

# SCIENTIFIC REPORTS



OPEN

## Effects of Coarse Graining and Saturation of Hydrocarbon Chains on Structure and Dynamics of Simulated Lipid Molecules

Pavel Buslaev<sup>1</sup> & Ivan Gushchin<sup>1,2</sup> 

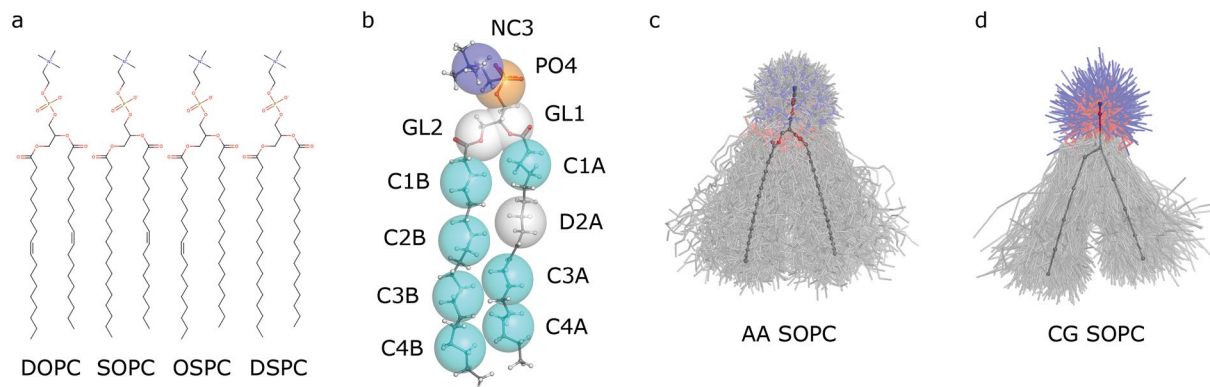
Molecular dynamics simulations are used extensively to study the processes on biological membranes. The simulations can be conducted at different levels of resolution: all atom (AA), where all atomistic details are provided; united atom (UA), where hydrogen atoms are treated inseparably of corresponding heavy atoms; and coarse grained (CG), where atoms are grouped into larger particles. Here, we study the behavior of model bilayers consisting of saturated and unsaturated lipids DOPC, SOPC, OSPC and DSPC in simulations performed using all atom CHARMM36 and coarse grained Martini force fields. Using principal components analysis, we show that the structural and dynamical properties of the lipids are similar, both in AA and CG simulations, although the unsaturated molecules are more dynamic and favor more extended conformations. We find that CG simulations capture 75 to 100% of the major collective motions, overestimate short range ordering, result in more flexible molecules and 5–7 fold faster sampling. We expect that the results reported here will be useful for comprehensive quantitative comparisons of simulations conducted at different resolution levels and for further development and improvement of CG force fields.

Nowadays, molecular dynamics (MD) simulations became indispensable in biological research. They are used to study the properties of solvents, small molecules, polymers, oligonucleotides and proteins. The simulations are especially fruitful in the field of membrane biology where they can be used to characterize the behavior of various lipidic phases: micelles, bilayers, vesicles, cubic and others, and to model physiological processes, such as membrane pore formation or membrane fusion and fission, and to probe the effects of additives<sup>1–4</sup>.

One of prerequisites of a successful simulation is employment of a correct set of simulation parameters, known as a force field (FF). Currently, several FFs for simulations of lipids are available<sup>5–20</sup>. Since the parameterization is often based on different principles, inevitably there are differences in the results of the simulations, which however become smaller in later versions of popular FFs<sup>21–29</sup>. Among the major differences between the FFs is the resolution level at which the lipids are simulated. First, there are all atom (AA) FFs, such as CHARMM36<sup>3</sup>, Lipid14<sup>6</sup>, OPLS-AA<sup>7,8</sup> and Stockholm lipids<sup>9</sup>. These FFs faithfully represent both structure and dynamics of lipid molecules, but require significant computational power for simulations of complex phenomena. Next, it is possible to treat some or all of the hydrogen atoms inseparably from the heavy atoms to which they are covalently bound, and consequently reduce the number of interacting particles and computational cost. This so-called united atom (UA) approach was implemented in, for example, Berger<sup>10</sup>, GROMOS<sup>11–13</sup>, CHARMM36-UA<sup>14</sup> and other FFs<sup>15,16</sup>, and results in 1.5–2x speedup of the calculations.

Unfortunately, many relevant physiological processes occur on spatiotemporal scales currently unreachable for atomistic FFs, and sampling is a problem in MD simulations<sup>30</sup>. Therefore, a set of even simpler coarse grained (CG) models has been developed. In such models, molecule's atoms are grouped into beads, so that the general geometry and connectivity are retained<sup>31–35</sup>. These models were used successfully to simulate proteins<sup>36,37</sup> as well as lipids. There are several CG FFs for lipid molecules that illustrate the force matching approach<sup>38–40</sup>, reproduce characteristic time scales from AA simulations and experiments<sup>41,42</sup> and concentrate on the phase behavior of lipid membranes<sup>43,44</sup> or interactions of proteins with lipids<sup>45</sup>. Perhaps the most popular lipid CG FF

<sup>1</sup>Moscow Institute of Physics and Technology, 141700, Dolgoprudny, Russia. <sup>2</sup>Institute of Complex Systems (ICS), ICS-6: Structural Biochemistry, Research Centre Jülich, 52425, Jülich, Germany. Correspondence and requests for materials should be addressed to P.B. (email: [pbuslaev@phystech.edu](mailto:pbuslaev@phystech.edu)) or I.G. (email: [ivan.gushchin@phystech.edu](mailto:ivan.gushchin@phystech.edu))



**Figure 1.** Structures of the studied molecules. **(a)** Chemical structures of DOPC (1,2-dioleoyl-*sn*-glycero-3-phosphocholine, 18:1c9 PC), SOPC (1-stearoyl-2-oleoyl-*sn*-glycero-3-phosphocholine, 18:0/18:1c9 PC), OSPC (1-oleoyl-2-stearoyl-*sn*-glycero-3-phosphocholine, 18:1c9/18:0 PC) and DSPC (1,2-distearoyl-*sn*-glycero-3-phosphocholine, 18:0 PC), respectively. **(b)** Martini mapping between all-atom and coarse grained structures for SOPC lipid molecule. Mappings for DOPC, SOPC, OSPC and DSPC headgroups and oleoyl and stearoyl chains are identical. **(c,d)** Samples of SOPC conformations observed in the AA and CG molecular dynamics simulations. 500 structure snapshots from each simulation are shown. The structures are aligned to the averaged structure.

is Martini<sup>17–19</sup>, now expanded to proteins<sup>46,47</sup>, DNA<sup>48</sup> and carbohydrates<sup>49</sup>. With Martini, it is possible to study extremely large and diverse systems such as plasma or thylakoid membranes<sup>50,51</sup>, cellulose microfibrils<sup>52</sup>, and many others.

CG FFs are very useful for extended simulations, and largely capture the major structural and dynamical features of atomistic FFs<sup>53–56</sup>. Inevitably, there are also some limitations. The short-range ordering is overestimated<sup>53,54</sup>. The lipid-lipid interactions in Martini are too weak, water-lipid repulsion is overestimated, and the temperature dependencies of thermodynamic quantities are weaker compared to AA FFs<sup>53</sup>. The dynamics is altered: convergence of configurational entropy and full sampling of internal motions of lipid tails in CG simulations can be reached faster than in atomistic simulations<sup>55,56</sup>. Dynamics of water particles and peptides in Martini simulations was found to be ~4 times faster than that in the AA and UA simulations<sup>19,57,58</sup>. Finally, uneven acceleration of dynamic processes was also observed in other CG simulations, and thus the characteristic time scales were deemed unrealistic<sup>59</sup>.

One of the defining factors of lipid molecule properties is the saturation of its hydrocarbon chains. The model lipids with different combinations of saturated and mono-unsaturated hydrocarbon chains are 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (18:1c9 PC, DOPC), 1-stearoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (18:0/18:1c9 PC, SOPC), 1-oleoyl-2-stearoyl-*sn*-glycero-3-phosphocholine (18:1c9/18:0 PC, OSPC) and 1,2-distearoyl-*sn*-glycero-3-phosphocholine (18:0 PC, DSPC) (Fig. 1a). Mono-unsaturated chains with *cis* double bonds have higher propensity to be disordered than the saturated ones, and consequently the more saturated chains such a lipid has, the higher is its melting temperature (256, 279, 282 and 328 K for DOPC, SOPC, OSPC and DSPC, respectively<sup>60</sup>), and the lower is its area per lipid (~73, ~70, ~70 and 67 Å<sup>2</sup> at for DOPC, SOPC, OSPC and DSPC, respectively, at 338 K<sup>61–70</sup>). These effects can be successfully reproduced in MD simulations of lipid bilayers<sup>63,64,71</sup>.

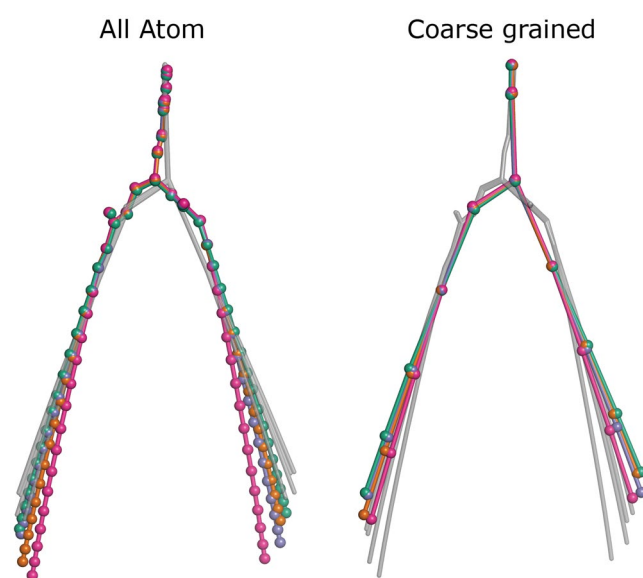
Recently, we applied the principal components analysis (PCA)<sup>72</sup> to compare several AA and UA lipid forcefields<sup>28</sup>. Here, we demonstrate how the PCA approach can be utilized to study the influence of coarse graining and saturation of hydrocarbon chains on structure and dynamics of lipid molecules in molecular dynamics simulations. We compare the properties of DOPC, SOPC, OSPC and DSPC molecules in atomistic and CG simulations at the single molecule level and describe the similarities and differences between the simulations. Most of the simulations are conducted at 338 K, above the DSPC melting temperature. However, this is an elevated temperature for most of the biological systems, and many of the experiments and simulations are performed at or below 310 K. Consequently, we probe the behavior of DOPC also at 323 K and 310 K, and analyze the influence of the simulation temperature on the lipid's conformations and dynamics.

