Supporting Information for Publication: Optimization of Slipids Force Field Parameters Describing Headgroups of Phospholipids

Fredrik Grote and Alexander P. Lyubartsev*

Department of Materials and Environmental Chemistry, Stockholm University, SE 106 91, Stockholm, Sweden

tel. +46-8161193

E-mail: alexander.lyubartsev@mmk.su.se

Quantum chemistry test calculations

The choice of quantum chemistry method was based on test calculations performed using different combinations of methods and basis sets, see Table S1. For parametrization of dihedral angles the crucial quantity to describe correctly is the relative energy between lipid conformations. We have chosen randomly four lipid headgroup conformations and computed energy difference between each pair of them by each of the method, evaluating also the computational time. It was found that accurate *ab-initio* methods (i.e. CCSD(T) and MP2) were too computationally expensive to be considered for $\sim$1000 single point energy calculations. In order to find a cheaper method that give accurate results we carried out single point energy calculations using DFT with B3LYP and B3P86 exchange-correlation functionals in combination with different basis sets as well as Hartree-Fock (HF) calculations. The accuracy of these cheaper methods was investigated by comparing energy differences between lipid conformations to the corresponding energy difference obtained using CCSD(T) and MP2. The result is shown in Figure S1. We found that DFT with different combinations of basis sets and exchange-correlation functionals gave energy differences between lipid conformations in better agreement with MP2 and CCSD(T) compared to HF. When comparing the DFT methods our results indicate that the B3P86 functional performs better than the B3LYP. For this functional we carried out calculations using the cc-pvqz and cc-pv5z basis set but since both gave similar results the smaller basis set (cc-pvqz) was chosen for the final calculations.

Table S1: Combinations of quantum chemistry methods and basis sets that were tested.

Method	Basis set
CCSD(T)	cc-pvqz
CCSD(T)	cc-pvdz
MP2	cc-pvqz
MP2	cc-pvdz
DFT B3P86	cc-pv5z
DFT B3P86	cc-pvqz
DFT B3LYP	cc-pv5z
DFT B3LYP	cc-pvqz
DFT B3LYP	cc-pvtz
HF	cc-pvqz

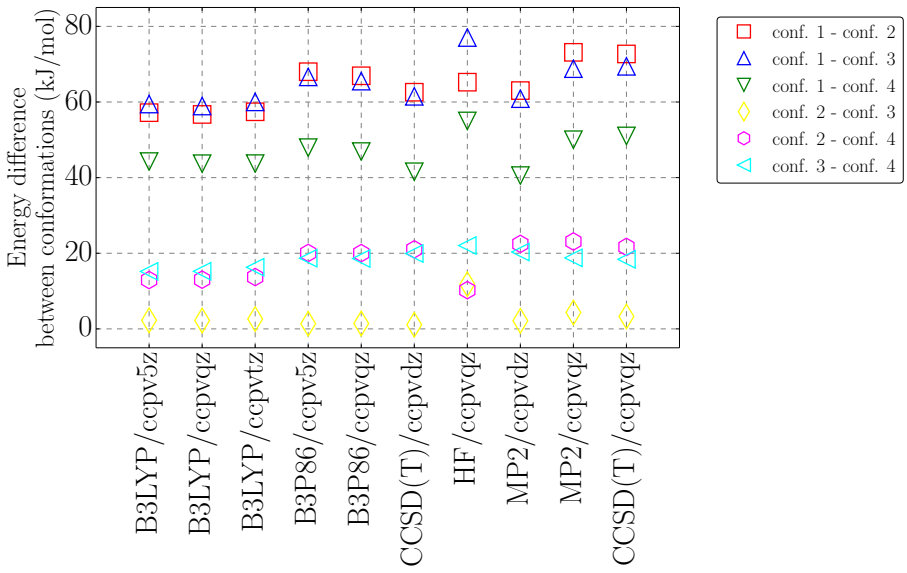

Figure S1: Energy differences between lipid conformations computed using different level quantum chemical theory.

Parameter fitting

In order to reduce overfitting we carried out regularized least square optimization for a range of regularization parameters (λ = 0,1,2,3,4 and 5). It was found that $\lambda = 2$ gives a Pearson coefficient of 0.644 and reduces the ratio of SSD for validation and training set from 1.16 to 1.13. Figure S2 shows scatter plots indicating the correlation between quantum chemical reference energies and force field energies fitted using different values of λ. Using $\lambda = 2$ reduces overfitting, by decrease of the SSD ratio, and gives a scatter plot similar to that obtained with $\lambda = 0$ (compare Figure S2 (a) and (b)). While a larger value for λ, e.g. $\lambda = 5$, reduces overfitting even more the scatter plot starts to deviate from that of $\lambda = 0$ (compare Figure S2 (a) and (c)). This gives confidence that $\lambda = 2$ is a reasonable compromise reducing overfitting while not significantly deteriorating the fit. Figure S3 shows a scatter plot indicating the correlation between force field energies and quantum chemical reference energies for the new and old parameter set. It is evident that the new parameters give better agreement with the quantum chemistry energy surface.

Parameter convergence

We optimized dihedral force field parameters using 2000 lipid headgroup conformations. In order to understand how the optimized parameters depend on the number of conformations used in the optimization, N, we fitted parameters for N ranging from 100 to 1000 in the training set and analyzed how the Pearson coefficient and ratio of SSD between validation and training set varied as function of N.

We observed that initial increase of the number of lipid fragments included in the optimization significantly decreases the Pearson coefficient and the SSD-ratio as shown in Figure S4. It was also found that increasing N beyond 700 does not change the quality of fit considerably and overfitting is no longer reduced. This gives confidence that the parameters have converged. However, we found that energy surfaces of individual dihedrals were fluctuating

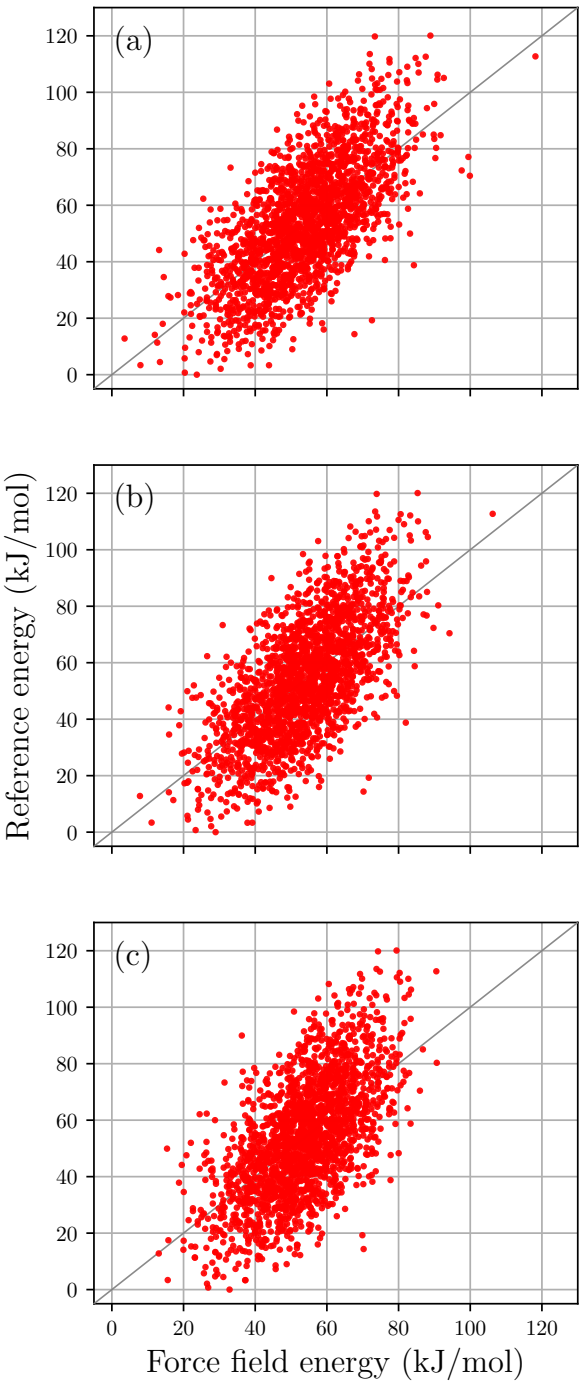

Figure S2: Scatter plot indicating the correlation between reference energies and force field energies fitted using (a) $\lambda = 0$, (b) $\lambda = 2$, (c) $\lambda = 5$

as function of N, see Figure S5. These fluctuations persist up to $N = 2000$ and can be attributed to limitations of our simple force field function for describing the more complicated quantum chemical energy landscape. It can be noted that, at least qualitatively, the shape

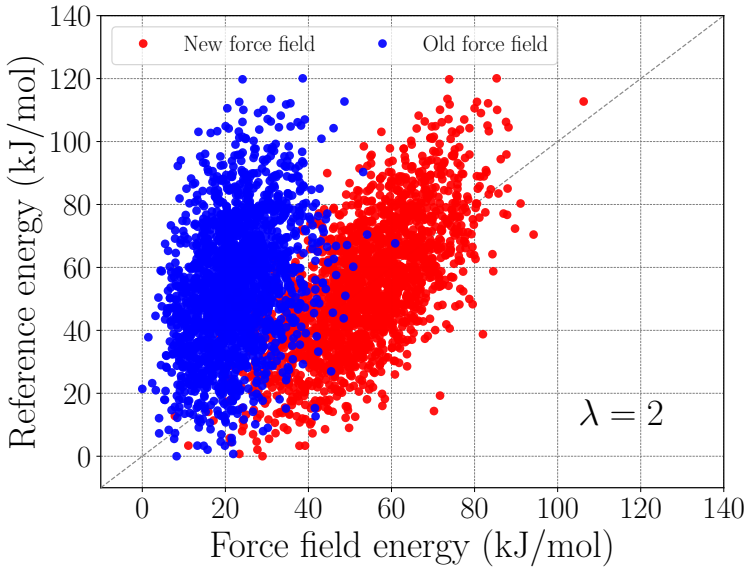

Figure S3: Scatter plot indicating the correlation between reference energies and force field energies for the old parameters and the new parameters fitted using $\lambda = 2$.

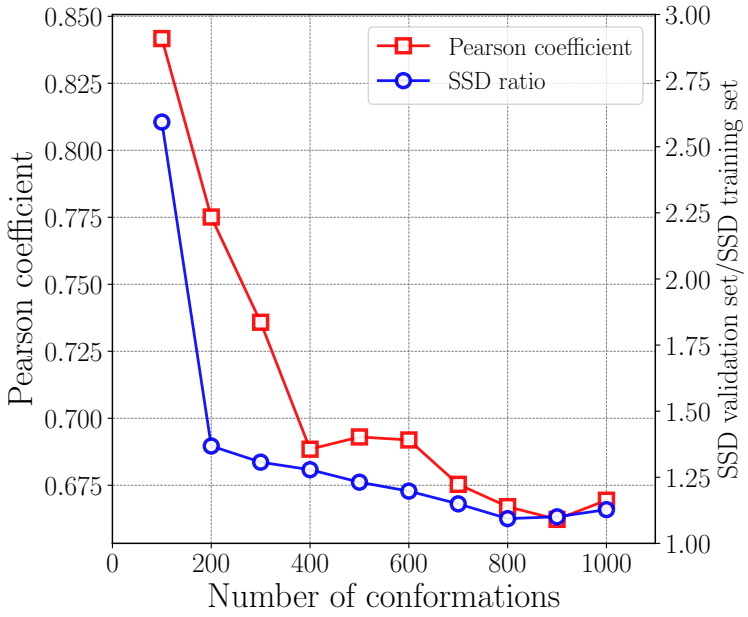

Figure S4: Convergence of SSD-ratio and Pearson coefficient with respect to the number of lipid conformations used in the optimization.

of the energy surfaces are essentially unchanged when increasing N from 1000 to 2000, i.e. although the exact level of minima and maximia change they remain high/low with respect to k_BT suggesting that the conformational distributions are not affected significantly. In order to obtain a better fit to the quantum chemical reference energies it would be necessary

to use a refined model for the dihedral terms in the force field including possible cross-terms. However, here we are limited to the standard functional forms that are implemented in MD codes.

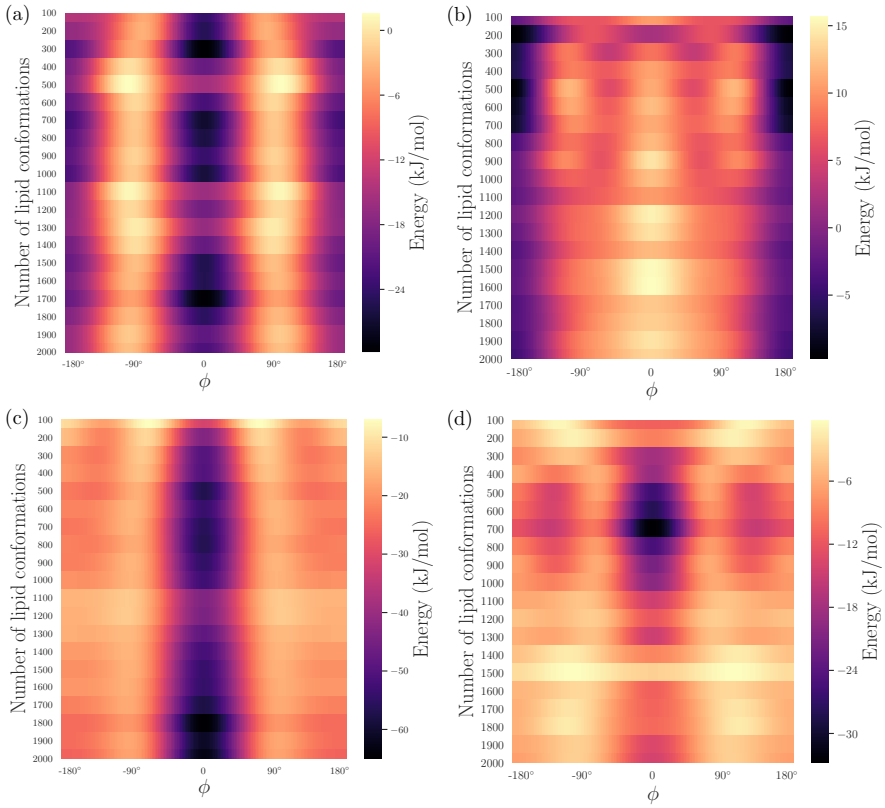

Figure S5: Energy surface as function of the number of lipid conformations; (a) PL-OSLP-CTL2-CTL1, (b) OSL-CTL1-CTL2-OSL, (c) CTL2-CTL1-OSL-CL and (d) CTL2-CTL1-CTL2-OSL.

Effect of cholesterol content

In addition to the three one-component bilayer systems we calculated order parameters for a POPC bilayer with different cholesterol content (20, 30 and 60 mol%). The result is shown in Figure S6. The effect on cholesterol on the headgroup order parameters is discussed in the section 'Cholesterol content' in the main text. Here we briefly discuss the results obtained for the order parameters in the lipid tails. It is evident that increase of cholesterol content leads

to ordering och C-H vectors in the lipid tails which is in agreement with NMR experiments.[1] It can also be noted that both the new and the old force field give order parameters in close agreement with each other.

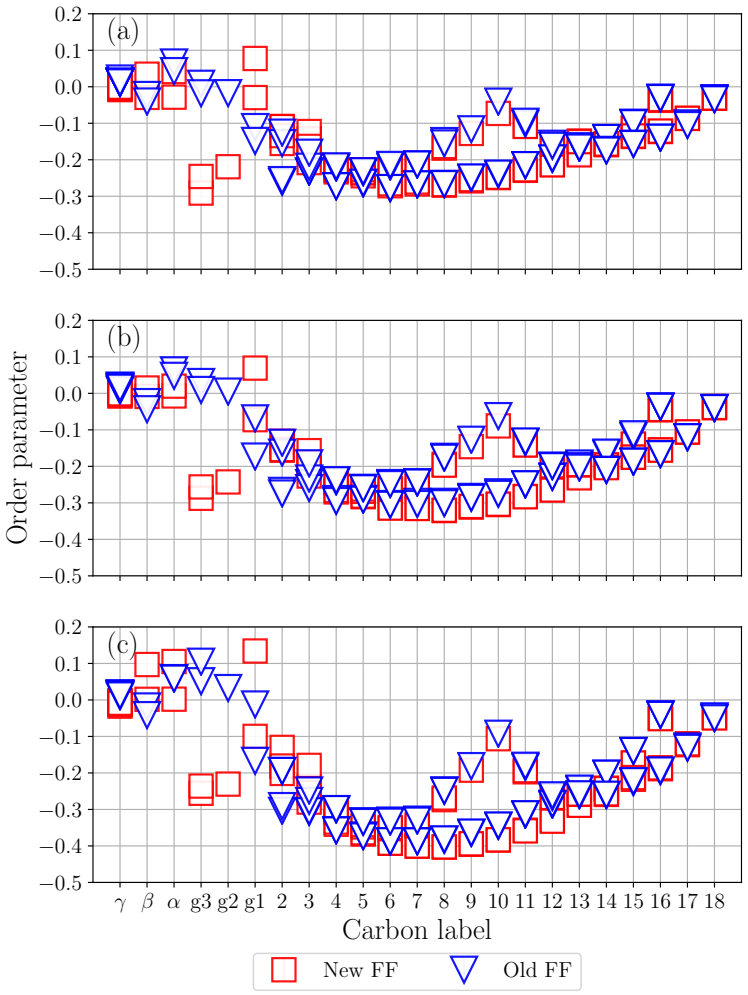

Figure S6: Order parameters for POPC simulated using the new and old force field at different cholesterol content (a) 20 mol%, (b) 30 mol% and (c) 60 mol%.

Effect of tuning Lennard Jones parameters

The fitted dihedral force field parameters underestimated the area per lipid by roughly 0.01-0.02 nm^2. This deviation is small but falls outside the experimental error bars. The area per lipid is sensitive to the effective radius of atoms which is determined by the Lennard Jones repulsion terms. We express the Lennard Jones potential according to Eqn. 1:

$$E_{LJ} = \frac{A}{r^{12}} - \frac{B}{r^6} \tag{1}$$

where $A = 4\varepsilon\sigma^{12}$ and $B = 4\varepsilon\sigma^6$. The parameter A defines the effective radius of the interacting atoms and B determines the strength of the dispersion attraction. In order to obtain better agreement with experimental area per lipid we therefore increased the Lennard Jones parameter A, keeping the value of B fixed in order to not change the dispersion interactions. We increased A for several different atoms in the lipid headgroup and investigated how the area per lipid was affected. It was found that increasing σ by 0.02 Åfor atom types CTL5 and HL and adjusting ε so that B terms remains unchanged gave better results for the area per lipid while keeping good agreement with other computed properties. The effect of this adjustment is shown in Figure S7 for POPC where Figure S7 (a) shows the area per lipid for the original Slipids force field and (b) and (c) show the result for the new force field before and after adjustment of Lennard-Jones parameters respectively. The improved agreement with experimental area per lipid is observed when comparing Figure S7 (b) and (c).

Gel phase

One important feature of the Slipids force field is that while it provides a good description of the biologically relevant L_α-phase it also correctly predicts formation of an ordered 'gel' phase at lower temperature.[2] In order to confirm that a gel phase is formed also with the new parameter set we reproduced the simulation of DPPC at 293 K from ref.[2] Figure S8 shows

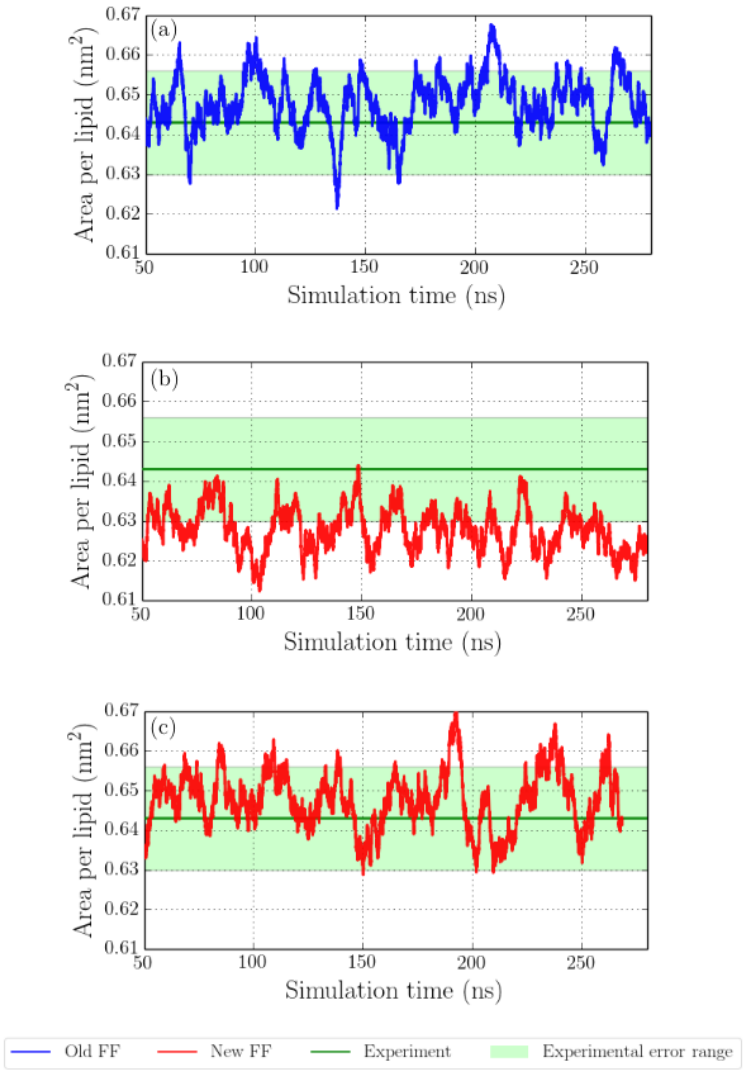

Figure S7: Area per lipid for POPC (a) Old FF, (b) New FF before tuning LJ parameters and (c) New FF after tuning LJ parameters.

area per lipid for DPPC as function of simulation time starting from the last configuration of our simulation carried out at 323 K. It is clear that when decreasing the temperature from 323 K to 293 K the area per lipid decreases to $\sim$0.49 nm^2. This is in agreement with the experimental value of 0.473 nm^2 that has recently been reported for the DPPC gel phase at 293 K.[3] Further evidence for the formation of a gel phase is provided by system snapshots shown in Figure S9. This gives some confidence that the new force field parameters are transferable to simulations of the gel phase and motivates a more detailed analysis of how well the new force field describes other lipid phases.

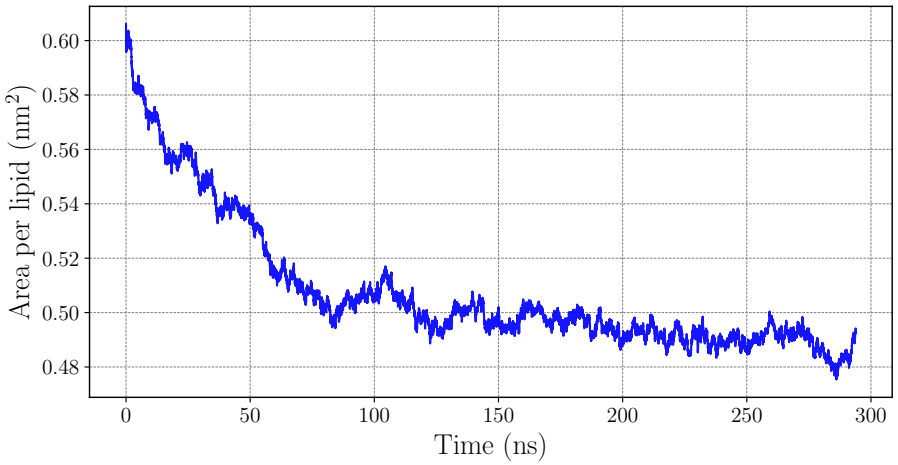

Figure S8: Area per lipid for simulation of DPPC at 293 K starting from the last frame of the simulation perofrmed at 323 K.

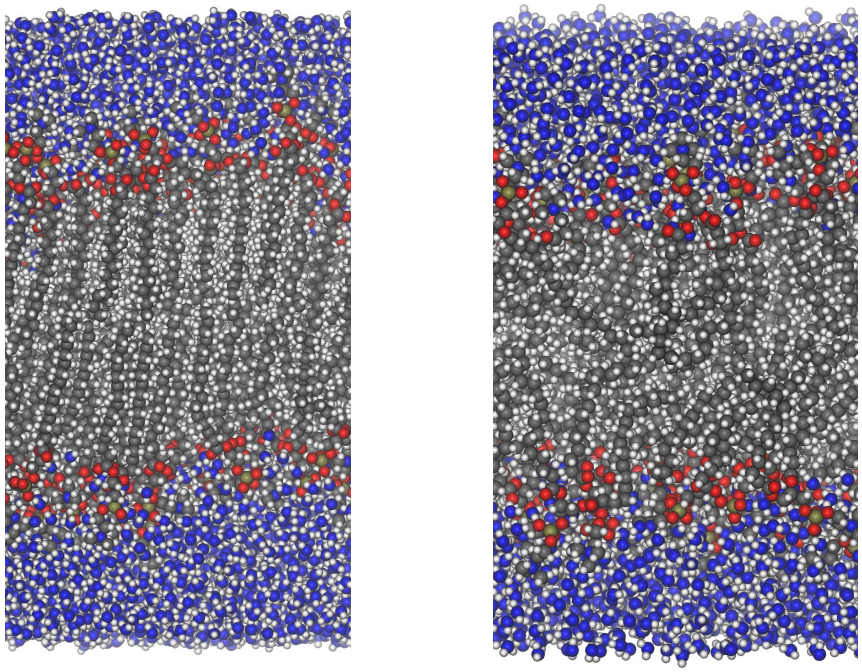

Figure S9: System snapshots for simulation of DPPC at (a) 293 K and (b) 323 K.

Effect of periodic boundary conditions on lateral diffusion coefficients

The lateral diffusion coefficients calculated for DMPC and POPC using the new force field underestimated the values measured by NMR in ref.[4] by a factor of 3.3 and 4.3 respectively.

This can be an artifact introduced by the periodic boundary conditions used in the simulations.[5] Venable et al.[6] have suggested an approach to correct for the periodic boundary condition effect based on the Saffman-Delbrük model for a single-leaflet spanning cylinder of radius R diffusing in a lipid bilayer of lateral size L immersed in a water layer of thickness $2H$. In this model the diffusion coefficient for an infinite non-periodic system, D^{∞}, and a system system under periodic boundary conditions, D^{PBC}, is described in terms of a set of parameters being the radius R of the cylinder, membrane viscosity η_m, interleaflet friction, b, and water viscosity, η_w. Using the online resource `https://diffusion.lobos.nih.gov` we calculated D^{∞} and D^{PBC} for R, H and L obtained from our DPPC MD simulations and the default values for the other parameters, i.e. $\eta_m = 5.47 \times 10^{-3}$ Poise, $\eta_w = 1.6 \times 10^{-7}$ Poise and $b = 10^7$ Poise/cm. We found that setting R to be consistent with the area per lipid gives $D^{PBC} = 4.95 \times 10^{-8}$ cm^2/s and $D^{\infty} = 1.52 \times 10^{-7}$ cm^2/s, i.e. $D^{\infty}/D^{PBC} = 3.1$. As pointed out in ref.[7] the effective hydrodynamic radius of the diffusing lipid can be smaller than what is predicted from the bilayer surface area. Setting $R = 0.9 \times \sqrt{A_{lipid}/\pi}$ gives $D^{\infty}/D^{PBC} = 2.9$ and $R = 0.8 \times \sqrt{A_{lipid}/\pi}$ gives $D^{\infty}/D^{PBC} = 2.8$. According to this model the diffusion under periodic boundary conditions is about three times slower compared to the infinite non-periodic system. This indicates that the diffusion coefficients reported in this work can be in closer agreement with the experimental values in ref.[4] than suggested by a direct comparison.

References

(1) Ferreira, T. M.; Coreta-Gomes, F.; Ollila, O. S.; Moreno, M. J.; Vaz, W. L.; Topgaard, D. Cholesterol and POPC Segmental Order Parameters in Lipid Membranes: Solid State 1 H–13 C NMR and MD Simulation Studies. *Phys. Chem. Chem. Phys.* **2013**, *15*, 1976–1989.

(2) 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* **2012**, *116*, 3164–3179.

(3) Nagle, J. F.; Cognet, P.; Dupuy, F. G.; Tristram-Nagle, S. Structure of Gel Phase DPPC Determined by X-ray Diffraction. *Chem. Phys. Lipids* **2019**, *218*, 168–177.

(4) Filippov, A.; Orädd, G.; Lindblom, G. Influence of Cholesterol and Water Content on Phospholipid Lateral Diffusion in Bilayers. *Langmuir* **2003**, *19*, 6397–6400.

(5) Camley, B. A.; Lerner, M. G.; Pastor, R. W.; Brown, F. L. Strong Influence of Periodic Boundary Conditions on Lateral Diffusion in Lipid Bilayer Membranes. *J. Chem. Phys.* **2015**, *143*, 12B604_1.

(6) Venable, R. M.; Ingólfsson, H. I.; Lerner, M. G.; Perrin Jr, B. S.; Camley, B. A.; Marrink, S. J.; Brown, F. L.; Pastor, R. W. Lipid and Peptide Diffusion in Bilayers: The Saffman–Delbruck Model and Periodic Boundary Conditions. *J. