thermal undulations of the bilayer has been removed from the experimental data by a deconvolution based on the known value of the bilayer elasticity.<sup>5</sup> Hence, a comparison with the essentially fluctuation-free result from the simulations of a small bilayer patch is permissible. In complementary simulations using the  $NP_{zz}\Sigma T$  ensemble (with an applied tension  $\Sigma = 0$ ), we found that the saturated area shrinks by 15–18% compared to experiments. This is a well-known artifact of the CHARMM27 force field<sup>55</sup> which we will explicitly address during the coarse-graining procedure. Even though the bilayer at the correct area per lipid is therefore under stress, a comparison of structure is nevertheless believed to be meaningful.<sup>56</sup> The reason is that stresses react far more sensitively to changes in interactions.

When computing the density and pressure profiles, we aligned the center of mass of the membrane for each frame individually before block averaging data. The pressure profiles were calculated at constant volume and fixed area of the simulation box. The instantaneous local pressure tensors were computed



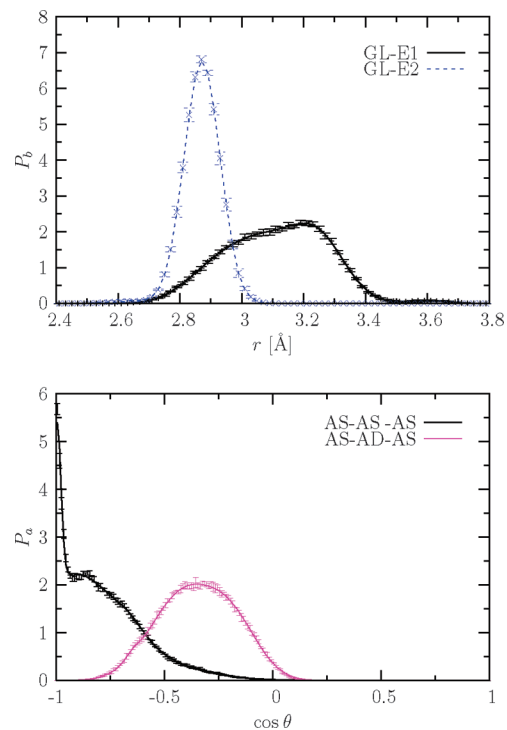
**Figure 2.** Structure of the POPC lipid molecule. The left panel shows the mapping scheme from AA to CG, the middle panel shows the topology of the lipid molecule in a CG simulation with multibody interactions (where the three-body angle interactions and four-body improper dihedral interactions are not explicitly shown), and the right panel shows the topology of the lipid molecule in which all the multibody interactions have been replaced with effective two-body interactions. Notice that the size of the CG beads in the middle and right panel does not represent the real excluded volume of each bead.

with the method presented by Lindahl and Edholm,<sup>57</sup> where the simulation space was partitioned into approximately 1 Å wide slabs along the average normal of the membrane (i.e., the  $z$ -direction), and the instantaneous pressure was computed from all kinetic contributions and pairwise interparticle interactions. Pairwise interactions include the covalent two-, three-, and four-body terms of the bonded forces (three- and four-body terms were decomposed into simple two-body forces in the calculation of the instantaneous pressure), as well as the nonbonded electrostatic and van der Waals forces. The contribution to the local pressure tensor from the kinetic energy and the covalent interactions were first computed by conducting simulations using PME as described above, with coordinates saved every 1000 fs. These saved coordinates were subsequently analyzed using the direct Coulomb interaction truncated at the cutoff distance of 20 Å. This truncation of the electrostatic potential is sufficiently long to reproduce a similar local pressure profile to the Ewald profile of bilayers.<sup>58</sup> All spatially resolved stress profiles were sampled over 20 ns after equilibration. For an anisotropic lipid bilayer system, the lateral membrane stress is obtained from the pressure tensor according to ref 59

$$\Sigma = - \int_{-\infty}^{\infty} dz [P_n - P_t(z)] \quad (1)$$

where  $P_n = \langle P_{zz} \rangle$  is the normal and  $P_t(z) = (1/2)\langle P_{xx}(z) + P_{yy}(z) \rangle$  the average lateral pressure. While  $P_n$  is constant for reasons of mechanical stability,  $P_t(z)$  depends on the depth inside the bilayer.

**2.2. CG Model.** Each POPC lipid molecule is mapped onto a structure consisting of 16 CG sites and 8 different bead types (see Figure 2). We will use the following notation: CH for the choline entity of the headgroup; PH for the phosphate group; GL for the glycerol backbone; E1 and E2 for the ester groups of the sn-1 and sn-2 tails of the lipid, respectively; AS, AD, and AE for the hydrocarbon groups  $-(\text{CH}_2-\text{CH}_2-\text{CH}_2)-$ ,  $-(\text{CH}_2=\text{CH}_2)-$ , and  $-(\text{CH}_2-\text{CH}_3)-$ , respectively (see Figure 2). This mapping is at the same resolution as previous work.<sup>39,42,49</sup> In addition, we distinguish the unsaturated alkyl groups AD with a double bond from the saturated AS groups similar to the improved MARTINI force field.<sup>43</sup> This is important, because lipid chain orientational order and bilayer phase behavior are sensitive to the level of alkyl chain unsaturation.<sup>60,61</sup> Figure 3 shows a comparison of the probability distribution of the



**Figure 3.** (a, top) Comparison of the distribution probabilities of bonds between AA simulations and fitting data via eq 6. (b, bottom) Comparison of the distribution probabilities of CG bond angles across a saturated and unsaturated CG bead between AA simulations and fitting data via eq 7.

intramolecular bond angle between AS–AD–AS and AS–AS–AS, from which it is obvious that this distinction is necessary to properly capture the structural difference between saturated and unsaturated tails.

The comparison of bond distributions between the CG sites E1 and E2 is also shown in Figure 3. Although E1 and E2 are different in bonded interactions, their nonbonded interacting potentials with other particles should not differ since they have the same chemical structure. We thus denote both E1 and E2 as ES when referring to their nonbonded interactions.

For each of the CG beads, the center of mass of the atom groups is defined as the CG center. We will also separate the interactions between CG beads into bonded and nonbonded interactions, i.e.

$$U^{\text{cg}}(r) = \sum U_{\text{b}}^{\text{cg}}(r) + \sum U_{\text{nb}}^{\text{cg}}(r) \quad (2)$$

Neither is implementing this in practice trivial nor is the answer unique, because the available atomistic pair correlation functions entangle the contributions from bonded and nonbonded forces. However, due to the physically distinct role played by both types of potentials, the effort is worthwhile, because it should improve transferability of the force field.<sup>62</sup>

For the intramolecular bonds, angles, and dihedrals, the potentials of mean force are obtained from their probability distributions:

$$U_{\text{b}}^{\text{cg}}(r) = -k_{\text{B}}T \ln \frac{P_{\text{b}}(r)}{r^2} + \text{const} \quad (3)$$

$$U_a^{\text{CG}}(\cos \theta) = -k_B T \ln P_a(\cos \theta) + \text{const} \quad (4)$$

$$U_d^{\text{CG}}(\psi) = -k_B T \ln P_d(\psi) + \text{const} \quad (5)$$

where  $k_B$  is the Boltzmann constant,  $T$  the temperature in Kelvin,  $r$  the bond length,  $\theta$  the angle between two intramolecular bonds, and  $\psi$  the intramolecular dihedral angle.

Figure 3 shows an example of the distributions of bonds and angles on the CG level obtained from AA simulations. The distribution of bonds can be successfully described as a sum of 2 to 5 Gaussian functions:

$$P_b(r) = \sum_i a_{b,i} \exp\left\{-\frac{(r - b_{b,i})^2}{c_{b,i}^2}\right\} \quad (6)$$

and similarly the angle distributions can be described by

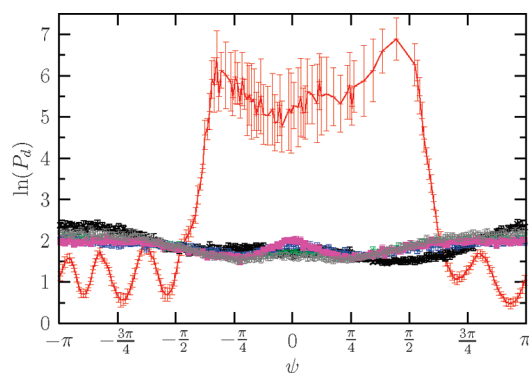
$$P_a(\cos \theta) = \sum_i a_{a,i} \exp\left\{-\frac{(\cos \theta - b_{a,i})^2}{c_{a,i}^2}\right\} \quad (7)$$

with suitable fitting coefficients of  $a_{b,i}$ ,  $b_{b,i}$ ,  $c_{b,i}$ , etc.

Figure 4 shows  $\ln[P_d(\psi)]$  for several intramolecular dihedral angles between CG groups obtained from AA simulations. Except for the improper dihedral PH–GL–E1–E2, all the other effective dihedrals vary on the order of the thermal energy  $k_B T$ , so we chose to neglect the latter in our CG model.

It is of course feasible to perform CG simulations with multibody interactions such as the ones in eqs 4 and 5. However, computing surface tension and pressure profiles of a simulation system with multibody interactions is nontrivial.<sup>63</sup> In order to simplify our CG force field, we therefore replaced all multibody intramolecular interactions with effective pairwise bonded interactions (see Figure 2, from middle panel to right panel). Using this CG model then leads to 21 types of bonded interactions and 28 types of nonbonded interactions (see Supporting Information for more details).

**2.3. CG Simulations.** We performed CG MD simulations in both the *NAT* and the *NΣT* ensemble, where  $N$  is the number of CG particles,  $A$  the box area,  $T$  the temperature, and  $\Sigma$  the lateral bilayer tension. The units used in our CG simulations are  $\zeta = 1 \text{ \AA}$  (length),  $\varepsilon = k_B T \approx 4.28 \times 10^{-21} \text{ J} \approx 2.58 \text{ kJ/mol}$



**Figure 4.** Selected distributions of dihedral angles. The red line shows the improper dihedral PH–GL\*–E1–E2; all other curves (corresponding to the dihedrals CH–PH–GL–E2, E1–AS–AS–AS, E2–AS–AS–AD, AD–AS–AS–AE, and AS–AS–AS–AS) show a variation that justifies their neglect in our CG simulation.

$\approx 0.617 \text{ kcal/mol}$  (energy) at  $T = 310 \text{ K}$ ,  $m = 1 \text{ Da}$  (mass),  $\tau = \zeta(m/\varepsilon)^{1/2} \approx 0.062 \text{ ps}$  (time). Notice that, although the units for length, temperature, and mass retain their physical meaning also in the CG system, the same does not hold for the time unit  $\tau$ . It is the appropriate characterization of the dynamics only if one considers *instantaneous* properties such as the kinetic energy, but it does *not* describe the correct *long-time* dynamics of the corresponding AA system, because there is no reason why coarse-graining would preserve *dynamical* properties. In fact, due to the strongly reduced molecular friction in a system with fewer degrees of freedom, the dynamics tends to be substantially speeded up, an effect which is usually considered beneficial. We will discuss this in more detail in section 3.

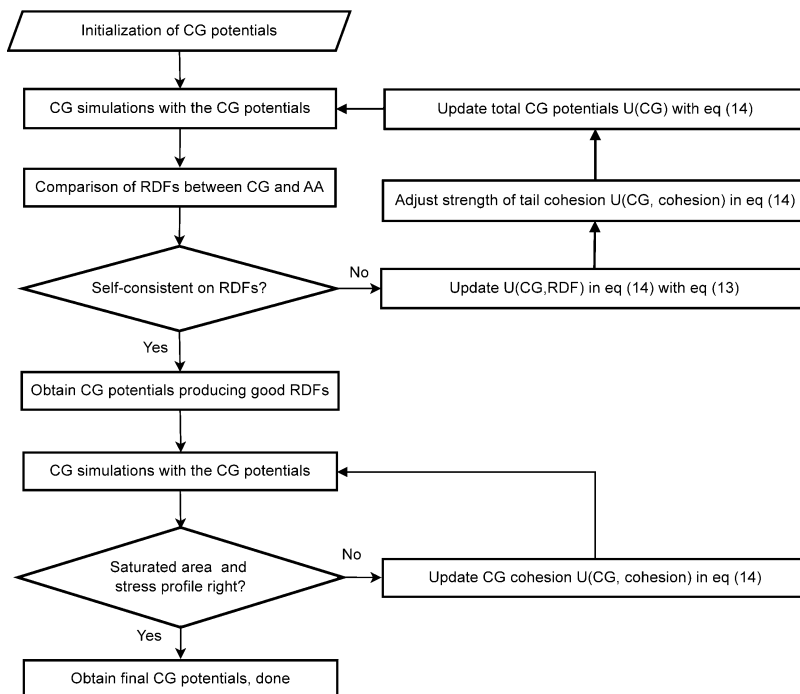
The constant temperature in all the CG simulations was achieved via a Langevin thermostat<sup>64</sup> (with a time step  $\delta t = 0.1\tau$  and a friction constant  $\Gamma = 0.2\tau^{-1}$ ). Constant-volume simulations were implemented using a quadratic cross-sectional area in bilayer  $xy$ -plane, i.e.,  $L_x = L_y$ . Periodic boundary conditions were applied in all three directions and  $L_z$  was set large enough so that the membrane will not interact with its periodic image in  $z$ -direction ( $L_z \approx 167\zeta$  in our simulations). In the *NΣT* ensemble, the constant tension conditions were achieved via a modified Andersen barostat,<sup>65</sup> allowing for box resizing only in  $x$ - and  $y$ -directions (with a box friction  $\Gamma_{\text{box}} = 4 \times 10^{-5}\tau^{-1}$  and a box mass  $Q = 10^{-5} - 10^{-4}m$ ). All CG simulations were performed using ESPResSo.<sup>66</sup>

**2.4. CG Force Field Derivation.** The difficulties encountered by Murtola et al.<sup>48</sup> when following the standard structure-based bottom-up CG strategies (see our Introduction) suggest that elimination of the solvent entails difficulties that need to be addressed explicitly. Indeed, after the water is removed the hydrophobic effect must be reincorporated *by hand* and is *not* automatically contained in the potentials derived from the AA pair correlation functions extracted from the bilayer state (a subset of phase space!). The reason for this is easy to understand: imagine a system of two particle types, A and B, with purely repulsive interactions but a  $\chi$ -parameter which drives demixing. The pair correlation functions between A particles inside the demixed A phase can be inverted to the original repulsive A–A-interactions, but of course in the absence of B particles the A particles would *not* aggregate. Rather, the absence of a confining pressure must be compensated by cohesive interactions which are adjusted to reproduce the A density of the segregated binary system in the B-free aggregated unary system.

Hence, we need to incorporate an additional effective cohesion between the lipids. We will then show how to adjust this interaction during the usual iteration loop for potential optimization such that (i) its strength correctly mimics lipid cohesion while (ii) it does not affect the known structure that we of course still aim to reproduce. As it turns out, this additional parameter conveniently permits us to correct the one obvious deficiency of the AA simulation, namely, that the area per lipid at zero tension is too small.

In our CG model we choose to exclude the intramolecular nonbonded interactions between the CG beads which have been covered by bonded interactions. Figure 5 shows the control diagram in deriving the CG force fields with the improved multiscale structure-match methodology. Let us now comment on some critical aspects in a bit more detail:

**1. Choosing good initial potentials.** For the bonded interactions, as shown in subsection 2.2, the initial estimates for the tabulated nonbonded potentials among CG sites  $\alpha$  and  $\beta$  were



**Figure 5.** Flow chart for deriving the solvent-free CG force field for our lipid bilayer membrane. The initialization of CG potentials includes two parts (see eq 14). In the optimization procedure, the Lorentz–Berthelot rule was applied to the alkyl CG beads but this is not explicitly shown inside.

based on the potentials of mean force measured in the AA reference system, i.e.

$$U_{\text{pmf},\alpha\beta}(r) = -k_{\text{B}}T \ln[g_{\alpha\beta}^{\text{AA}}(r)] \quad (8)$$

$$U_{\text{n},\alpha\beta,0}^{\text{CG}}(r) = \lambda_0 U_{\text{pmf},\alpha\beta}(r) \quad (9)$$

where  $g_{\alpha\beta}^{\text{AA}}$  is the RDF between CG types  $\alpha$  and  $\beta$  measured in the AA simulation (where RDFs were tabulated every 0.1 Å), the subscript n represents nonbonded interactions, 0 represents an initial estimate, and  $\lambda_0$  is an adaptive coefficient typically within the range 0.15–0.25.

We computed the AA bond distributions from the atomistic conformations sampled by simulating a single lipid molecule in vacuum and avoided “double counting” of the nonbonded interactions (i.e., erroneously incorporating some of their effects into the bonded interactions) by switching off the intralipid nonbonded interactions for a pair of atoms if they are not explicitly covered via bonded interaction potentials on the CG level (which includes bond and angle potentials).<sup>62,67</sup> These long-range interactions will later be incorporated through CG intramolecular nonbonded interactions.

Similar to the previous SM CG work,<sup>39</sup> we tabulated the nonbonded potentials and matched them to the RDFs of the CG sites of the reference system (AA simulations). For the latter, we used the iterative Boltzmann technique (see below).

For most of the nonbonded interactions, the bilayer state under a saturated area taken from experiment was chosen as our reference systems for RDFs (see Figure 6a) since we aim to reproduce the structure of such a bilayer system. However, the sampled phase space of this reference system has an extremely small probability for the hydrophilic CG beads (e.g., CH) meeting any of the hydrophobic beads (alkyl beads AS, AD, and AE). To alleviate this problem, we cut the bonded interactions between each pair of the alkyl CG representatives

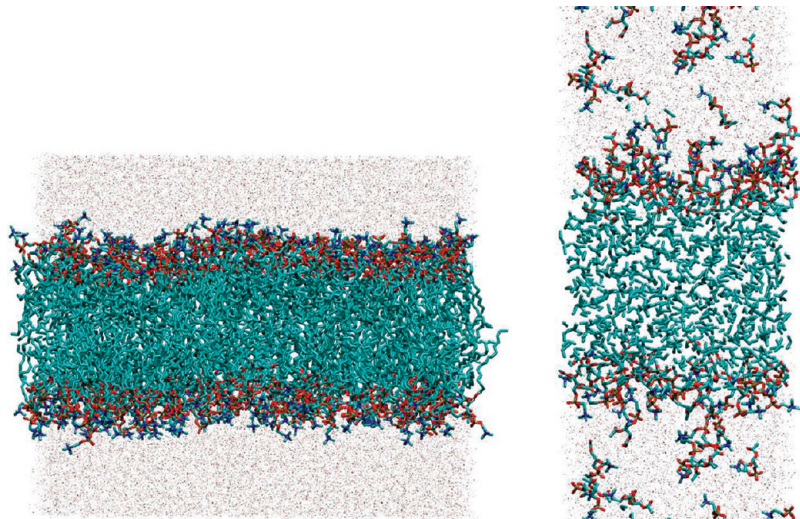
in AA simulations, resulting in another reference system for the initial potentials as shown in Figure 6b in which CG equivalents of head–tail contacts are sampled much more frequently. The initial estimate of nonbonded potentials between CH, PH and AS, AD, AE were obtained from this chopped-up system rather than the intact bilayer. This trick of splitting the target molecule into fragment molecules has been successfully applied in CG modeling of azobenzene molecules with anisotropic structures,<sup>62,67</sup> but has not been tried in the previous CG strategies for membranes.<sup>39,40,42,45</sup> With this initial estimate of the potentials, the excluded volume in the nonbonded interactions between head and tail CG beads is much more accurate. Since the CG potentials are state dependent, the difference in average density between the two reference systems might lead to a variation of the long-range part of the effective interactions, but we can correct this in the subsequent optimization loops where the bilayer membrane is the reference system.

**2. Applications of Lorentz–Berthelot rules.** The Lorentz–Berthelot combining rules have been found valid for simple mixtures of much greater degree of asymmetry.<sup>68</sup> Considering the similarity of the three types of alkyl CG beads in chemical specificity, the assumption that their nonbonded interactions obey Lorentz–Berthelot rules is thus reasonable.

In the parameter derivation, we separate the potentials into five sets according to CH–alkyl, PH–alkyl, GL–alkyl, ES–alkyl, and alkyl–alkyl interactions, as these five sets likely differ in their functional form. For the alkyl–alkyl interactions, using Lorentz–Berthelot rules and choosing the AS–AS interaction as a reference, we define

$$U_{ij}(r) = \frac{\sqrt{\varepsilon_i \varepsilon_j}}{\varepsilon_{\text{AS}}} U_{\text{AS-AS}} \left( \frac{\sigma_i + \sigma_j}{2\sigma_{\text{AS}}} r \right) \quad (10)$$

where  $i, j \in \{\text{AS}, \text{AD}, \text{AE}\}$  and  $\sigma$  and  $\varepsilon$  represent the diameter of exclude volume and potential depth, respectively. For the



**Figure 6.** Snapshots of AA simulation of (a, left) a bilayer with 288 lipids; (b, right) a fragment system of 72 lipids where the AA bonds representing the alkyl CG bonds were cut for each lipid.

**TABLE 1: Parameters Used in the Potential Mixing Rules from Eqs 10 and 11**

CG type	AD	AE	CH	PH	GL	ES
$\sigma/\sigma_{AS}$	0.97	0.96	1.2	1.0	0.9	1.0
$\epsilon/\epsilon_{AS}$	0.44	1.1025	—	—	—	—

other four sets of potentials, we use a slightly modified Lorentz–Berthelot mixing rule, namely

$$U_{ij}(r) = \sqrt{\frac{\epsilon_j}{\epsilon_{AS}}} U_{i-AS} \left( \frac{\sigma_i + \sigma_j}{\sigma_i + \sigma_{AS}} r \right) \quad (11)$$

where  $i \in \{\text{CH, PH, GL, ES}\}$  and  $j \in \{\text{AS, AD, AE}\}$ . Table 1 summarizes the Lorentz–Berthelot parameters in eqs 10 and 11. By use of these rules, the total number of nonbonded interactions that need to be optimized was cut down by a factor of 2 (from 28 to 15), which constitutes a significant technical advantage.

**3. Iterative Boltzmann inversion.** The initial guess for the potentials will not reproduce the target RDFs from AA simulations. One way to iteratively achieve convergence is to use the difference in the AA PMF and the CG PMF in iteration  $i$  as a correction to the PMF used for the next CG simulation  $i + 1$ . This so-called iterative Boltzmann inversion method thus computes<sup>69,70</sup>

$$\Delta U_{n,\alpha\beta,i}^{\text{CG}}(r) = -k_B T \ln[g_{\alpha\beta,i}^{\text{CG}}(r)] - U_{\text{pmf},\alpha\beta}(r) \quad (12)$$

$$U_{n,\alpha\beta,i+1}^{\text{CG}}(r) = U_{n,\alpha\beta,i}^{\text{CG}}(r) - \lambda_{\text{iter},\alpha\beta,i} \Delta U_{n,\alpha\beta,i}^{\text{CG}}(r) \quad (13)$$

where the coefficient  $\lambda_{\text{iter},\alpha\beta,i}$  improves convergence. It typically has values between 0.05 and 0.25 and can be optimized based on the resulting difference in RDFs between a trial short-time CG and AA simulations. In order to calculate the derivatives of these potentials, we fit  $\Delta U_{n,\alpha\beta,i}^{\text{CG}}(r)$  to splines after each iteration.<sup>71</sup> The updated CG nonbonded potentials  $U_{n,\alpha\beta,i+1}^{\text{CG,rdf}}(r)$  were then tabulated every 0.01 Å and truncated at 15 Å, and were used in a new simulation. This iteration was continued until the maximum of the differences between the RDFs was smaller than 0.25.

#### 4. Phenomenological cohesion mimicking the hydrophobic effect.

As explained above, we need to introduce an additional cohesive interaction potential to drive lipid aggregation. Borrowing ideas from top-down CG methods,<sup>33,35,36</sup> we mimic the hydrophobic effect of water via a weak long-range attractive interactions. The cohesion was added to alkyl tails and interfacial beads (PH, GL, ES) of lipid based on the previous top-down CG experience. For these CG sites, the effective force field consists of two parts, namely

$$U^{\text{CG}} = U^{\text{CG,RDF}} + U^{\text{CG,cohesion}} \quad (14)$$

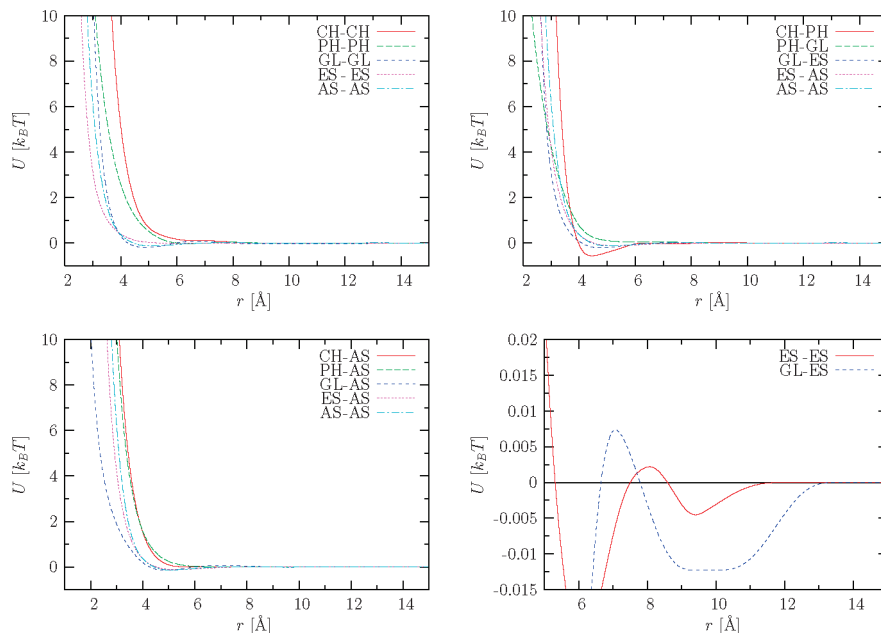
The effective force and potential describing the cohesion are respectively

$$f_{\text{cohesion}}(r) = -\frac{2\epsilon}{w_c} \cos^2(r^*), \quad r_c < r < r_c + w_c \quad (15)$$

$$U_{\text{cohesion}}(r) = \begin{cases} -\epsilon & r \leq r_c \\ -\frac{\epsilon}{2} \left[ 1 - \frac{2}{\pi} [r^* + \sin(r^*) \cos(r^*)] \right] & r_c < r < r_c + w_c \\ 0 & r \geq r_c + w_c \end{cases} \quad (16)$$

where for convenience we defined the scaled distance  $r^* = (\pi/w_c)[r - (r_c + 1/2w_c)]$ . This potential is qualitatively the same as the one proposed in refs 35 and 36, but it differs quantitatively by also making the force a differentiable function of distance, which simplifies later spline fitting. The cohesion strength  $\epsilon$  and its range  $[r_c, r_c + w_c]$  now need to be adjusted. The necessary tuning is based on (i) requiring bilayer stability, (ii) monitoring the lateral stress profile, and (iii) comparing the RDFs of alkyl beads from CG simulations with the reference AA system. The latter two requirements do not play a role in top-down schemes, but they matter for our bottom-up scheme.

We started by adding a weak cohesion to the hydrocarbon tails of lipids and optimize the range of this cohesion and its strength by matching the RDFs between CG and AA simulations. Afterward, we included cohesion terms between the other types of CG beads and also optimized their ranges and strengths



**Figure 7.** Selected nonbonded interacting potentials between CG beads.

according to the saturated area per lipid, the lateral stress profile; at the same time, we tried to keep the difference in RDFs between CG and AA as small as possible. An appropriate range of this cohesion is  $r_c = 8\text{--}10\text{ \AA}$  and  $r_c + w_c = 12\text{--}15\text{ \AA}$ . Both the range and strength of cohesion are CG-type dependent. Its strength turned out to be remarkably small:  $\epsilon$  is within  $0.006\text{--}0.02k_B T$ , but without this small attraction the bilayer falls apart. We also found that the optimized strength of cohesion is stronger for the interfacial CG beads GL and ES ( $\epsilon = 0.01\text{--}0.2k_B T$ ) compared to the hydrocarbon beads ( $\epsilon = 0.006k_B T$ ). This finding is consistent with the previous CG study in a top-down scheme,<sup>33</sup> but in our case it comes out as a result of the optimization process and has not been imposed from the start.

The adjustment of  $f_{\text{cohesion}}$  was included in both iteration loops (see Figure 5). In practice, in the first loop, after getting corrected  $U^{\text{CG, RDF}}$ , we tested several cohesion values and chose the  $U^{\text{CG, cohesion}}$  reproducing the long-range density (the tail part of the RDFs) in AA simulations. The range, however, was not modified, and cohesion only acted between tail beads. This best  $U^{\text{CG, cohesion}}$  is then included in the update of  $U^{\text{CG}}$ . In addition, recall that in the optimization process based on RDFs, rather than remedying incorrect bonded distributions with nonbonded interactions (or vice versa), we worked hard to separately adjust the two. This also helps now, when we add an additional nonbonded interaction which we do not want to affect bonded forces. Overall, the number of iterations via RDFs was around 20 in the case leading to overall well-converged potentials with the CG RDFs agreeing with AA simulations.

In the second iteration loop (see Figure 5), we include all beads except CH in the adjustment of cohesion, and optimize the partitioning of the cohesion in order to match both the saturated area per lipid and stress profile (this required around eight iterations). If locally the bilayer tension is too large/small, the cohesion is locally reduced/increased. After the relative strength has converged, we adjust the overall cohesion strength by a global prefactor that aims to reproduce the correct area. Finally, we made an overall adjustment of the CG force field together by considering both RDFs and the saturated area per lipid.

Figure 7 shows the effective potentials of the solvent-free CG model. Figures 8 and 9 show several selected total RDFs

and intermolecular RDFs respectively. For detailed reference, the Supporting Information shows a comparison of total RDFs, intermolecular RDFs, and intramolecular RDFs among all CG types (see Figures S1, S2, and S3, respectively). All bonded and nonbonded distributions obtained from AA simulations are quantitatively reproduced by our CG force field. Recall that all CG potentials corresponding to CG sites AD and AE were obtained from mixing rules of eqs 10 and 11 using the parameters in Table 1, rather than from individual iterative optimizations. The observed consistency in RDFs between CG and AA simulations also for these interactions indicates that using Lorentz–Berthelot-like combining rules, if carefully handled, are a powerful and reliable way to derive effective force fields for CG sites with similar chemical property.

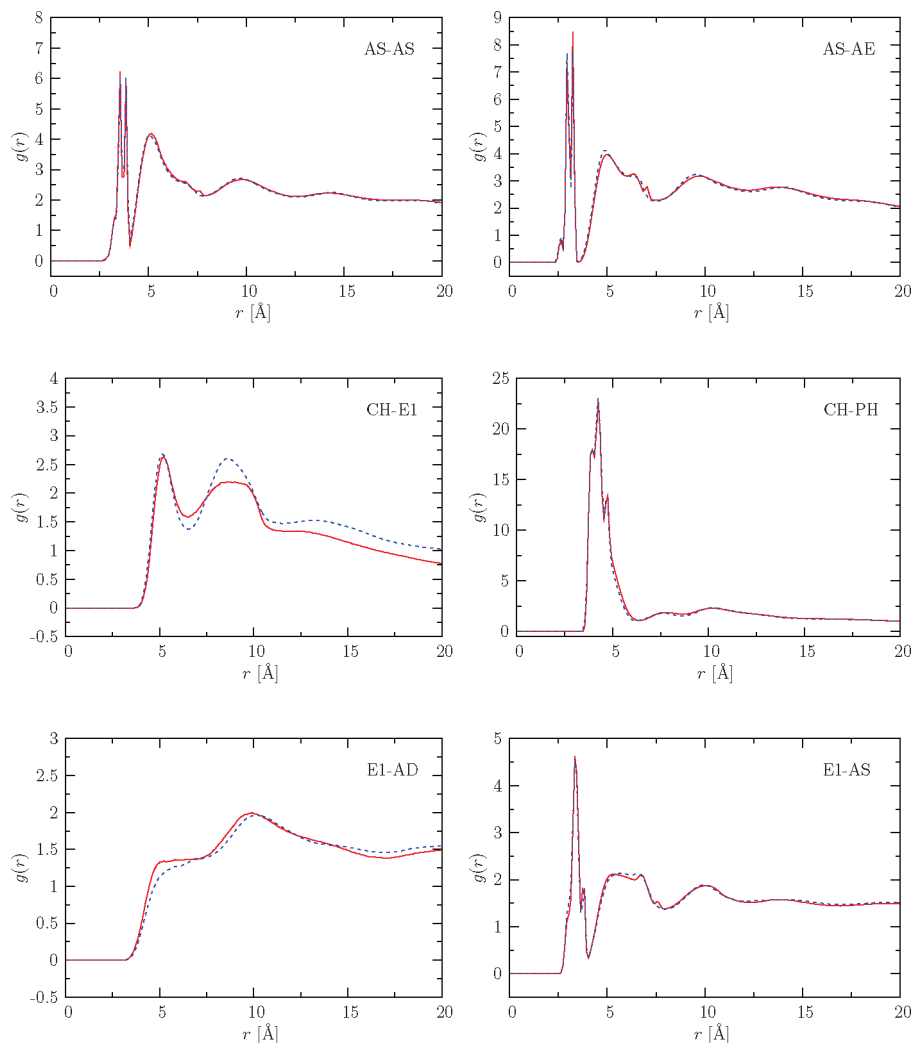
### 3. Performance and Discussion

**3.1. Structure.** A comparison of the cross-bilayer mass density profile (which is now more natural than the electron density profile) between the AA and our CG model is shown in Figure 10. Both AA and CG simulations were performed in the NVT ensemble with the saturated area per lipid fixed to the value  $A_{\text{lipid}} = 68.3\text{ \AA}^2$  reported in experiments.<sup>9</sup> Considering the error bars, the density profiles are mutually consistent, indicating a good reproduction of structure by the solvent-free CG force field.

To evaluate the lipid alignment with the bilayer normal, we computed the orientational order parameter of each lipid bond as

$$S_{\text{bond}} = \frac{1}{2} \langle 3 \cos^2 \theta - 1 \rangle = \langle P_2(\cos \theta) \rangle \quad (17)$$

where  $\theta$  is the angle between the unit vector along some particular CG bond and the average bilayer normal. A completely aligned bond has  $S_{\text{bond}} = 1$ , a completely random (isotropic) bond has  $S_{\text{bond}} = 0$ , and a bond which is perfectly perpendicular to the bilayer normal has  $S_{\text{bond}} = -1/2$ . The comparison of the orientational order parameter of each CG lipid bonds between CG and AA simulations shows again a good overall agreement (see Figure 11) even though the tail order appears to be slightly underestimated in our CG model. Notice, however, that the reduction in order in the



**Figure 8.** Comparison of the total radial distribution functions  $g(r)$  between AA (red solid line) and CG (blue dashed line) simulations among six selected CG types. The total  $g(r)$  includes both intermolecular and intramolecular contributions. The figure for CH–E1 displays the worst-converging case with the largest difference of  $g(r)$  among all the CG types. Both the AA and CG results were computed from block average of a simulation of a bilayer with 288 lipids.

unsaturated sn-2 tail is captured very well. This shows that the CG force field is able to reproduce important membrane features corresponding to the lipid bond order, which matters for instance for bilayer phase transitions, in which this order changes significantly.

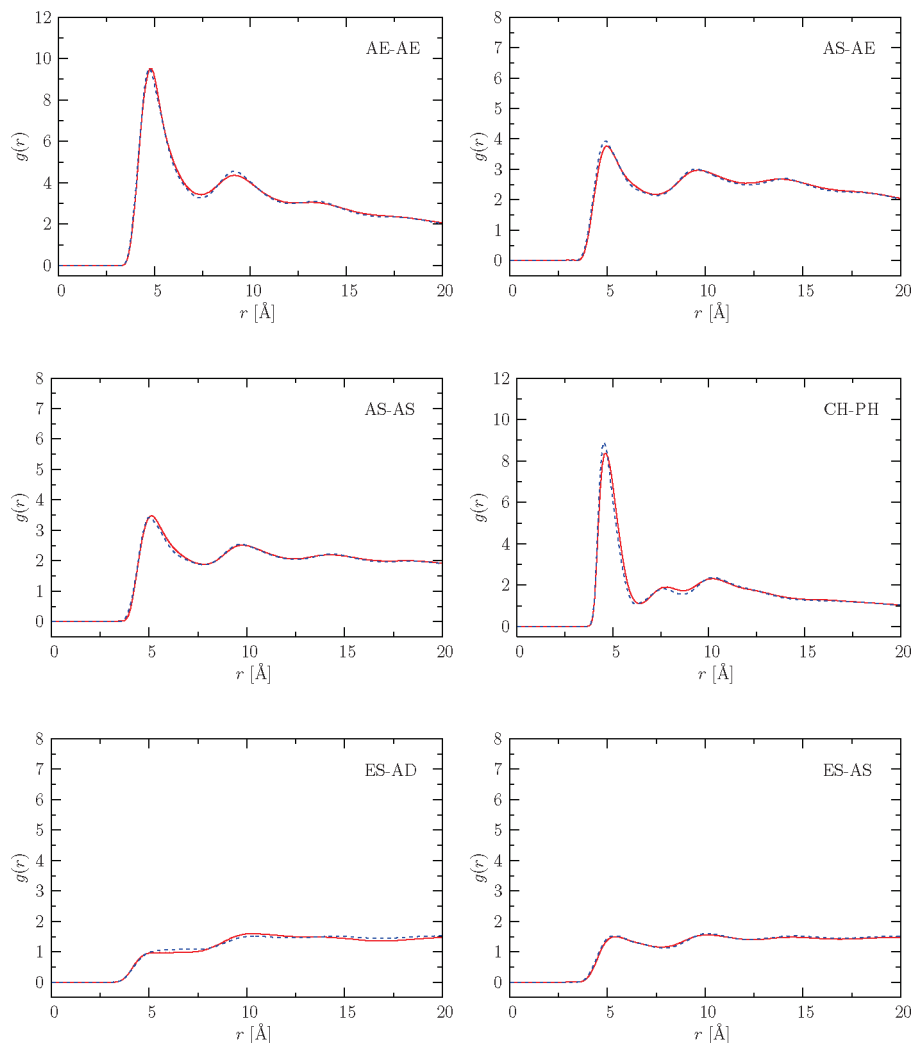
**3.2. Saturated Area Per Lipid.** The solvent-free CG simulations were performed both in the  $NAT$  and in the  $N\Sigma T$  ensembles. Figure 12 shows the pressure profile from an  $NAT$  CG simulation. The normal component of the pressure of the lipid membrane is constant (as required by mechanical stability) and in particular equal to zero in our CG case (as expected in the absence of explicit solvent). The surface tension of the CG membrane is thus obtained by eq 1 with  $P_n = 0$ , i.e., by the integral over the tangential stress distribution. A comparison of the stress profile  $P_t(z) - P_n$  between solvent-free CG simulations and AA simulations is shown in Figure 12. Although the overall curve shape is similar between CG and AA, the amplitude in the AA simulations is bigger than in the CG case. The reduced CG stress in the center region of the CG lipid membrane is due to the added cohesion recreating the hydrophobic effect. The smaller outer peaks within a distance of 20–30 Å from the center of the CG bilayer also stem from the absence of an explicit solvent; this region of our simulation where the transition to the solvent (or rather its absence) actually occurs naturally is most susceptible to such artifacts, and we

did not find a way to entirely eliminate the discrepancy in the stress profile without at the same time adversely affecting other observables, such as the RDFs.

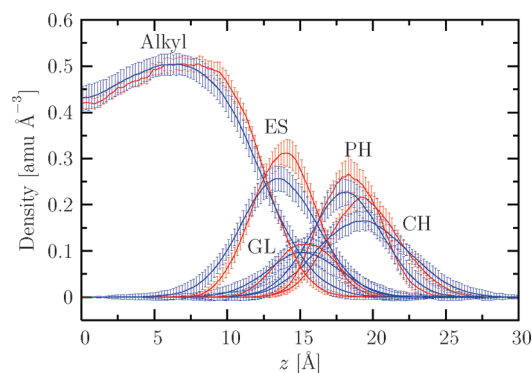
Notice that, with the CHARMM27 force field, the AA simulations in the  $NP_{zz}\Sigma T$  ensemble produced a much too small saturated area per lipid compared to experiment. If we write  $d_A$  for the hydrocarbon thickness in an  $NP_{zz}AT$  atomistic simulation (with  $A$  chosen as the correct saturated area from experiment<sup>9</sup>), and  $d_{\Sigma_0}$  for the (incorrect) hydrocarbon thickness from a zero-tension atomistic simulation, we can define the scaling parameter  $s = d_A/d_{\Sigma_0}$  and stretch  $P_t(z)$  into  $P_t^s(z) = P_t(z/s)$ . Both  $P_t(z) - P_n$  from an  $NP_{zz}AT$  atomistic simulation and  $P_t^s(z) - P_n$  from a zero-tension atomistic simulation are shown in Figure 12.

Figure 13 shows the time evolution of the projected (simulation box) area in a CG simulation. The average of the data after equilibration (past  $3 \times 10^3 \tau$  in Figure 13, corresponding to an uncorrected time of 20 ns, i.e., without including the speed-up factor associated with the reduced number of degrees of freedom) yields a saturated area per lipid  $A = 69.8 \pm 0.5 \text{ \AA}^2$  from an  $N\Sigma T$  CG simulation, consistent with the experimental value  $A = 68.3 \pm 1.5 \text{ \AA}^2$  per lipid.<sup>9</sup> This verifies that our solvent-free CG force field reproduces the saturated area per lipid from experiments.



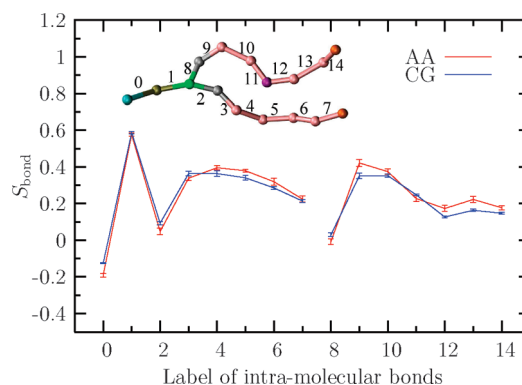


**Figure 9.** Comparison of the intermolecular radial distribution functions  $g_{\text{inter}}(r)$  between AA (red solid line) and CG (blue dashed line) simulations among six selected CG types. Both the AA and CG results were computed from block average of a simulation of a bilayer with 288 lipids. The results of AS–AS, AE–AE, and AS–AE show the efficiency of the Lorentz–Berthelot combining rules in our applications. Comparison of  $g_{\text{inter}}(r)$  of CH–PH and ES–AS with the  $g(r)$  in Figure 8 illustrates how strongly the bonded and nonbonded interactions are entangled.



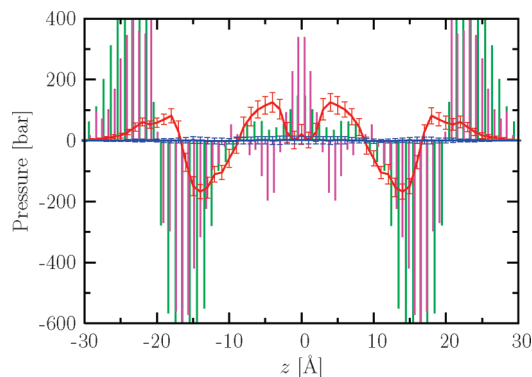
**Figure 10.** Comparison of the density profile between AA (red solid line) and CG (blue dashed line) simulations of a bilayer, where both AA and CG error bars were obtained from block averaging.

**3.3. Stability and Self-assembly.** To test the thermodynamic stability of a CG bilayer, we tested whether our CG lipids would self-assemble into this state. We thus followed the time evolution of a system that was started from 288 lipids randomly dispersed in a box of size  $(99.17 \text{ \AA})^3$  (see Figure 14), which means that the initial configuration was generated by choosing both the position and orientation of every lipid randomly (of course without mutual overlap). To avoid getting kinetically trapped

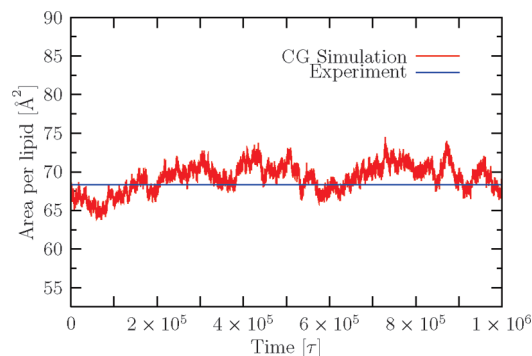


**Figure 11.** Comparison of the orientational order parameter  $P_2$  of intramolecular bonds between AA (red) and CG (blue) simulations. The bond numbers listed on the horizontal axis are indexed in the inset. Notice that the bonds 11 and 12 around the unsaturated CG group (and their subsequent tail bonds) show a significantly lower bond order than their corresponding partner bonds in the fully saturated tail.

in states which display very slow healing of large-scale bilayer defects, we used parallel tempering with replica exchanges between eight temperatures (279, 294.5, 310, 325.5, 341, 356.5, 372, and 387.5 K). A sequence of snapshots taken from the system at the temperature of interest,  $T = 310 \text{ K}$ , (see Figure



**Figure 12.** Comparison of the pressure profile between AA and CG simulations of a bilayer with 288 lipids. The value of  $P_t(z) - P_n$  for  $NP_{zz}$  AT and  $NP_{zz}\Sigma T$  ( $\Sigma = 0$ ) in AA simulations are shown with green and pink impulses, respectively. Both  $P_t(z) - P_n$  (red line) and  $P_n$  (blue line) for the CG simulations are shown as well.



**Figure 13.** Projected area per lipid as a function of time for an  $N\Sigma T$  ( $\Sigma = 0$ ) CG simulation of a bilayer with 288 lipids. The blue line shows the measured value from experiment.

14) proves that our lipids spontaneously aggregate into a membrane, or in other words, that the bilayer state is the *thermodynamically* stable conformation of our lipids. This is a nontrivial statement even if one considers that our force field was parametrized from the bilayer state, because there is no automatic way of guaranteeing that other states would not have an even lower free energy. However, given the bilayer-derived origin of our CG interactions, neither the thermodynamic properties of the initial nonbilayer phase nor the kinetics or thermodynamics of the transition into the bilayer equilibrium should be regarded as being quantitatively correct.

**3.4. Elastic Property. 1. Stretching Modulus.** In simulations, the zero-tension stretching modulus of a membrane,  $K_A$ , can be estimated from the thermal fluctuations of the membrane area:<sup>72</sup>

$$K_A = \frac{k_B T \langle A \rangle_{N\Sigma T}}{\langle A^2 \rangle_{N\Sigma T} - \langle A \rangle_{N\Sigma T}^2} \quad (18)$$

where  $A$  is the projected area of membrane and  $\Sigma$  is its surface tension, chosen to be zero in this case. From the projected area fluctuation at equilibrium (after  $3 \times 10^5 \tau$  in Figure 13), we obtained  $K_A = 108 \pm 20$  mN/m.

In experiments, the expansion modulus  $K_A$  of membranes was obtained from the measured relation between the apparent area,  $A$ , of vesicles and the applied tension:<sup>7,73</sup>

$$K_A = A_0 \frac{\partial \Sigma}{\partial A} \quad (19)$$

where  $A_0$  is the saturated area of the membrane at zero surface tension. Notice that close to the stress-free state the surface tension is not proportional to area dilation; instead, due to the effect of thermal shape undulations the logarithm of the tension becomes a linear function of area dilation, with a slope proportional to the elastic bending modulus  $\kappa$ . A Hookean stress-strain relation with a corresponding linear modulus  $K_{app}$  therefore only exists at slightly more elevated tension (above approximately 0.5 mN/m).<sup>7,73</sup> But since in simulations a bilayer often consists of only a few hundred lipids, with very few possibilities for shape undulations, eq 19 can be used to measure  $K_A$  even in the near-zero-tension regime (see the plot of surface tensor versus projected area in Figure 15). By fitting the data in Figure 15 to eq 19, we obtained  $K_A = 140 \pm 15$  mN/m. The value of  $K_A$  from “mechanical” measurements is larger than the value obtained from “thermal” measurements by about 30%, but within our error bars both values are not mutually inconsistent as they have been in a previous study.<sup>72</sup> Our value of  $K_A$  is somewhat smaller than typical experimental values for phospholipid bilayers (230–265 mN/m),<sup>7,73</sup> but not distressingly so.

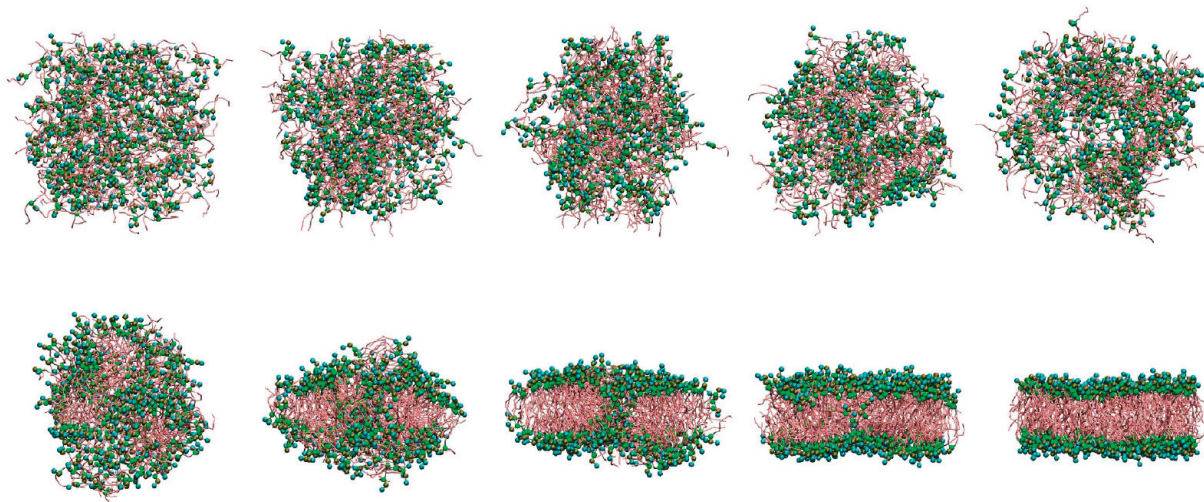
**2. Bending Modulus.** Evans and Rawicz proposed that the bending modulus  $\kappa$  can be obtained from the stretching modulus  $K_A$  through the relation  $\kappa = K_A h^2 / b$ , where  $h$  is the membrane thickness and  $b$  is a numerical elastic ratio. This relation was obtained by a scaling analysis based on the molecular theory of chain packing,<sup>3</sup> but it also follows directly just from simple continuum elasticity considerations.<sup>74</sup> The tricky bit is the numerical value of  $b$ . For a polymer-brush bilayer, where two monolayers are held together by hydrophobic interactions, the theoretical analysis<sup>75</sup> proposed  $b = 24$ . For an unbounded isotropic-elastic bilayer, where the two monolayers can slide freely against each other, the expected value is  $b = 48$ .<sup>27,75</sup> For a pair of isotropic layers that cannot slide with respect to each other,  $b = 12$ .<sup>75</sup> In experiments, the value of  $b = 24$  is typically used to infer the stretching modulus from the bending modulus.<sup>76</sup> Using this value, and the membrane hydrocarbon thickness  $h = 27.1$  Å, we find that the bending modulus of our CG POPC membrane at  $T = 310$  K is  $\kappa = (4.3 \pm 0.5) \times 10^{-20}$  J. Compared with experimental data  $\kappa = 8.5 \times 10^{-20}$  J,<sup>9,76</sup> our thus calculated bending modulus is about a factor 2 too small. But since the value of  $b$  is model-dependent, and also the precise value for the membrane thickness  $h$  is not completely obvious (but it enters quadratically), further checks are needed.

A more model-independent way to determine  $\kappa$  can be derived from measuring the membrane undulation spectrum. The elastic energy of a free-standing membrane, in the absence of external perturbation, can be written as<sup>77</sup>

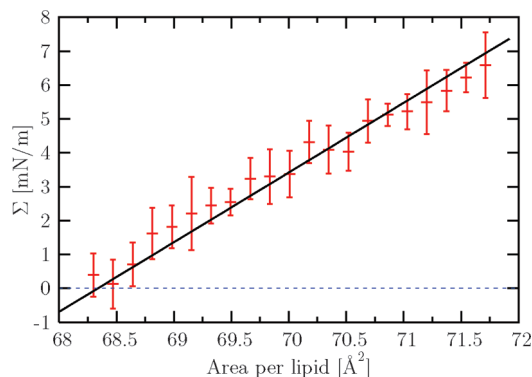
$$E = \frac{1}{2} \int_{A_p} d\vec{r} [\Sigma |\nabla h_{\vec{r}}|^2 + \kappa |\nabla^2 h_{\vec{r}}|^2] \quad (20)$$

where  $\kappa$  and  $\Sigma$  are the bending modulus and lateral tension respectively,  $A_p$  is the projected area of the membrane,  $h_{\vec{r}}$  describes the height of the membrane relative to some reference plane (i.e., “Monge gauge”). Expanding  $h_{\vec{r}}$  in Fourier modes

$$h_{\vec{r}} = \sum_{\vec{q}} h_{\vec{q}} e^{i\vec{q} \cdot \vec{r}} \quad (21)$$



**Figure 14.** Self-assembly of lipids into a bilayer starting from a random dispersion at a temperature  $T = 310$  K captured from a parallel tempering simulation. From top left to bottom right the snapshots taken at times  $t/\tau = 0, 600, 800, 950, 1000, 16\,000, 21\,000, 38\,000, 50\,000, 64\,000$ , respectively. Initially, the lipids aggregate quickly into a *locally* lamellar structure, presumably corresponding to a local minimum in the free energy (see the top five snapshots from  $t = 0$  to  $t = 1000\tau$ ). After this, it takes a very substantial time to eliminate kinetically trapped large-scale defects and arrive at the global minimum in free energy, which corresponds to a defect-free fluid bilayer (see the bottom snapshots from  $t = 16\,000\tau$  to  $t = 64\,000\tau$ ).



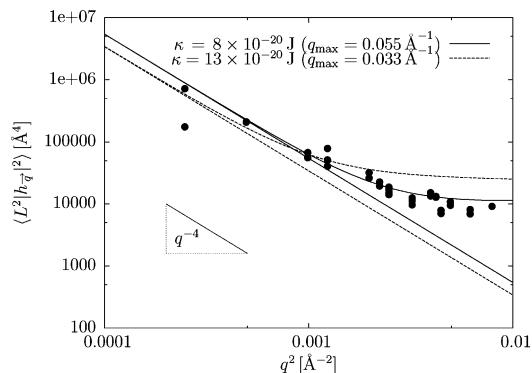
**Figure 15.** Bilayer tension  $\Sigma$  as a function of projected area per lipid from NVT CG simulations of a bilayer with 288 lipids. The solid line is a linear fit to the Hookean model  $\Sigma = K_A(A_{\text{lipid}} - A_{\text{lipid},0})/A_{\text{lipid},0}$ .

where  $\vec{q} = (2\pi/L)(n_x\vec{i} + n_y\vec{j})$ , and applying the equipartition theorem, we obtain the power spectrum of modes<sup>32,34,36,47</sup>

$$\langle L^2 |h_{\vec{q}}|^2 \rangle = \frac{k_B T}{\kappa q^4 + \Sigma q^2} \quad (22)$$

Here  $q$  is the magnitude of the vector  $\vec{q}$  and  $L$  is the side length of the square membrane. Notice that under zero lateral tension, the  $\Sigma q^2$  term vanishes. Figure 16 shows the undulation spectrum of a square membrane with 4608 lipids ( $\langle L \rangle \approx 40$  nm) which is sampled in the  $N(\Sigma=0)T$  ensemble for the simulation time of  $9.8 \times 10^6\tau$ . Even though we have simulated a fairly large membrane, there are still only a few modes which are in the asymptotic regime, and the lowest- $q$  mode might not yet be fully equilibrated. Hence, we need a cautious extrapolation procedure when fitting the spectrum. We decided to look at  $q^4 \langle L^2 |h_{\vec{q}}|^2 \rangle$  for low  $q$  and fit it to the functional form

$$q^4 \langle L^2 |h_{\vec{q}}|^2 \rangle = \frac{k_B T}{\kappa} + c_1 q^{c_2} \quad (23)$$

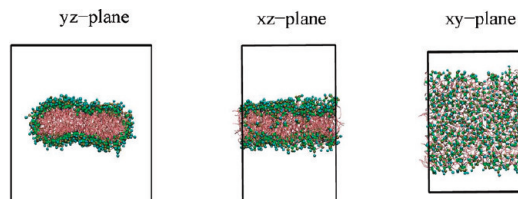


**Figure 16.** Undulation spectrum of a bilayer membrane with 4608 lipids sampled in  $N(\Sigma = 0)T$  ensemble for  $9.8 \times 10^6\tau$ . The fits of the data to the functional form eq 23 are shown with the upper limit of  $q$  chosen as  $q_{\text{max}} = 0.055 \text{ \AA}^{-1}$  (solid line) and  $q_{\text{max}} = 0.033 \text{ \AA}^{-1}$  (dash line). The fitted  $q^{-4}$ -only terms (without protrusion term) are explicitly shown as well (the straight lines in logarithm  $x$ - and  $y$ -coordinates).

see Figure 16. Notice that the  $c_1 q^{c_2}$  term ( $c_2 \approx 2$ ) corresponds to a protrusion contribution.<sup>78</sup> Depending on the upper limit of  $q$  chosen for the fit, we obtain values for  $\kappa$  in the range  $8 \times 10^{-20} \text{ J}$  ( $q_{\text{max}} = 0.055 \text{ \AA}^{-1}$ ,  $c_2 \rightarrow 2.1$ ) up to  $13 \times 10^{-20} \text{ J}$  ( $q_{\text{max}} = 0.033 \text{ \AA}^{-1}$ ,  $c_2 \rightarrow 1.9$ ). This value for  $\kappa$  agrees better with experiments than the estimate based on the stretching modulus.

**3.5. Line Tension and Rupture Tension.** In our CG model, the hydrophobic effect of the aqueous environment is implicitly accounted for by a phenomenological cohesion between non-headgroup beads, which is optimized to reproduce the neat bilayer state, but not an open edge. Hence, both the line tension and, as a consequence, the overall stability of the bilayer could be adversely affected. To test whether the membrane is overstabilized by the cohesion, we computed the line tension for CG membranes.

The line tension can be determined from a computation of the free energy as a function of pore size<sup>32,79,80</sup> or from the stress-strain relation of a bilayer.<sup>36,81</sup> Here we use a simpler method. By spanning a membrane across the  $x$ -direction of a periodic box which is not large enough to also span the  $y$ -direction, two stable linear open edges appear (see Figure 17).



**Figure 17.** Snapshots in equilibrium of a CG simulation of bilayer with 288 POPC lipids for computing line tension, viewed along  $x$ - (left panel),  $y$ - (center panel), and  $z$ -axes (right panel).

These edges will exert a force equal to twice the line tension along the  $x$ -direction. Notice that since the width of the membrane can relax by choosing the distance between the two open edges appropriately, no additional force due to surface tension exists. Hence, the line tension  $\gamma$  is given by the simple formula

$$\gamma = -\frac{1}{2}\langle\sigma_{xx}\rangle L_y L_z \quad (24)$$

where  $\sigma_{xx}$  is the  $xx$ -component of the stress tensor, and  $L_y$  and  $L_z$  are the side lengths of the simulation box in  $y$ - and  $z$ -directions, respectively. After warmup and an additional equilibration time of  $2 \times 10^5 \tau$ , averages were obtained by simulating up to  $10^6 \tau$  and measuring the stress  $\sigma_{xx}$  every  $2\tau$  (after which the stress autocorrelation has essentially decayed to zero). The measured value is  $\gamma = 29 \pm 12$  pN.

Typical experimental line tensions for a phosphatidylcholine (PC) bilayer membrane have been reported in the range of 6.5–30 pN.<sup>82–85</sup> All-atom simulations find  $(12 \pm 9)$ – $(35 \pm 10)$  pN for a DMPC membrane.<sup>86</sup> Our value is well within the range of both experiments and all-atom simulations of PC membranes. Moreover, with a simple theory,<sup>36</sup> the line tension  $\gamma$  can be translated into a rupture tension

$$\Sigma_{\text{rupture}} = 3 \left( \frac{\pi K_A \gamma^2}{2A_0} \right)^{1/3} \quad (25)$$

where  $A_0$  is the saturated area of the membrane. Using the value  $A_0 = (40 \text{ nm})^2$ , the area we used in the simulations where we determined the undulation spectrum, we find that the rupture tension is around 14 mN/m, showing again that we have not overstabilized the bilayer phase.

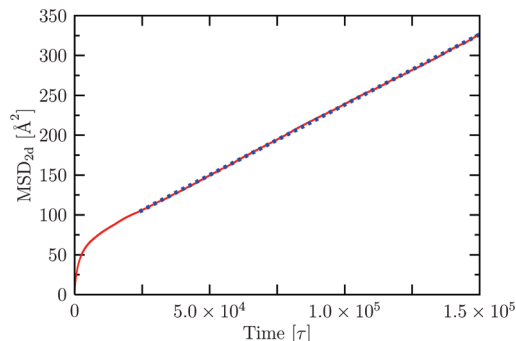
**3.6. Diffusion.** The dynamics of the solvent-free bilayer was examined by monitoring the in-plane lateral diffusion of lipids. The long-time diffusion coefficient is defined as

$$D = \lim_{t \rightarrow \infty} \left\{ \frac{1}{4} \frac{d}{dt} \langle \text{MSD}_{2d}(t) \rangle \right\} = \lim_{t \rightarrow \infty} \left\{ \frac{1}{4N} \frac{d}{dt} \left\langle \sum_{i=1}^N (\Delta \mathbf{r}_i)^2 \right\rangle \right\} \quad (26)$$

where

$$(\Delta \mathbf{r}_i)^2 = \{ [\mathbf{r}_i(t) - \mathbf{r}_{\text{cm}}(t)] - [\mathbf{r}_i(0) - \mathbf{r}_{\text{cm}}(0)] \}^2 \quad (27)$$

and  $\mathbf{r}_i(t)$  and  $\mathbf{r}_{\text{cm}}(t)$  are the (two-dimensional) in-plane coordinates of lipid  $i$  and the center of mass of the membrane at time  $t$ , respectively, and  $N$  is the number of lipids. Figure 18 shows



**Figure 18.** Two-dimensional mean-squared displacement from CG simulation (red solid line). The blue dashed line represent a linear fit from  $2.5 \times 10^4 \tau$  to  $1.5 \times 10^5 \tau$ .

the in-plane displacement of lipids as a function of time from our CG simulation. It can be seen from this figure that the CG lipid displacement curve exhibits a constant slope when  $t$  exceeds about  $2.5 \times 10^4 \tau$ . Fitting the slope yields the diffusion coefficient  $D = 70.86 \pm 0.77 \mu\text{m}^2/\text{s}$ , using the uncorrected time mapping  $\tau = 0.062$  ps. Compared with typical experimental values  $D = 1.7$ – $7.8 \mu\text{m}^2/\text{s}$ <sup>87</sup> and AA simulations  $D = 2.6$ – $3.9 \mu\text{m}^2/\text{s}$ ,<sup>87</sup> the effective lateral diffusive dynamics is faster by 1–2 orders of magnitude. This is of course neither surprising nor disconcerting, considering the reduced degrees of freedom in the CG model. Rather, this factor is a measure of the model-intrinsic speed-up between the “naive” time scale  $\zeta(m/\epsilon)^{1/2} = 0.062$  ps derived from natural length, mass, and energy, and the actual long-term time scale which belongs to the physical system our CG model is supposed to represent.

This speed-up is one of the two reasons why our CG model gains computational efficiency. The other reason is the fact that due to the different potentials, fewer degrees of freedom, and the absence of water even a specified bare time can be simulated faster on the CG level than on the AA level. Specifically, simulating a membrane with 288 lipids for 1 ns (uncorrected time in units of  $\tau$ ) on our cluster (Xeon E5430 2.66 GHz chips in parallel with infiniband connection) takes 8 CPUh with our CG model (using ESPReso) but 504 CPUh with a fully atomistic model (using NAMD). Combining both factors, we expect an overall speed-up in computational efficiency of approximately 3–4 orders of magnitude between our CG model and a comparable AA simulation. Of course, once one begins to look at bilayers which are curved and thus need to be embedded in substantially more solvent, this factor will be significantly bigger.

## 4. Conclusions

On the basis of an improved structure-matching methodology, we presented a solvent-free coarse-grained model for quantitative simulations of a POPC lipid membrane. When deriving the CG force field, data from both all-atom simulations and experiments were carefully chosen as a reference. The effective CG potentials were optimized through iterative loops, aimed to reproduce both RDFs and the saturated area per lipid of the bilayer.

The solvent-free CG force field drives a random lipid dispersion into a fluid bilayer via self-assembly at the experimental temperature. This shows that the bilayer phase is indeed the thermodynamically stable state. In particular, this also assures us that a bilayer will be stable under many kinds of structural perturbations (open edges, inserted or adsorbed proteins, strong curvature, fusion intermediates, etc.), meaning

that one can venture to study such situations without having to fear that the membrane, due to its exotic solvent free representation, will fall apart if not meticulously kept in a pristine bilayer conformation. At the same time, the reproduction of a physically meaningful line tension shows that our phenomenological cohesion potential does not overstabilize the bilayer state.

To examine the performance of the solvent-free CG force field, we explored many physical properties of the bilayer membrane, including the cross-bilayer density profile, the  $P_2$  order parameter of each lipid bond, the pressure profile, and the stretching and bending moduli. Comparison of these properties with experiments and fully atomistic simulations shows a quantitative or semiquantitative consistency, with the notable caveat that elastic properties seem to be slightly too soft. The speed-up factor derived from the reduced number of degrees of freedom is approximately 1–2 orders of magnitude, as determined from matching diffusion coefficients. Combined with the more efficient integration of the equations of motion this results in an overall speed-up of 3–4 orders of magnitude compared to AA simulations while still keeping a large amount of chemical specificity. This will be highly valuable for any study that involves mesoscale membranes, or even their interactions with equally coarse-grained proteins in implicit solvent.<sup>88–92</sup>

**Acknowledgment.** We thank J. F. Nagle and S. Tristram-Nagle for their useful comments on the force fields of fully atomistic simulations and for sharing their experimental data of electron densities of fully hydrated POPC bilayers, and T. Bereau for many helpful discussions on the ESPResSo package. The work was financially supported by an NIH grant (P01 AG032131). The computations were supported in part by the National Science Foundation through TeraGrid resources provided by the Pittsburgh Supercomputing Center and the National Center for Supercomputing Applications (NCSA).

**Supporting Information Available:** Comparison of total RDFs, intermolecular RDFs, and intramolecular RDFs among all CG types. The CG force field is available as a downloadable package “cgforcefield” (compressed and published online as “jp102543j\_si\_003.zip”). This material is available free of charge via the Internet at <http://pubs.acs.org>.

## References and Notes

- Alberts, B.; Johnson, A.; Lewis, L.; Raff, M.; Roberts, K.; Walter, P. *Mol. Biol. Cell*, 4th ed.; Garland Science: New York, 2002.
- Karp, G. *Cell and Molecular Biology: Concepts and Experiments*, 5th ed.; Wiley: Hoboken, NJ, 2007.
- Lipowsky, R.; Sackmann, E., Eds. *Structure and Dynamics of Membranes From Cells to Vesicles*, 1st ed.; Elsevier Science B.V.: Amsterdam, The Netherlands, 1995; Vol. 1.
- Yeagle, P. *The Structure of Biological Membranes*, 2nd ed.; CRC Press LLC: Boca Raton, FL, 2005.
- Nagle, J. F.; Tristram-Nagle, S. *Biochim. Biophys. Acta* **2000**, *1469* (3), 159–195.
- Petrache, H.; Dodd, S.; Brown, M. *Biophys. J.* **2000**, *79* (6), 3172–3192.
- Evans, E.; Heinrich, V.; Ludwig, F.; Rawicz, W. *Biophys. J.* **2003**, *85* (4), 2342–2350.
- Veatch, S. L.; Polozov, I. V.; Gawrisch, K.; Keller, S. L. *Biophys. J.* **2004**, *86* (5), 2910–2922.
- Kučerka, N.; Tristram-Nagle, S.; Nagle, J. J. *Membr. Biol.* **2006**, *208* (3), 193–202.
- McGillivray, D. J.; Valincius, G.; Vanderah, D. J.; Febo-Ayala, W.; Woodward, J. T.; Heinrich, F.; Kasianowicz, J. J.; Lösche, M. *Biointerphases* **2007**, *2* (1), 21–33.
- Miller, C. E.; Majewski, J.; Watkins, E. B.; Mulder, D. J.; Gog, T.; Kuhl, T. L. *Phys. Rev. Lett.* **2008**, *100* (5), 058103.
- Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed.; Academic Press: Boston, 2002.
- Allen, M.; Tildesley, D. *Computer Simulation of Liquids*, 2nd ed.; Oxford University Press: New York, 2002.
- Binder, K. *Monte Carlo and Molecular Dynamics Simulations in Polymer Science*, 1st ed.; Oxford University Press: New York, 1995.
- Ferrario, M.; Ciccotti, G.; Binder, K., Eds. *Computer simulations in condensed matter: from materials to chemical biology, Lecture Notes in Physics*, 1st ed.; Springer-Verlag: Berlin, 2006.
- Berendsen, H. *Simulating the Physical World: Hierarchical Modeling from Quantum Mechanics to Fluid Dynamics*, 1st ed.; Cambridge University Press: Cambridge, UK, 2007.
- Tieleman, D.; Marrink, S.-J. *J. Am. Chem. Soc.* **2006**, *128* (38), 12462–12467.
- Faraldo-Gómez, J.; Forrest, L.; Baaden, M.; Bond, P.; Domene, C.; Patargias, G.; Cuthbertson, J.; Sansom, M. *Proteins: Struct., Funct., Bioinf.* **2004**, *57* (4), 783–791.
- Saiz, L.; Bandyopadhyay, S.; Klein, M. *Biosci. Rep.* **2002**, *22* (2), 151–173.
- Ash, W.; Zlomislic, M.; Oloo, E.; Tieleman, D. *Biochim. Biophys. Acta, Biomembr.* **2004**, *1666* (1–2), 158–189.
- Feller, S., Ed. *Computational modeling of membrane bilayers*, 1st ed.; Academic: London, 2008.
- Venturoli, M.; Sperotto, M.; Kranenburg, M.; Smit, B. *Phys. Rep.—Rev. Sect. Phys. Lett.* **2006**, *437* (1–2), 1–54.
- Müller, M.; Katsov, K.; Schick, M. *Phys. Rep.—Rev. Sect. Phys. Lett.* **2006**, *434* (5–6), 113–176.
- Shillcock, J.; Lipowsky, R. *J. Phys.: Condens. Matter* **2006**, *18* (28), S1191–S1219.
- Deserno, M. *Macromol. Rapid Commun.* **2009**, *30* (9–10), 752–771.
- Goetz, R.; Lipowsky, R. *J. Chem. Phys.* **1998**, *108* (17), 7397–7409.
- Goetz, R.; Gompper, G.; Lipowsky, R. *Phys. Rev. Lett.* **1999**, *82* (1), 221–224.
- Venturoli, M.; Smit, B. *Phys. Chem. Commun.* **1999**, *10*, 45–49.
- Drouffe, J.; Maggs, A.; Leibler, S. *Science* **1991**, *254* (5036), 1353–1356.
- Noguchi, H.; Takasu, M. *Phys. Rev. E* **2001**, *64* (4), 041913.
- Noguchi, H.; Takasu, M. *J. Chem. Phys.* **2001**, *115* (20), 9547–9551.
- Farago, O. *J. Chem. Phys.* **2003**, *119* (1), 596–605.
- Brannigan, G.; Philips, P.; Brown, F. *Phys. Rev. E* **2005**, *72* (1), 011915.
- Wang, Z.-J.; Frenkel, D. *J. Chem. Phys.* **2005**, *122* (23), 234711.
- Cooke, I.; Kremer, K.; Deserno, M. *Phys. Rev. E* **2005**, *72* (1), 011506.
- Cooke, I.; Deserno, M. *J. Chem. Phys.* **2005**, *123* (22), 224710.
- Ilyia, G.; Deserno, M. *Biophys. J.* **2008**, *95* (9), 4163–4173.
- Rzepele, A. J.; Sengupta, D.; Goga, N.; Marrink, S. J. *Faraday Discuss.* **2010**, *144*, 431–443.
- Shelley, J.; Shelley, M.; Reeder, R.; Bandyopadhyay, S.; Klein, M. *J. Phys. Chem. B* **2001**, *105* (19), 4464–4470.
- Lyubartsev, A. *Eur. Biophys. J. Biophys. Lett.* **2005**, *35* (1), 53–61.
- Rekviq, L.; Kranenburg, M.; Vreede, J.; Hafskjold, B.; Smit, B. *Langmuir* **2003**, *19* (20), 8195–8205.
- Marrink, S.-J.; de Vries, A.; Mark, A. *J. Phys. Chem. B* **2004**, *108* (2), 750–760.
- Marrink, S.-J.; Risselada, H.; Yefimov, S.; Tieleman, D.; de Vries, A. *J. Phys. Chem. B* **2007**, *111* (27), 7812–7824.
- Monticelli, L.; Kandasamy, S.; Periole, X.; Larson, R.; Tieleman, D.; Marrink, S.-J. *J. Chem. Theory Comput.* **2008**, *4* (5), 819–834.
- Izvekov, S.; Voth, G. J. *J. Phys. Chem. B* **2005**, *109* (7), 2469–2473.
- Noid, W.; Chu, J.-W.; Ayton, G.; Voth, G. J. *J. Phys. Chem. B* **2007**, *111* (16), 4116–4127.
- Brannigan, G.; Brown, F. *J. Chem. Phys.* **2004**, *120* (2), 1059–1071.
- Murtola, T.; Falck, E.; Karttunen, M.; Vattulainen, I. *J. Chem. Phys.* **2007**, *126* (7), 075101.
- Izvekov, S.; Voth, G. J. *J. Phys. Chem. B* **2009**, *113* (13), 4443–4455.
- Phillips, J.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26* (16), 1781–1802.
- Feller, S.; MacKerell, A. J. *J. Phys. Chem. B* **2000**, *104* (31), 7510–7515.
- MacKerell, A.; Banavali, N.; Foloppe, N. *Biopolymers* **2000**, *56* (4), 257–265.
- Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14* (1), 33–39.
- Essmann, U.; Perera, L.; Berkowitz, M.; Darden, T.; Lee, H.; Pedersen, L. *J. Chem. Phys.* **1995**, *103* (19), 8577–8593.
- Benz, R.; Castro-Roman, F.; Tobias, D.; White, S. *Biophys. J.* **2005**, *88* (2), 805–817.

- (56) Armen, R.; Uitto, O.; Feller, S. *Biophys. J.* **1998**, *75* (2), 734–744.
- (57) Lindahl, E.; Edholm, O. *J. Chem. Phys.* **2000**, *113* (9), 3882–3893.
- (58) Sonne, J.; Hansen, F.; Peters, G. *J. Chem. Phys.* **2005**, *122* (12), 124903.
- (59) Rowlinson, J.; Widom, B. *Molecular Theory of Capillarity*, 1st ed.; Clarendon Press: Oxford, UK, 1982.
- (60) Petrache, H.; Salmon, A.; Brown, M. *J. Am. Chem. Soc.* **2001**, *123* (50), 12611–12622.
- (61) Coolbear, K.; Keough, K. *Biochim. Biophys. Acta* **1983**, *732* (3), 531–540.
- (62) Peter, C.; Delle Site, L.; Kremer, K. *Soft Matter* **2008**, *4* (4), 859–869.
- (63) Heinz, H.; Paul, W.; Binder, K. *Phys. Rev. E* **2005**, *72* (6), 066704.
- (64) Grest, G.; Kremer, K. *Phys. Rev. A* **1986**, *33* (5), 3628–3631.
- (65) Kolb, A.; Dünweg, B. *J. Chem. Phys.* **1999**, *111* (10), 4453–4459.
- (66) Limbach, H.-J.; Arnold, A.; Mann, B.; Holm, C. *Comput. Phys. Commun.* **2006**, *174* (8), 704–727.
- (67) Peter, C.; Kremer, K. *Soft Matter* **2009**, *5* (22), 4357–4366.
- (68) Harismiadis, V.; Koutras, N.; Tassios, D.; Panagiotopoulos, A. *Fluid Phase Equilib.* **1991**, *65*, 1–18.
- (69) Soper, A. K. *Chem. Phys.* **1996**, *202* (2–3), 295–306.
- (70) Reith, D.; Pütz, M.; Müller-Plathe, F. *J. Comput. Chem.* **2003**, *24* (13), 1624–1636.
- (71) De Boer, C. A. *A Practical Guide to Splines*, 1st ed.; Springer-Verlag: New York, 2001.
- (72) Feller, S.; Pastor, R. *J. Chem. Phys.* **1999**, *111* (3), 1281–1287.
- (73) Evans, E.; Rawicz, W. *Phys. Rev. Lett.* **1990**, *64* (17), 2094–2097.
- (74) Helfrich, W. *Elasticity and thermal undulations of fluid films of amphiphiles, in Liquids at Interfaces*; Charvolin, J., Joanny, J. F., Zinn-Justin, J., Eds.; Elsevier Science Publishers B.V.: Amsterdam, The Netherlands, 1990.
- (75) Rawicz, W.; Olbrich, K.; McIntosh, T.; Needham, D.; Evans, E. *Biophys. J.* **2000**, *79* (1), 328–339.
- (76) Pan, J.; Tristram-Nagle, S.; Kučerka, N.; Nagle, J. *Biophys. J.* **2008**, *94* (1), 117–124.
- (77) Safran, S. *Statistical thermodynamics of surfaces, interfaces and membranes*; Westview Press: Boulder, CO, 1994.
- (78) Lipowsky, R.; Grotehans, S. *Biophys. Chem.* **1994**, *49* (1), 27–37.
- (79) Wang, Z.-J.; Frenkel, D. *J. Chem. Phys.* **2005**, *123* (15), 154701.
- (80) Wohler, J.; den Otter, W.; Edholm, O.; Briels, W. *J. Chem. Phys.* **2006**, *124* (15), 154905.
- (81) Tolpekina, T.; den Otter, W.; Briels, W. *J. Chem. Phys.* **2004**, *121* (23), 12060–12066.
- (82) Taupin, C.; Dvolaitzky, M.; Sauterey, C. *Biochemistry* **1975**, *14* (21), 4771–4775.
- (83) Genco, I.; Gliozzi, A.; Relini, A.; Robello, M.; Scalas, E. *Biochim. Biophys. Acta* **1993**, *1149* (1), 10–18.
- (84) Zhelev, D.; Needham, D. *Biochim. Biophys. Acta* **1993**, *1147* (1), 89–104.
- (85) Karatekin, E.; Sandre, O.; Guitouni, H.; Borghi, N.; Puech, P.; Brochard-Wyart, F. *Biophys. J.* **2003**, *84* (3), 1734–1749.
- (86) Jiang, F.; Bouret, Y.; Kindt, J. *Biophys. J.* **2004**, *87* (1), 182–192.
- (87) Bockmann, R.; Hac, A.; Heimburg, T.; Grubmüller, H. *Biophys. J.* **2003**, *85* (3), 1647–1655.
- (88) Irbäck, A.; Sjunnesson, F.; Wallin, S. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97* (25), 13614–13618.
- (89) Head-Gordon, T.; Brown, S. *Curr. Opin. Struct. Biol.* **2003**, *13* (2), 160–167.
- (90) Ding, F.; Borreguero, J.; Buldyrey, S.; Stanley, H.; Dokholyan, N. *Proteins: Struct., Funct., Genet.* **2003**, *53* (2), 220–228.
- (91) Derreumaux, P.; Mousseau, N. *J. Chem. Phys.* **2007**, *126* (2), 025101.
- (92) Bereau, T.; Deserno, M. *J. Chem. Phys.* **2009**, *130* (23), 235106.

JP102543J