## Results and Discussion

**Performed simulations.** We performed molecular dynamics simulations of lipids DOPC, SOPC, OSPC, and DSPC that have different combinations of monounsaturated (18:1c9) oleoyl and saturated (18:0) stearoyl chains at positions *sn*-1 and *sn*-2 (Fig. 1a), at two different levels of resolution: AA and CG, using CHARMM36<sup>5</sup> and Martini<sup>17–19</sup> force fields, respectively. Details of the simulations are presented in Table 1. Simulations AA1-3 and CG1-3 probe the effects of temperature on dynamics of individual lipid molecules, whereas AA1,4-6 and CG1,4-6 probe the similarities and differences between the studied lipids. Chemical structures, observed conformations and Martini mapping between the AA and CG models are presented in Fig. 1. Since macroscopic properties of DOPC, SOPC, OSPC and DSPC in atomistic and CG simulations have been described elsewhere<sup>5,9,18,65,68–70</sup>, we focus on the properties of individual molecules.

Simulation	Lipid	Temperature, K	No. of lipids	No. of waters *	Force field	Water model	Duration, ns
AA1	DOPC	338	128	5120	CHARMM36	CHARMM TIP3P	200
AA2	DOPC	323	128	5120	CHARMM36	CHARMM TIP3P	200
AA3	DOPC	310	128	5120	CHARMM36	CHARMM TIP3P	200
AA4	SOPC	338	128	5120	CHARMM36	CHARMM TIP3P	200
AA5	OSPC	338	128	5120	CHARMM36	CHARMM TIP3P	200
AA6	DSPC	338	128	5120	CHARMM36	CHARMM TIP3P	200
CG1	DOPC	338	128	1276	Martini	Polarizable Martini	200
CG2	DOPC	323	128	1276	Martini	Polarizable Martini	200
CG3	DOPC	310	128	1276	Martini	Polarizable Martini	200
CG4	SOPC	338	128	1276	Martini	Polarizable Martini	200
CG5	OSPC	338	128	1276	Martini	Polarizable Martini	200
CG6	DSPC	338	128	1276	Martini	Polarizable Martini	200

**Table 1.** Overview of the simulations presented in this study. \*For Martini simulations, the number of polarizable water particles. Each Martini water particle corresponds to 4 water molecules.



**Figure 2.** Influence of hydrocarbon chain saturation on the average structure of lipid molecule. DOPC is shown in maroon, SOPC is in orange, OSPC is in violet and DSPC is in magenta. Average structures at another level of resolution are shown in grey. While the headgroup structures are well conserved, the positions of hydrophobic chains differ. Oleoyl (18:1c9) chains occupy the outermost position and stearoyl (18:0) chains are the innermost.

**Average structures.** First, we calculated the average atomic positions (Fig. 2). Notably, both in AA and CG simulations, the unsaturated acyl chains occupy the outermost position while the saturated acyl chains occupy the innermost one (Fig. 2). This is most probably due to the fact that the unsaturated hydrocarbon chains are less ordered<sup>63</sup> and deviate from the normal to the bilayer. As for the individual atoms, in the AA simulations, positions of DOPC oleoyl chains are almost identical to the positions of SOPC and OSPC oleoyl chains. Similarly, positions of DSPC stearoyl chains are almost identical to those of SOPC and OSPC stearoyl chains. The average CG structures correspond to the average AA structures, in accordance with the Martini atoms-to-beads mapping, and reveal the same trends.

**Principal components analysis.** To gain further insight into the effects of coarse graining and saturation of hydrocarbon chains, we studied the lipid conformations observed in the obtained trajectories using principal components analysis<sup>28,72</sup>. The analysis results in a linear transformation of the conformational space via identification of new orthogonal basis vectors, called principal components (PCs), such that the variance of the input data along PC1 is the greatest possible, the variance along PC2 is the second greatest, etc. Each of the principal components defines a certain collective motion of the molecule's atoms.

Similarly to our previous analysis of AA and UA DOPC force fields, we observed that the covariance matrices of atomic displacements are dominated by few large-amplitude motions (Figure S2)<sup>28</sup>. Four major components account for ~50% of the structural variation, and 90% of the structural variation is covered by 14 components

	AA DOPC, 338 K	AA SOPC, 338 K	AA OSPC, 338 K	AA DSPC, 338 K	CG DOPC, 338 K	CG SOPC, 338 K	CG OSPC, 338 K	CG DSPC, 338 K
AA DOPC, 338 K	1.000	0.997	0.997	0.992	0.956	0.952	0.954	0.947
AA SOPC, 338 K	0.997	1.000	0.992	0.994	0.957	0.958	0.953	0.952
AA OSPC, 338 K	0.997	0.992	1.000	0.995	0.956	0.950	0.957	0.949
AA DSPC, 338 K	0.992	0.994	0.995	1.000	0.958	0.957	0.958	0.956
CG DOPC, 338 K	0.956	0.957	0.956	0.958	1.000	0.996	0.996	0.992
CG SOPC, 338 K	0.952	0.958	0.950	0.957	0.996	1.000	0.990	0.995
CG OSPC, 338 K	0.954	0.953	0.957	0.958	0.996	0.990	1.000	0.995
CG DSPC, 338 K	0.947	0.952	0.949	0.956	0.992	0.995	0.995	1.000

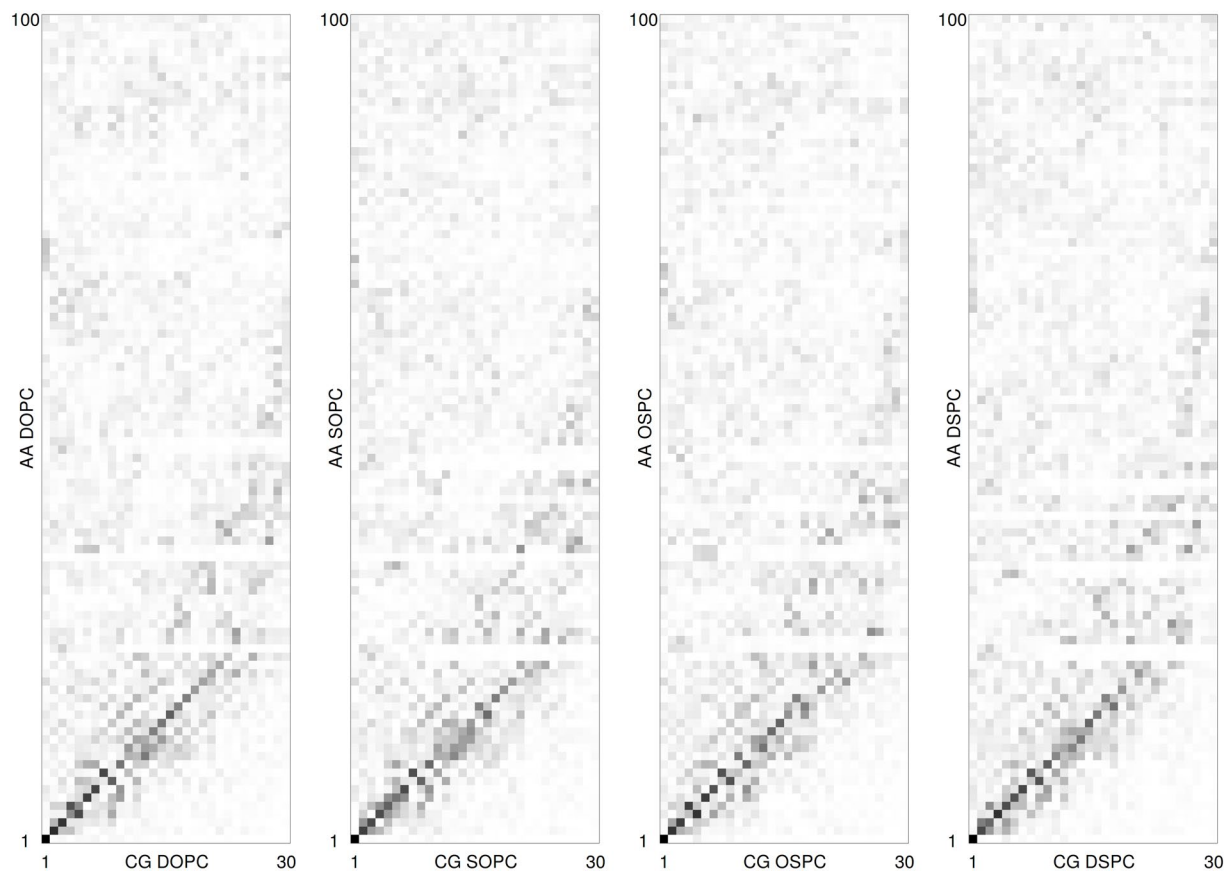
**Figure 3.** Pearson correlation coefficients between the covariance matrices of atomic displacements in respective simulations<sup>28</sup> (please find the description in Methods). The correlation coefficients are color-coded to highlight the similarity and dissimilarity between the simulations.

for the AA simulations, and by 11 components for the CG simulations. This slightly lower dimensionality of the CG configurational space is not unexpected<sup>55,56</sup>. As for the dependence on the hydrophobic chain saturation, the eigenvalue distributions are similar for the four lipids, although there is a clear trend that for more saturated hydrophobic chains the contribution of the major principal components becomes higher (Figure S2).

In all of the simulations the first, largest-amplitude principal component (PC1) corresponds to the scissoring motion of the hydrophobic tails (Figure S3). The others, lower-amplitude components, are less collective<sup>28</sup> and account for various bending and twisting motions of lipid molecules (Figure S3). Whereas in the atomistic simulations the principal components almost always include displacements both in the plane of the scissoring PC1 motion and perpendicular to it, in CG simulations the displacements are almost always either completely in plane or completely perpendicular to it (Figure S3).

**Comparison of the simulations.** Direct comparison of the obtained AA and CG trajectories to each other using PCA is problematic since the AA and CG lipid representations have different number of particles. This can be overcome in two ways. First, for most of the analyses we convert the AA trajectories into Martini CG representation. Second, for comparisons of AA collective motions with CG collective motions, we conduct PCA on the AA trajectories, and only then convert the obtained principal components into CG representation.

A very general comparison of the conformational ensembles obtained in different simulations can be performed by calculating the Pearson correlation coefficients (PCCs) of the covariance matrix elements<sup>28</sup>. Analysis of the coarse grained trajectories reveals that AA simulations produce very similar covariance values that are however different from those obtained in CG simulations (Fig. 3). When compared in the common AA basis, AA covariance matrices produce PCC values similar to those obtained in CG basis (Figure S4). Comparison of present DOPC simulations at 310 K with those from our previous work<sup>28</sup> reveals better correspondence between Martini and latest AA force fields (CHARMM36, Lipid14 and Slipids) than between Martini and Berger/GROMOS family UA forcefields to which Martini was originally compared<sup>17,18</sup> (Figure S5). Finally, both in AA and CG simulations, DOPC-SOPC differences are the same as DOPC-OSPC, SOPC-DSPC and OSPC-DSPC differences, and DOPC-DSPC differences are the same as SOPC-OSPC differences (Figs 3 and S4). Interestingly, covariance matrices for DOPC simulated at different temperatures (310–338 K) show very little change both in AA and CG simulations (Figure S6). Correspondence between the covariance matrices obtained in AA and CG simulations is slightly better for the CG simulations conducted at 338 K (Figure S6).



**Figure 4.** Comparison of the coarse grained PCA eigenvectors obtained in the AA simulations with the ones obtained in the CG simulations. Dot product matrices of the corresponding eigenvector sets are shown. Black and white correspond to the dot product equal to one and zero, respectively.

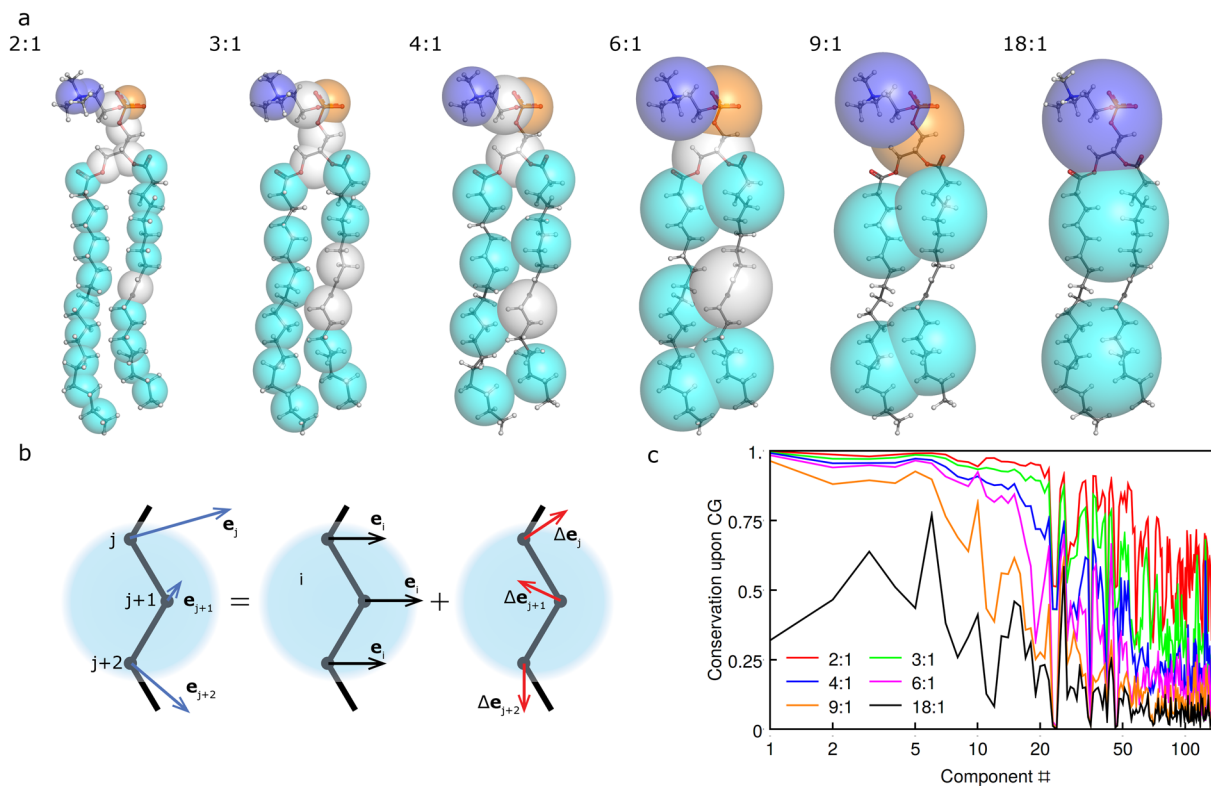
To find the differences in behavior of the four lipids, we compared the principal components at each level of resolution. The dot product matrices are almost diagonal, and the DOPC-DSPC and OSPC-SOPC differences are, as expected, larger than those between DOPC and SOPC/OSPC, or between SOPC/OSPC and DSPC (Figure S7).

Next, we compared collective motions of each lipid at different resolution levels. Comparison of the coarse grained AA principal components with those from the CG simulations again reveals close correspondence and diagonal structure of the dot-product matrices, which are however significantly noisier, and non-major atomistic collective motions are not captured very well in the CG simulations (Fig. 4).

**General effects of reduction in dimensionality.** Comparison of atomistic collective motions to those observed in CG simulations merits additional considerations. Development of each CG model comprises two major steps: 1) choice of the preferred CG model architecture and appropriate mapping; 2) parameterization of the constructed model<sup>33,34</sup>. Since the number of degrees of freedom in CG models is always smaller than in atomic ones, some of atomic motions are inevitably eliminated. Therefore, part of the differences between the AA and CG simulations is inherent to the chosen architecture of the CG model and does not depend on its parameterization. We note that the loss of information associated with dimensionality reduction is usually not proportional to the number of degrees of freedom that are eliminated since the contribution and amplitudes of motions along different degrees of freedom are not equal. This is expected, because the goal of coarse graining is to reduce the computational complexity without losing the details of interest.

To find how well the principal components are retained upon coarse graining, we analyze several hypothetical mappings with different atoms-to-beads ratios (Table S1, Fig. 5a). Displacement of each atom  $j$  in the atomic model can be represented as a displacement of the corresponding bead  $i$  in the CG model plus motion of the atom relative to the center of the bead (Fig. 5b). Motions of atoms relative to each other are eliminated in the CG model. The ratio of retained RMSD of the beads positions in the CG model to the RMSD of the atoms positions in the atomic model can be used to measure how well corresponding motion is conserved in the CG model. The ratio also should take into account different weights of different atoms in particular atom-to-bead mapping. With these considerations, we arrive at a formula described in Methods.

Recently, Foley *et al.* applied the relative entropy approach<sup>73</sup> to evaluate the impact of CG mapping upon information in coarse-grained models<sup>74</sup>. We note that their approach is complementary to ours. Indeed, our approach provides the information about a particular motion irrespective of associated thermodynamic properties, whereas the relative entropy approach provides the thermodynamic information about the whole system



**Figure 5.** Loss of motion details upon coarse graining. **(a)** Hypothetical mappings with varying atom-to-bead ratios. Details of the mappings may be found in Supplementary Table S1. **(b)** Decomposition of atomic displacements into bead displacements and displacements of atoms relative to each other. **(c)** Conservation of SOPC collective motions (principal components) upon hypothetical coarse grainings using different atom-to-bead mappings.

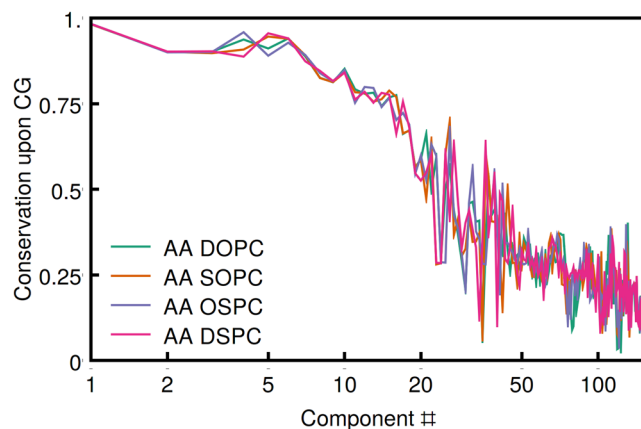
(so-called mapping entropy) and relies on the chosen potential and thermodynamic properties of the reference (for example, atomistic) model<sup>73,74</sup>.

Using SOPC as an example, we find that with the 2:1, 3:1, 4:1 and 6:1 mappings all of the major collective motions (up to PC18) are essentially conserved, whereas smaller amplitude motions are progressively eliminated (Fig. 5c). With the 9:1 mapping even the major PCs become affected. Finally, in the case of the extreme 18:1 coarse graining, similar to the model by Cooke *et al.*<sup>75</sup>, only the bending degrees of freedom are partially conserved, with other motions being almost completely eliminated (Fig. 5c). Overall, these results show that construction of meaningful CG models retaining most of the lipid collective motions is possible when up to ~6 heavy atoms are mapped to each CG bead.

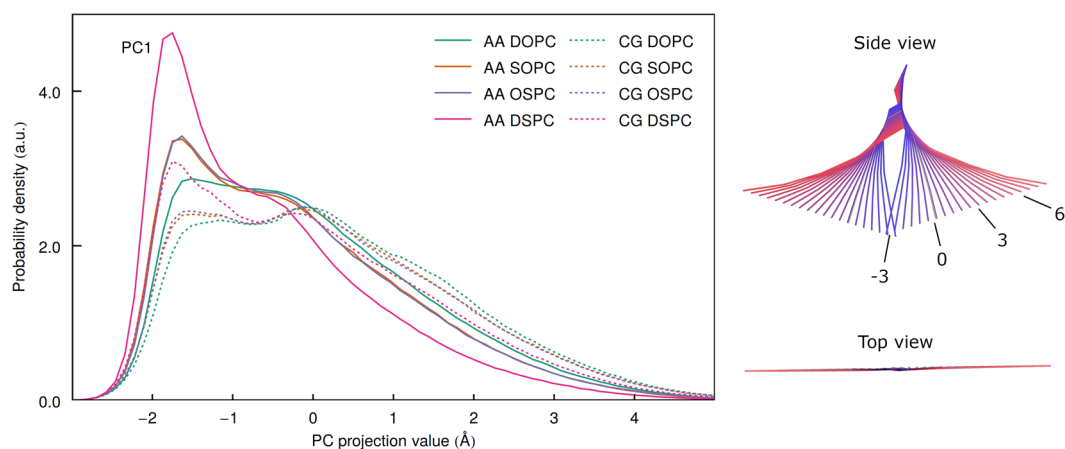
In Martini, each bead on average represents 4 heavy atoms<sup>17</sup>, however the bead definitions were recently updated to include up to 9 atoms with individual weights ranging from  $1/13^{\text{th}}$  to  $1/3^{\text{rd}}$ <sup>76</sup>. We find that, similarly to the hypothetical CG models analyzed earlier, Martini representation retains 75% to 100% of the major DOPC, SOPC, OSPC and DPPC collective motions but significantly reduces the minor ones (Fig. 6). The degree of motion conservation upon coarse graining follows the same pattern as the collectivity of principal components (calculated previously<sup>28</sup>).

While principal components are orthogonal by definition, coarse graining and consequent reduction in dimensionality of the conformational space results in loss of some motions (displacements of the atoms within one CG bead relative to each other) and non-zero overlap of the remaining ones. Indeed, comparison of the coarse-grained AA principal components reveals almost complete loss of smaller-amplitude motions and emergence of spurious similarities between the larger-amplitude ones (Figure S8).

**Comparison of configurational spaces.** Next, we analyze the conformational properties of the lipids by comparing the distributions of the PC projections in the common CG PCA basis (Figs 7 and 8 and S9–S10). As was observed before, the distribution of PC1 projection values is clearly non-Gaussian<sup>28</sup>. There are three peaks of varying amplitudes that correspond to the conformations of the lipid molecule where the hydrocarbon chains are in direct contact; separated by a single chain from another molecule; and separated by two and more other chains. Both in AA and CG simulations the first peak in the distribution is the highest for the DSPC, intermediate for SOPC/OSPC and the lowest for DOPC, and the distribution is the narrowest for DSPC, intermediate for SOPC/OSPC and widest for DOPC lipid molecules. The distributions for SOPC and OSPC are almost identical. Overall, this is in agreement with the observation that more unsaturated lipid molecules are more flexible and disordered<sup>63</sup>, and have higher equilibrium RMSD (Fig. 9). At the same time, for each lipid the first peak in the



**Figure 6.** Conservation of the motions corresponding to different PCs in atomistic simulations upon Martini coarse graining (please find the description in Methods). The major motions are very well conserved whereas the lower-amplitude PCs are mostly eliminated.

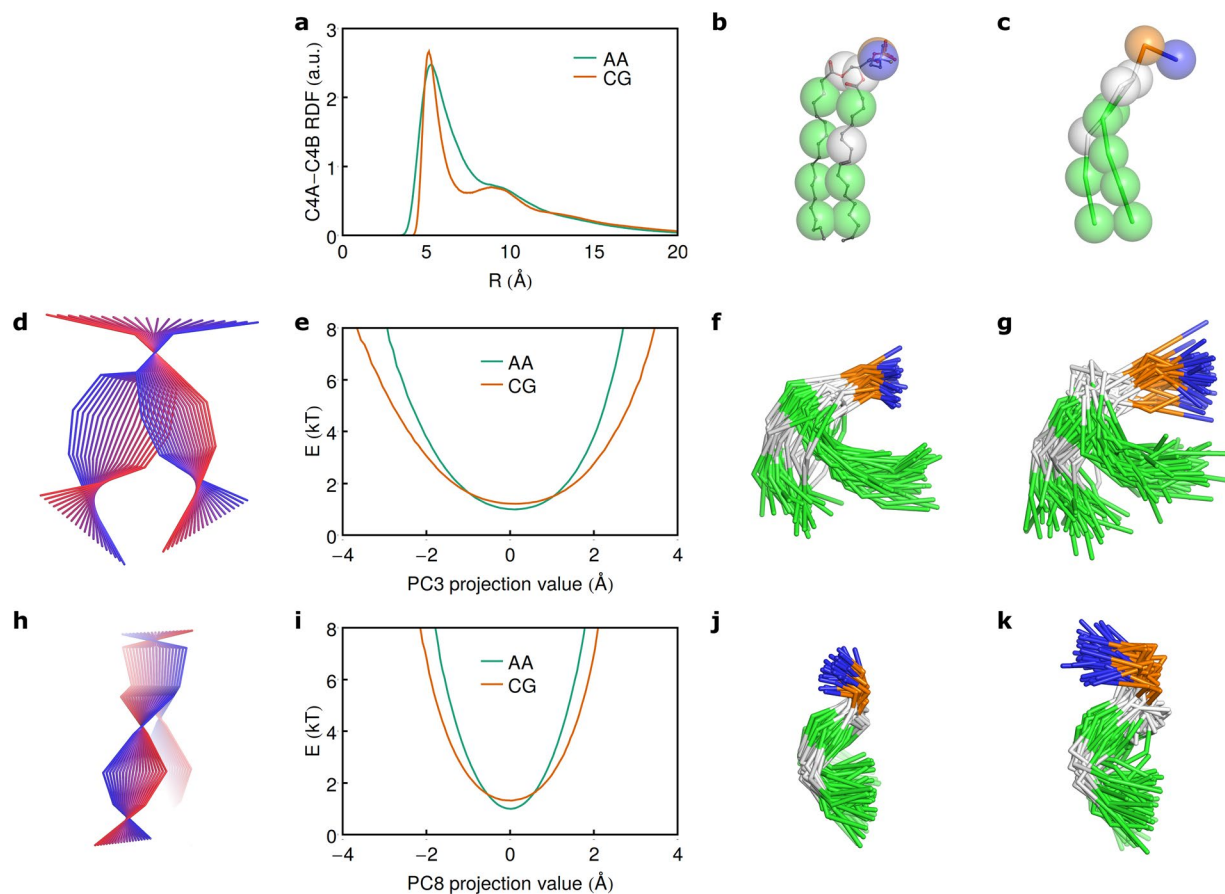


**Figure 7.** Influence of coarse graining and saturation of hydrocarbon chains on the scissoring motion of hydrophobic chains. Distributions of the PC1 projections in the same PCA eigenvector basis are shown on the left and corresponding structures are shown on the right.

PC1 distribution is the highest in AA and the lowest in CG simulations, while the probability distribution is the narrowest in AA simulations and the widest in CG simulations. Perhaps, this is a consequence of higher number of interaction sites and resulting higher friction and restriction of motions in more detailed FFs<sup>77</sup>. Similar overestimation of flexibility in CG simulations was observed before<sup>53,55,56</sup>.

PC1 distributions prompted us to calculate intra- and interlipid radial distribution functions for the ends of hydrocarbon chains (Figures S11–12). The distributions of distances within one molecule (Figure S11) are understandably similar to distributions of projections on PC1 (Fig. 7). The first- and second-order RDF peaks in the CG simulations are much sharper (Figures S11–12). This has been noted before<sup>53,54</sup> and probably reflects the fact that CG chains cannot intertwine as easily as the chains in atomic representation. Indeed, comparison of the conformations with closest chain-end beads in atomic and CG simulations reveals exactly that (Fig. 8b,c). Similarly, other bead-bead RDFs in CG simulations might diverge from the ones in the atomic simulations, and thus CG potentials might need to be corrected as suggested elsewhere<sup>40,54,78</sup>.

The distributions for other PCs are closer to Gaussian, although some of them are clearly asymmetric (for example, PCs 5, 6, 7 and 10, Figure S9). Comparisons reveal that in many cases CG simulations result in much broader distributions both for in plane (for example, PC3, Fig. 8d) and out of plane (for example, PC8, Fig. 8h) motions. Evidently, lipid molecules in Martini simulations are much more flexible, and can access highly bent and possibly protruding conformations not accessible in atomic simulations (see examples in Fig. 8d–k). While no direct data on conformations of lipid molecules are available, single molecule protrusions have been observed both in simulations<sup>79</sup> (the frequency of protrusions depends strongly on force field<sup>21</sup>) and fluorescence experiments with pyrene-labeled lipids<sup>80,81</sup>. Still, the flexibility of the lipid molecules might be overestimated in CG simulations, as was observed before<sup>53,55,56</sup>.



**Figure 8.** Examples of the extreme conformations from the AA and CG simulations, as exemplified by SOPC. (a) Radial distribution functions for the C4A and C4B beads belonging to the same molecule. AA simulations were coarse grained to obtain the data. (b,c) Conformations with the minimal C4A-C4B distances observed in AA and CG simulations (3.4 Å and 4.2 Å, respectively). In the atomic simulations, the inter-bead distance can become smaller due to intertwining of the hydrocarbon chains. (d) Collective motion associated with PC3. (e) Effective energy landscape for PC3. (f,g) Conformations with the smallest projection values on PC3 observed in AA and CG simulations (projection values from  $-3.5$  to  $-3$  Å and from  $-4.4$  to  $-3.9$  Å, respectively). In the CG simulations, the molecule is significantly more flexible. (h) Collective motion associated with PC8. (i) Effective energy landscape for PC8. (j,k) Conformations with the smallest projection on PC8 observed in AA and CG simulations (projection values from  $-2.2$  to  $-1.9$  Å and from  $-2.5$  to  $-2.3$  Å, respectively). In the CG simulations, the molecule is more flexible.

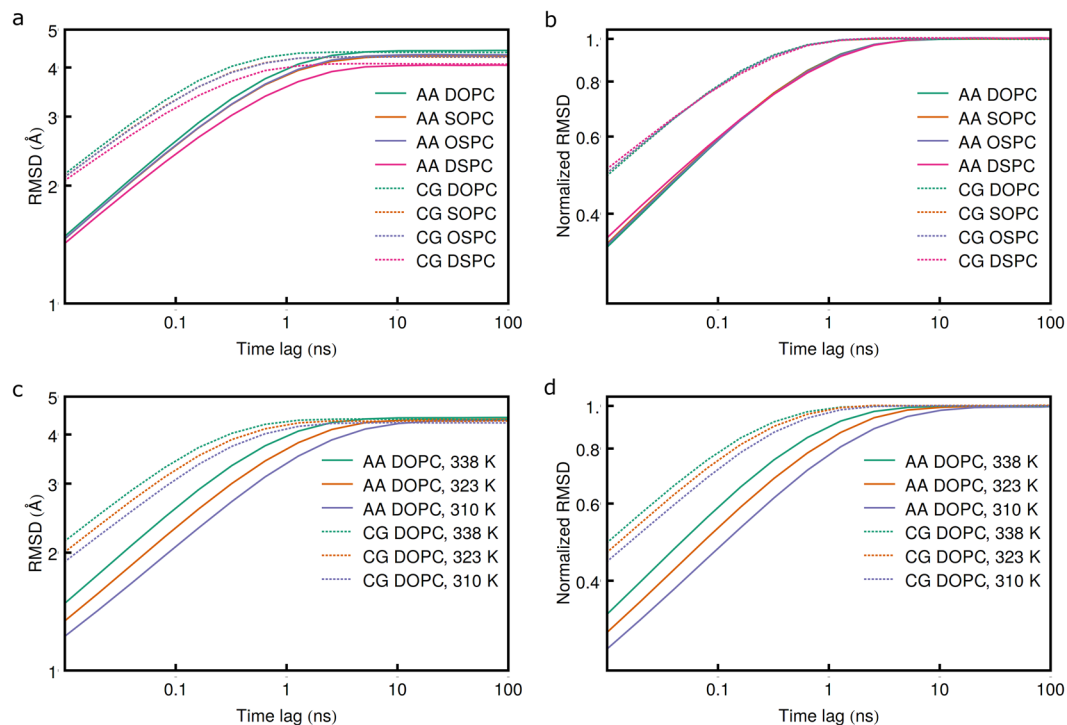
**Effects of hydrocarbon chain saturation on conformations.** Comparison of DOPC, SOPC, OSCP and DSPC configurational spaces in the same CG PCA basis reveals mostly similar distributions, with some collective motions revealing more differences (for example, PCs 1,5–10, Figure S9) than others (PCs 2–4, Figure S9).

Comparison of the lipids in atomistic PCA basis (Figures S13 and S14) reveals similar patterns. The differences and similarities in distributions depend on the nature of corresponding collective motion. PC1, which corresponds to scissoring motion (Figure S3), reveals more extended DOPC, intermediate and similar SOPC and OSCP, and more compact DSPC distributions (Figure S13). PC3, corresponding to mostly in-plane bending of the whole lipid molecule (Figure S3), reveals almost no differences between the lipids (Figure S13). On the contrary, PC6 is able to distinguish all the four lipids from each other (Figure S13).

Overall, the obtained results suggest that the effects of monounsaturated chains are additive. Pearson correlation coefficient (PCC) of the DOPC and SOPC/OSCP covariance matrices is the same as PCC of the SOPC/OSCP and DSPC covariance matrices, and PCC of the DOPC and DSPC covariance matrices is the same as PCC of the SOPC and OSCP covariance matrices. Distributions of SOPC and OSCP PC1 and most of other projections are similar, and are intermediate between those of DOPC and DPPC. At the same time, it should be noted that positional isomers of phosphatidylcholine lipids (such as SOPC and OSCP) have slightly different phase transition temperatures<sup>82</sup>, and thus contributions of the *sn*-1 and *sn*-2 chains to the lipids' physical properties are not identical.

**Effects of temperature on conformations.** Comparing AA and CG simulations of DOPC conducted at 310, 323 and 338 K, we observe very little differences in configurational spaces (Figures S15–S17). The lipids are slightly more ordered at lower temperatures (Figure S15), as could be expected, but no other significant effects of





**Figure 9.** Average RMSD (left) and normalized RMSD (right) of lipid heavy atoms or particles positions as a function of time lag. Note that the structures were aligned prior to RMSD calculation and thus there is no contribution from diffusion.

temperature on the available configurational space are observed. Covariance matrices show very little change with temperature both in AA and CG simulations (Figure S6).

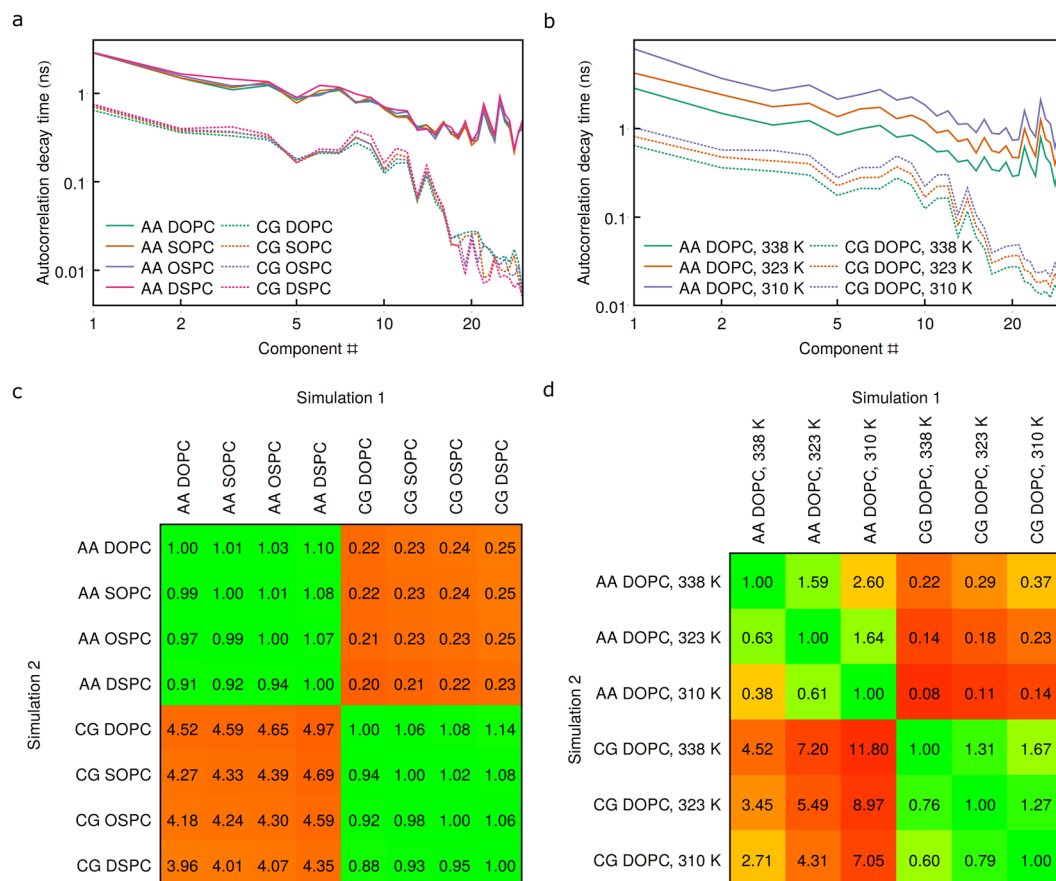
**Convergence of the simulations.** 200 ns long trajectories of systems consisting of  $\sim 100$  lipids with snapshots taken each 10 ps contain  $\sim 2 \cdot 10^6$  individual conformations. As we discussed earlier, such simulation lengths are sufficient for relaxation of even the slowest motions and convergence of the observed variables<sup>28</sup>, with characteristic autocorrelation decay times converging to within 5% (data not shown).

We begin the analysis and comparison of characteristic time scales in the present simulations by studying the dependence of structure RMSD of individual lipid molecules on time lag (Fig. 9, note that the molecules were aligned prior to RMSD calculation, as described in Methods). Overall, coarse graining results in 4.6 (at 338 K) to 7.2 (at 310 K) times faster equilibration. DOPC RMSD values plateau at the highest levels, whereas SOPC/OSPC and DSPC RMSD values are intermediate and lowest, correspondingly. RMSD values for SOPC and OSPC are almost identical. At the same temperature DOPC, SOPC, OSPC and DSPC equilibrate at similar rates, however more unsaturated lipids equilibrate slightly faster. Fitting the normalized RMSD with exponential function  $c_1 \tau^{-c_2}$  in the linear region of the plot reveals  $c_2$  values of 0.214, 0.207, 0.205 and 0.193 in the atomistic simulations, and 0.240, 0.233, 0.234 and 0.222 in the CG simulations, for DOPC, SOPC, OSPC and DSPC, respectively. Compared to the simulations conducted at 310 K, in atomistic simulations, RMSD converges 1.56 and 2.46 times faster at 323 K and 338 K, respectively, and in CG simulations, RMSD converges 1.26 and 1.59 times faster at 323 K and 338 K, respectively. The influence of temperature will be discussed below in more detail.

While studying RMSD time dependence is instructive, it does not provide definitive answers, because some particular motions may equilibrate much slower than the RMSD of the molecule as a whole<sup>28</sup>. Therefore, we investigated the dynamics and convergence of simulations in the coordinates defined by PCs.

**Characteristic equilibration time scales.** First of all, characteristic time scales (the time in which the autocorrelation decays in  $e^2$  times) in different AA simulations conducted at the same temperature are very close to each other and are below 10 ns (Figs 10 and S18). As was observed for RMSD (Fig. 9), SOPC and OSPC equilibrate slightly slower than DOPC, and DSPC equilibrates notably slower. This slowdown is not uniform both in AA and CG simulations: for example, for PC1, DSPC is  $\sim 1\%$  slower than DOPC in AA simulations (17% slower in CG); for PC3 DSPC is  $\sim 32\%$  slower than DOPC in AA simulations (26% slower in CG). Similar trends are observed when PCA is conducted only on atomistic simulations in AA basis (Figure S19). Surprisingly, all of the collective motions are accelerated very homogeneously with rising temperature (Figs 10 and S19).

In the CG simulations the autocorrelation decays  $\sim 4.5$  times faster at 338 K,  $\sim 5.5$  times faster at 323 K and  $\sim 7$  times faster at 310 K, compared to AA simulations (Fig. 10). The speedup is roughly homogenous for the major principal components, but becomes more significant for lower-amplitude motions (Fig. 10), perhaps, due to filtering out of these motions in CG representation (Fig. 6). Therefore, the dynamical processes are accelerated



**Figure 10.** Characteristic autocorrelation decay times for the analysis in the joint PC basis, and corresponding speedups (geometric mean ratios, please find the description in Methods). The ratios are color-coded to highlight the similarities and dissimilarities between the simulations.

unevenly in CG simulations, similarly to what was observed before<sup>59</sup>. The acceleration is different at different temperatures (Fig. 10), as has been observed for various thermodynamic properties<sup>53</sup>. The obtained values of ~7 fold acceleration at 310 K are slightly higher than what was observed previously, whereas 4–5 fold acceleration at 338 K is similar<sup>19, 57, 59</sup>.

## Conclusions

In this work, we studied the single molecule behavior of saturated and unsaturated lipids DOPC, SOPC, OSPC and DSPC in atomistic and CG MD simulations by means of principal components analysis. The analysis was effective at identification and pinpointing similarities and differences between the simulations. We found that the conformational spaces available to the four lipids significantly overlap, although there is a notable difference: lipids with more unsaturated hydrocarbon chains favor more extended conformations (Figs 2, 7 and 9). The dynamics of the four lipids is very similar when studied at the same temperature, with unsaturated lipids equilibrating slightly faster (Fig. 10). At lower temperatures, all of the motions are homogeneously slowed down (Fig. 10). Finally, we demonstrated how the results of the AA simulations can be compared directly to the results of the CG simulations, and showed that the latter capture 75 to 100% of the major conformational changes (Figs 4–6). Our approach allowed us to identify the conformations that are observed in AA simulations and not observed in CG ones, and vice versa (Fig. 8). Overall, the sampling was ~4–7 times faster in the CG simulations (Fig. 10). The CG sampling speedups were roughly the same for all of the major motions, but increased for lower amplitude motions (Fig. 10a,b) that are weakly conserved in CG simulations (Figs 4–6). We expect that the results reported here will be useful for efficient quantitative comparisons of simulations conducted at different levels of resolution and for further development and improvement of Martini and other CG lipid force fields. Comparison of the average structures might be used to improve CG mappings, and comparison of covariance matrices, projection distributions and characteristic timescales might be used to optimize the restraints and force constants.

## Materials and Methods

**Performed simulations.** Details of the performed simulations are summarized in Table 1. The simulations were performed using GROMACS version 5.1<sup>83</sup>. The simulations AA1–6 were performed using CHARMM36 force field and CHARMM TIP3P water model<sup>84, 85</sup>. The simulations CG1–6 were performed using Martini force field version 2.2<sup>17–19</sup>, and polarizable water model<sup>86</sup> where 4 water molecules are treated as a polar solvent particle.

4-bead representation of oleic and stearic acids was used<sup>87</sup>. Initial structures and topologies for simulations AA6 and CG6 were prepared using the CHARMM-GUI Membrane Builder<sup>88–92</sup>. Initial structures for simulations AA1–5, CG1–5 were generated from the AA6 and CG6 initial structures. Unsaturated bonds were introduced into the structure and topology files manually where needed. AA and CG systems were energy-minimized and equilibrated for 50 ns preceding the production runs. Evolution of the areas per lipid in the simulated systems is shown in Figure S1.

**Simulation conditions.** Parameters in the AA and CG simulations were chosen to match the original reports by J. Lee *et al.*<sup>89</sup> and the recommended parameters from the Martini web site<sup>76,93</sup> (Martini\_v2.x\_common updated July 15<sup>th</sup>, 2015), respectively. In the AA and CG simulations the leapfrog integrator with the time step sizes of 2 fs and 20 fs, respectively, was used. The system snapshots were collected every 10 ps. Periodic boundary conditions were applied in all 3 directions. The centers of mass of the bilayer and solvent were fixed. In the AA simulations covalent bonds to hydrogen atoms were constrained using the SHAKE algorithm<sup>94</sup>. The simulations AA1,4–6, CG1,4–6 were conducted at the reference temperature of 338 K. The simulations AA2 and CG2 were conducted at 323 K, and AA3 and CG3 at 310 K. The simulations were conducted using the reference pressure of 1 bar. Lipid and water molecules were coupled to the temperature baths separately. The AA simulations were performed using Nosé-Hoover temperature coupling method<sup>95</sup> with the coupling constant of 1 ps<sup>-1</sup> and a semi-isotropic Parrinello-Rahman barostat<sup>96</sup> with the relaxation time of 5 ps. The CG simulations were performed with the velocity rescale thermostat<sup>97</sup> with the coupling constant of 1 ps<sup>-1</sup> and the semiisotropic Parrinello-Rahman barostat<sup>96</sup> with the relaxation time of 12 ps.

The nonbonded pair list was updated every 20 steps with the cutoffs of 1.2 and 1.4 nm in the AA and CG simulations, respectively. Force-based switching function with the switching range of 1.0–1.2 nm and particle mesh Ewald (PME) method with 0.12 nm Fourier grid spacing and 1.2 nm cutoff were used for treatment of the van der Waals and electrostatic interactions in the AA simulations. Potentials shifted to zero at the cutoffs of 1.1 nm and a reaction-field potential with  $\epsilon_{rf} = \infty$  were used for treatment of the van der Waals and electrostatic interactions in the CG simulations, as recommended<sup>93</sup>.

**Analysis.** *Structure alignment.* All of the analyses of the lipid molecule conformations were preceded by non mass-weighted least squares alignment of the analyzed structures to the reference average structure (see Fig. 1c and d for the conformations from the simulations AA4 and CG4). The reference structure was obtained in two steps as described previously<sup>28</sup>. First, the structures from the simulation AA1 were aligned to an arbitrarily chosen structure, and the first average structure has been obtained. Second, the structures were realigned to the first average structure, resulting in the second average structure. The second average structure was not affected by the choice of the reference structure in the first step, and it was used as a reference structure for all other alignments. The AA structures were aligned using the heavy atoms. The CG structures were aligned to the coarse grained reference average structure. To highlight the differences in the positions of the hydrophobic chains, the average structures in the Fig. 2 were aligned to each other using the headgroup atoms and the first 8 atoms of the acyl chains (for AA simulations) and using the headgroup beads and the first 2 beads of the acyl chains (for CG simulations).

*Coarse graining of atomistic structures.* Atomic structures were coarse grained using *backward*<sup>98</sup>:

$$C = \omega_{ij}A, \quad (1)$$

where A is a  $n \times 3$  matrix of atomistic coordinates in Cartesian space, C is a  $m \times 3$  matrix of coordinates of coarse grained particles, and  $\omega_{ij}$  is the CG mapping matrix. The mappings were based on the data from the Martini web site<sup>76</sup>. Atomic displacements, corresponding to principal components, were obtained similarly:

$$e_i^{cg} = \sum_{j=1}^n \omega_{ij} e_j^{aa}, \quad (2)$$

where  $e_i^{aa}$  is the vector in the atomistic representation and  $e_i^{cg}$  is the corresponding vector in the CG space.

*Pearson correlation coefficient.* Pearson correlation coefficients of the covariance matrices  $M^i$  and  $M^j$  obtained in simulations  $i$  and  $j$ , correspondingly, were calculated as follows:

$$r_{ij} = \frac{\text{cov}(M^i, M^j)}{\sigma M^i \cdot \sigma M^j} \quad (3)$$

where  $\text{cov}(M^i, M^j)$  is the covariance of the elements of  $M^i$  and  $M^j$ , and  $\sigma M^i$  and  $\sigma M^j$  are the standard deviations of the elements of  $M^i$  and  $M^j$ . The covariance matrices for calculation of the Pearson correlation coefficients were obtained using the coarse grained trajectories (Figs 3, S4–S6 and S19).

*Principal component analysis.* In AA representation, PCA was performed on the positions of heavy atoms. In the CG representations, PCA was performed on the positions of the beads. PCA<sup>72</sup> was performed using the *covar* and *anaeig* utilities of the GROMACS software package<sup>83</sup> as described previously<sup>28</sup>. *covar* was used to calculate the atomic displacement covariance matrices, their eigenvalues and eigenvectors, and *anaeig* was used to project the molecular dynamics trajectories onto the obtained eigenvector bases. Three types of analyses were performed. First, PCA was performed for each simulation separately (Figs 3, 4, 6 and S2–S8). Second, PCA was performed jointly for simulations AA1–6 and CG1–6 in the CG representation (Figs 7, 8, 10, S9–S10 and S15–S18). Finally, PCA was performed jointly for AA1–6 in the AA representation (Figs S13–S14 and S19). For comparisons of the

results of the simulations for different lipids and topologies the eigenvalues were normalized by dividing by the number of particles in the model. Conformational spaces were compared using dot product matrices of eigenvectors mapped to the CG space (Figs 4 and S8) or eigenvectors in the native AA or CG basis (Figure S7).

**Conservation upon coarse graining.** Conservation of PCA eigenvectors upon coarse graining (Figs 5c and 6) was calculated as follows:

$$\alpha = \frac{\text{RMSD}^{\text{cg}}}{\text{RMSD}^{\text{aa}}} = \frac{\sqrt{\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^n \omega_{ij} \mathbf{e}_{ji}^{\text{aa-cg}^2}}}{\sqrt{\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^n \omega_{ij} \mathbf{e}_j^{\text{aa}^2}}} = \frac{\sqrt{\frac{1}{m} \sum_{i=1}^m \mathbf{e}_i^{\text{cg}^2}}}{\sqrt{\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^n \omega_{ij} \mathbf{e}_j^{\text{aa}^2}}} \leq 1 \quad (4)$$

where  $\mathbf{e}_i^{\text{aa}}$  is the  $n$ -dimensional vector of atomic displacements in the all atom representation,  $\mathbf{e}_{ji}^{\text{aa-cg}}$  is the  $n$ -dimensional vector of atomic displacements, where displacement of each atom  $j$  is the same as displacement of the corresponding bead  $i$ ,  $\mathbf{e}_i^{\text{cg}}$  is the  $m$ -dimensional vector of bead displacements upon the coarse graining procedure, and  $\omega_{ij}$  is the CG mapping matrix. Eliminated atomic motions  $\mathbf{e}_j^{\text{el}}$  constitute the kernel (null space) of the CG mapping matrix:

$$\sum_{j=1}^n \omega_{ij} \mathbf{e}_j^{\text{el}} = 0 \quad (5)$$

**Radial distribution function.** The radial distribution functions were calculated using the *rdf* utility of the GROMACS software package<sup>83</sup>.

**Effective potentials.** Effective potentials associated with the conformational changes along the eigenvectors were calculated as follows (Figs 8e,i, S9–S10 and S13, S14, S16 and S17):

$$E(x) = -k_B T \log P(x) \quad (6)$$

where  $P(x)$  is the probability density and  $x$  is the value of the projection on corresponding PC eigenvector.

**Root-mean-square-deviation.** Root-mean-square deviations (RMSD), presented in Fig. 9, were calculated as follows:

$$\text{RMSD}(\tau) = \frac{1}{N_L} \sum_{i=1}^{N_L} \langle \text{RMSD}_i(t, t + \tau) \rangle_t \quad (7)$$

where  $N_L$  is the number of the lipid molecules present in the system, and  $\langle \text{RMSD}_i(t, t + \tau) \rangle_t$  is the root-mean-square deviation of the  $i$ -th molecule heavy atoms positions between the time moments  $t$  and  $t + \tau$ , averaged over the time  $t$ . Normalized RMSD (NRMSD) was calculated as follows:

$$\text{NRMSD}(\tau) = \frac{\text{RMSD}(\tau)}{\text{RMSD}_{\text{eq}}} \quad (8)$$

where  $\text{RMSD}_{\text{eq}}$  is the equilibrium value of RMSD at  $\tau \rightarrow \infty$ . NRMSD was fitted with exponential function in the linear region of the plot (normalized RMSD values below 0.7):

$$\text{NRMSD}(\tau) = c_1 \tau^{c_2} \quad (9)$$

**Autocorrelation.** Autocorrelation  $R$  of the projections on principal components was calculated as follows:

$$R(\tau) = \frac{\langle (p(t) - \langle p(t) \rangle_t)(p(t + \tau) - \langle p(t) \rangle_t) \rangle_t}{\langle (p(t) - \langle p(t) \rangle_t)^2 \rangle_t} \quad (10)$$

where  $p(t)$  is the projection value at the time  $t$  and  $\langle \bullet \rangle_t$  is averaging over  $t$ .

**Characteristic time scales.** As autocorrelation decay could not be adequately fitted with power or exponential function<sup>28</sup>, we calculated characteristic times  $\tau$  needed for autocorrelation to decrease in  $e^2$  times from the starting value, where  $e$  is the natural logarithm base (Figs 10 and S18–S19). For simplicity, the starting value for autocorrelation was assumed to be 1.0. The time dependence values were linearly extrapolated where needed.

The eigenvalue-weighted geometric means of the characteristic time scale ratios in the simulations  $i$  and  $j$ , measured in common PC basis, were calculated as follows:

$$R(i, j) = \exp \left( \sum_{k=1}^{N_F} \beta E_k \log \frac{\tau_k^i}{\tau_k^j} \right) \quad (11)$$

where  $N_F$  is the number of the degrees of freedom for lipid model,  $E_k$  is the  $k$ -th eigenvalue of the covariance matrix,  $\beta$  is the normalization factor chosen so that  $\sum_{k=1}^{N_F} \beta E_k = 1$ , and  $\tau_k^i$  and  $\tau_k^j$  are the characteristic time scales of the distributions convergence or autocorrelation decay for the  $k$ -th PC in the simulations  $i$  and  $j$ .

## References

- Bennett, W. F. D. & Tieleman, D. P. Computer simulations of lipid membrane domains. *Biochim. Biophys. Acta BBA - Biomembr.* **1828**, 1765–1776 (2013).
- Marrink, S. J., de Vries, A. H. & Tieleman, D. P. Lipids on the move: Simulations of membrane pores, domains, stalks and curves. *Biochim. Biophys. Acta BBA - Biomembr.* **1788**, 149–168 (2009).
- Lindahl, E. & Sansom, M. S. Membrane proteins: molecular dynamics simulations. *Curr. Opin. Struct. Biol.* **18**, 425–431 (2008).
- Venturoli, M., Maddalena Sperotto, M., Kranenburg, M. & Smit, B. Mesoscopic models of biological membranes. *Phys. Rep.* **437**, 1–54 (2006).
- Klauda, J. B. *et al.* Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *J. Phys. Chem. B* **114**, 7830–7843 (2010).
- Dickson, C. J. *et al.* Lipid14: The Amber Lipid Force Field. *J. Chem. Theory Comput.* **10**, 865–879 (2014).
- Maciejewski, A., Pasenkiewicz-Gierula, M., Cramariuc, O., Vattulainen, I. & Rog, T. Refined OPLS All-Atom Force Field for Saturated Phosphatidylcholine Bilayers at Full Hydration. *J. Phys. Chem. B* **118**, 4571–4581 (2014).
- Kulig, W., Pasenkiewicz-Gierula, M. & Róg, T. Cis and trans unsaturated phosphatidylcholine bilayers: A molecular dynamics simulation study. *Chem. Phys. Lipids* **195**, 12–20 (2016).
- Jämbeck, J. P. M. & Lyubartsev, A. P. Derivation and Systematic Validation of a Refined All-Atom Force Field for Phosphatidylcholine Lipids. *J. Phys. Chem. B* **116**, 3164–3179 (2012).
- Berger, O., Edholm, O. & Jähnig, F. Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys. J.* **72**, 2002–2013 (1997).
- Chiu, S.-W., Pandit, S. A., Scott, H. L. & Jakobsson, E. An Improved United Atom Force Field for Simulation of Mixed Lipid Bilayers. *J. Phys. Chem. B* **113**, 2748–2763 (2009).
- Kukul, A. Lipid Models for United-Atom Molecular Dynamics Simulations of Proteins. *J. Chem. Theory Comput.* **5**, 615–626 (2009).
- Poger, D., Van Gunsteren, W. F. & Mark, A. E. A new force field for simulating phosphatidylcholine bilayers. *J. Comput. Chem.* **31**, 1117–1125 (2010).
- Lee, S. *et al.* CHARMM36 United Atom Chain Model for Lipids and Surfactants. *J. Phys. Chem. B* **118**, 547–556 (2014).
- Ulmschneider, J. P. & Ulmschneider, M. B. United Atom Lipid Parameters for Combination with the Optimized Potentials for Liquid Simulations All-Atom Force Field. *J. Chem. Theory Comput.* **5**, 1803–1813 (2009).
- Tjörnhammar, R. & Edholm, O. Reparameterized United Atom Model for Molecular Dynamics Simulations of Gel and Fluid Phosphatidylcholine Bilayers. *J. Chem. Theory Comput.* **10**, 5706–5715 (2014).
- Marrink, S. J., de Vries, A. H. & Mark, A. E. Coarse Grained Model for Semiquantitative Lipid Simulations. *J. Phys. Chem. B* **108**, 750–760 (2004).
- 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. *J. Phys. Chem. B* **111**, 7812–7824 (2007).
- Marrink, S. J. & Tieleman, D. P. Perspective on the Martini model. *Chem. Soc. Rev.* **42**, 6801–6822 (2013).
- Högberg, C.-J., Nikitin, A. M. & Lyubartsev, A. P. Modification of the CHARMM force field for DMPC lipid bilayer. *J. Comput. Chem.* **29**, 2359–2369 (2008).
- Pluhackova, K. *et al.* A Critical Comparison of Biomembrane Force Fields: Structure and Dynamics of Model DMPC, POPC, and POPE Bilayers. *J. Phys. Chem. B* **120**, 3888–3903 (2016).
- Siu, S. W. L., Vácha, R., Jungwirth, P. & Böckmann, R. A. Biomolecular simulations of membranes: Physical properties from different force fields. *J. Chem. Phys.* **128**, 125103 (2008).
- Lyubartsev, A. P. & Rabinovich, A. L. Force Field Development for Lipid Membrane Simulations. *Biochim. Biophys. Acta BBA - Biomembr.* **1858**, 2483–2497 (2016).
- Piggot, T. J., Piñeiro, Á. & Khalid, S. Molecular Dynamics Simulations of Phosphatidylcholine Membranes: A Comparative Force Field Study. *J. Chem. Theory Comput.* **8**, 4593–4609 (2012).
- Lyubartsev, A. P. & Rabinovich, A. L. Recent development in computer simulations of lipid bilayers. *Soft Matter* **7**, 25–39 (2010).
- Skjerveik, Á. A. *et al.* Simulation of lipid bilayer self-assembly using all-atom lipid force fields. *Phys. Chem. Chem. Phys.* **18**, 10573–10584 (2016).
- Botan, A. *et al.* Toward Atomistic Resolution Structure of Phosphatidylcholine Headgroup and Glycerol Backbone at Different Ambient Conditions. *J. Phys. Chem. B* **119**, 15075–15088 (2015).
- Buslaev, P., Gordeliy, V., Grudinin, S. & Gushchin, I. Principal Component Analysis of Lipid Molecule Conformational Changes in Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **12**, 1019–1028 (2016).
- Ollila, O. H. S. & Pabst, G. Atomistic resolution structure and dynamics of lipid bilayers in simulations and experiments. *Biochim. Biophys. Acta BBA - Biomembr.* **1858**, 2512–2528 (2016).
- Neale, C. & Pomès, R. Sampling errors in free energy simulations of small molecules in lipid bilayers. *Biochim. Biophys. Acta BBA - Biomembr.* **1858**, 2539–2548 (2016).
- Bennun, S. V., Hoopes, M. I., Xing, C. & Faller, R. Coarse-grained modeling of lipids. *Chem. Phys. Lipids* **159**, 59–66 (2009).
- Brini, E. *et al.* Systematic coarse-graining methods for soft matter simulations – a review. *Soft Matter* **9**, 2108–2119 (2013).
- Saunders, M. G. & Voth, G. A. Coarse-graining methods for computational biology. *Annu. Rev. Biophys.* **42**, 73–93 (2013).
- Noid, W. G. Perspective: Coarse-grained models for biomolecular systems. *J. Chem. Phys.* **139**, 090901 (2013).
- Ingólfsson, H. I. *et al.* The power of coarse graining in biomolecular simulations. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **4**, 225–248 (2014).
- Tozzini, V. Coarse-grained models for proteins. *Curr. Opin. Struct. Biol.* **15**, 144–150 (2005).
- Clementi, C. Coarse-grained models of protein folding: toy models or predictive tools? *Curr. Opin. Struct. Biol.* **18**, 10–15 (2008).
- Lu, L. & Voth, G. A. Systematic Coarse-graining of a Multicomponent Lipid Bilayer. *J. Phys. Chem. B* **113**, 1501–1510 (2009).
- Izvekov, S. & Voth, G. A. Multiscale Coarse-Graining of Mixed Phospholipid/Cholesterol Bilayers. *J. Chem. Theory Comput.* **2**, 637–648 (2006).
- Izvekov, S. & Voth, G. A. A Multiscale Coarse-Graining Method for Biomolecular Systems. *J. Phys. Chem. B* **109**, 2469–2473 (2005).
- Orsi, M., Haubertin, D. Y., Sanderson, W. E. & Essex, J. W. A Quantitative Coarse-Grain Model for Lipid Bilayers. *J. Phys. Chem. B* **112**, 802–815 (2008).
- Orsi, M. & Essex, J. W. The ELBA Force Field for Coarse-Grain Modeling of Lipid Membranes. *PLOS ONE* **6**, e28637 (2011).
- Meyer, F. de & Smit, B. Effect of cholesterol on the structure of a phospholipid bilayer. *Proc. Natl. Acad. Sci.* **106**, 3654–3658 (2009).
- Rodgers, J. M., Sørensen, J., de Meyer, F. J.-M., Schiøtt, B. & Smit, B. Understanding the Phase Behavior of Coarse-Grained Model Lipid Bilayers through Computational Calorimetry. *J. Phys. Chem. B* **116**, 1551–1569 (2012).
- Shih, A. Y., Arkhipov, A., Freddolino, P. L. & Schulten, K. Coarse Grained Protein–Lipid Model with Application to Lipoprotein Particles. *J. Phys. Chem. B* **110**, 3674–3684 (2006).
- Monticelli, L. *et al.* The MARTINI Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theory Comput.* **4**, 819–834 (2008).

47. Periolo, X., Cavalli, M., Marrink, S.-J. & Ceruso, M. A. Combining an Elastic Network With a Coarse-Grained Molecular Force Field: Structure, Dynamics, and Intermolecular Recognition. *J. Chem. Theory Comput.* **5**, 2531–2543 (2009).
48. Usitalo, J. J., Ingólfsson, H. I., Akhshi, P., Tieleman, D. P. & Marrink, S. J. Martini Coarse-Grained Force Field: Extension to DNA. *J. Chem. Theory Comput.* **11**, 3932–3945 (2015).
49. López, C. A. *et al.* Martini Coarse-Grained Force Field: Extension to Carbohydrates. *J. Chem. Theory Comput.* **5**, 3195–3210 (2009).
50. Ingólfsson, H. I. *et al.* Lipid Organization of the Plasma Membrane. *J. Am. Chem. Soc.* **136**, 14554–14559 (2014).
51. van Eerden, F. J., de Jong, D. H., de Vries, A. H., Wassenaar, T. A. & Marrink, S. J. Characterization of thylakoid lipid membranes from cyanobacteria and higher plants by molecular dynamics simulations. *Biochim. Biophys. Acta BBA - Biomembr.* **1848**, 1319–1330 (2015).
52. López, C. A. *et al.* MARTINI Coarse-Grained Model for Crystalline Cellulose Microfibers. *J. Phys. Chem. B* **119**, 465–473 (2015).
53. Baron, R. *et al.* Comparison of Thermodynamic Properties of Coarse-Grained and Atomic-Level Simulation Models. *ChemPhysChem* **8**, 452–461 (2007).
54. Eichenberger, A. P., Huang, W., Riniker, S. & van Gunsteren, W. F. Supra-Atomic Coarse-Grained GROMOS Force Field for Aliphatic Hydrocarbons in the Liquid Phase. *J. Chem. Theory Comput.* **11**, 2925–2937 (2015).
55. Baron, R., de Vries, A. H., Hünenberger, P. H. & van Gunsteren, W. F. Comparison of Atomic-Level and Coarse-Grained Models for Liquid Hydrocarbons from Molecular Dynamics Configurational Entropy Estimates. *J. Phys. Chem. B* **110**, 8464–8473 (2006).
56. Baron, R., de Vries, A. H., Hünenberger, P. H. & van Gunsteren, W. F. Configurational Entropies of Lipids in Pure and Mixed Bilayers from Atomic-Level and Coarse-Grained Molecular Dynamics Simulations. *J. Phys. Chem. B* **110**, 15602–15614 (2006).
57. Goga, N. *et al.* Benchmark of Schemes for Multiscale Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **11**, 1389–1398 (2015).
58. Ramadurai, S. *et al.* Influence of Hydrophobic Mismatch and Amino Acid Composition on the Lateral Diffusion of Transmembrane Peptides. *Biophys. J.* **99**, 1447–1454 (2010).
59. Fritz, D., Koschke, K., Harmandaris, V. A., Vegt, N. F. A. vander & Kremer, K. Multiscale modeling of soft matter: scaling of dynamics. *Phys. Chem. Chem. Phys.* **13**, 10412–10420 (2011).
60. Koynova, R. & Caffrey, M. Phases and phase transitions of the phosphatidylcholines. *Biochim. Biophys. Acta* **1376**, 91–145 (1998).
61. Kučerka, N., Tristram-Nagle, S. & Nagle, J. F. Structure of Fully Hydrated Fluid Phase Lipid Bilayers with Monounsaturated Chains. *J. Membr. Biol.* **208**, 193–202 (2006).
62. Nagle, J. F. & Tristram-Nagle, S. Structure of lipid bilayers. *Biochim. Biophys. Acta BBA - Rev. Biomembr.* **1469**, 159–195 (2000).
63. Poger, D. & Mark, A. E. On the Validation of Molecular Dynamics Simulations of Saturated and cis-Monounsaturated Phosphatidylcholine Lipid Bilayers: A Comparison with Experiment. *J. Chem. Theory Comput.* **6**, 325–336 (2010).
64. Poger, D., Caron, B. & Mark, A. E. Validating lipid force fields against experimental data: Progress, challenges and perspectives. *Biochim. Biophys. Acta BBA - Biomembr.* **1858**, 1556–1565 (2016).
65. Jämbeck, J. P. M. & Lyubartsev, A. P. An Extension and Further Validation of an All-Atomistic Force Field for Biological Membranes. *J. Chem. Theory Comput.* **8**, 2938–2948 (2012).
66. Petrache, H. I., Dodd, S. W. & Brown, M. F. Area per Lipid and Acyl Length Distributions in Fluid Phosphatidylcholines Determined by 2H NMR Spectroscopy. *Biophys. J.* **79**, 3172–3192 (2000).
67. Kučerka, N., Nieh, M.-P. & Katsaras, J. Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature. *Biochim. Biophys. Acta BBA - Biomembr.* **1808**, 2761–2771 (2011).
68. Martinez-Seara, H. *et al.* Effect of Double Bond Position on Lipid Bilayer Properties: Insight through Atomistic Simulations. *J. Phys. Chem. B* **111**, 11162–11168 (2007).
69. Martinez-Seara, H., Róg, T., Karttunen, M., Vattulainen, I. & Reigada, R. Why is the sn-2 Chain of Monounsaturated Glycerophospholipids Usually Unsaturated whereas the sn-1 Chain Is Saturated? Studies of 1-Stearoyl-2-oleoyl-sn-glycero-3-phosphatidylcholine (SOPC) and 1-Oleoyl-2-stearoyl-sn-glycero-3-phosphatidylcholine (OSPC) Membranes with and without Cholesterol. *J. Phys. Chem. B* **113**, 8347–8356 (2009).
70. Zhuang, X., Makover, J. R., Im, W. & Klauda, J. B. A systematic molecular dynamics simulation study of temperature dependent bilayer structural properties. *Biochim. Biophys. Acta BBA - Biomembr.* **1838**, 2520–2529 (2014).
71. Chiu, S. W., Jakobsson, E., Subramaniam, S. & Scott, H. L. Combined Monte Carlo and Molecular Dynamics Simulation of Fully Hydrated Dioleoyl and Palmitoyl-oleoyl Phosphatidylcholine Lipid Bilayers. *Biophys. J.* **77**, 2462–2469 (1999).
72. David, C. & Jacobs, D. Principal Component Analysis: A Method for Determining the Essential Dynamics of Proteins. In *Protein Dynamics* (ed. Livesay, D. R.) 193–226, doi:[https://doi.org/10.1007/978-1-62703-658-0\\_11](https://doi.org/10.1007/978-1-62703-658-0_11) (Humana Press, 2014).
73. Shell, M. S. Coarse-Graining with the Relative Entropy. in *Advances in Chemical Physics* (eds Rice, S. A. & Dinner, A. R.) 395–441x, doi:<https://doi.org/10.1002/9781119290971.ch5> (Humana Press, 2014).
74. Foley, T. T., Shell, M. S. & Noid, W. G. The impact of resolution upon entropy and information in coarse-grained models. *J. Chem. Phys.* **143**, 243104 (2015).
75. Cooke, I. R., Kremer, K. & Deserno, M. Tunable generic model for fluid bilayer membranes. *Phys. Rev. E* **72**, 011506 (2005).
76. Martini Web Site. Available at: <http://md.chem.rug.nl/> (Accessed: 5th September 2016).
77. Padding, J. T. & Briels, W. J. Systematic coarse-graining of the dynamics of entangled polymer melts: the road from chemistry to rheology. *J. Phys. Condens. Matter* **23**, 233101 (2011).
78. Lyubartsev, A. P. & Laaksonen, A. Calculation of effective interaction potentials from radial distribution functions: A reverse Monte Carlo approach. *Phys. Rev. E* **52**, 3730–3737 (1995).
79. Tahir, M. A., Van Lehn, R. C., Choi, S. H. & Alexander-Katz, A. Solvent-exposed lipid tail protrusions depend on lipid membrane composition and curvature. *Biochim. Biophys. Acta BBA - Biomembr.* **1858**, 1207–1215 (2016).
80. Kyrychenko, A. Using fluorescence for studies of biological membranes: a review. *Methods Appl. Fluoresc.* **3**, 042003 (2015).
81. Holopainen, J. M., Lehtonen, J. Y. A. & Kinnunen, P. K. J. Evidence for the Extended Phospholipid Conformation in Membrane Fusion and Hemifusion. *Biophys. J.* **76**, 2111–2120 (1999).
82. Davis, P. J., Fleming, B. D., Coolbear, K. P. & Keough, K. M. Gel to liquid-crystalline transition temperatures of water dispersions of two pairs of positional isomers of unsaturated mixed-acid phosphatidylcholines. *Biochemistry (Mosc.)* **20**, 3633–3636 (1981).
83. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
84. Durell, S. R., Brooks, B. R. & Ben-Naim, A. Solvent-Induced Forces between Two Hydrophilic Groups. *J. Phys. Chem.* **98**, 2198–2202 (1994).
85. Neria, E., Fischer, S. & Karplus, M. Simulation of activation free energies in molecular systems. *J. Chem. Phys.* **105**, 1902–1921 (1996).
86. Yesylevskyy, S. O., Schäfer, L. V., Sengupta, D. & Marrink, S. J. Polarizable Water Model for the Coarse-Grained MARTINI Force Field. *PLOS Comput Biol* **6**, e1000810 (2010).
87. Wassenaar, T. A., Ingólfsson, H. I., Böckmann, R. A., Tieleman, D. P. & Marrink, S. J. Computational Lipidomics with insane: A Versatile Tool for Generating Custom Membranes for Molecular Simulations. *J. Chem. Theory Comput.* **11**, 2144–2155 (2015).
88. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865 (2008).
89. Lee, J. *et al.* CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **12**, 405–413 (2016).

90. Qi, Y. *et al.* CHARMM-GUI Martini Maker for Coarse-Grained Simulations with the Martini Force Field. *J. Chem. Theory Comput.* **11**, 4486–4494 (2015).
91. Wu, E. L. *et al.* CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. *J. Comput. Chem.* **35**, 1997–2004 (2014).
92. Brooks, B. R. *et al.* CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **30**, 1545–1614 (2009).
93. de Jong, D. H., Baoukina, S., Ingólfsson, H. I. & Marrink, S. J. Martini straight: Boosting performance using a shorter cutoff and GPUs. *Comput. Phys. Commun.* **199**, 1–7 (2016).
94. Ryckaert, J.-P., Ciccotti, G. & Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).
95. Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.* **81**, 511–519 (1984).
96. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
97. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
98. Wassenaar, T. A., Pluhackova, K., Böckmann, R. A., Marrink, S. J. & Tieleman, D. P. Going Backward: A Flexible Geometric Approach to Reverse Transformation from Coarse Grained to Atomistic Models. *J. Chem. Theory Comput.* **10**, 676–690 (2014).

## Acknowledgements

The study was supported by the Supercomputing Center of Lomonosov Moscow State University and JURECA supercomputer of the Research Center Jülich, and performed in the framework of ERA.Net RUS Plus grant (ID 323, Russian Federal Target Program ‘Research and Development’ contract 14.587.21.0011, RFMEFI58715X0011).

## Author Contributions

P.B. and I.G. performed the research and wrote the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at doi:10.1038/s41598-017-11761-5

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017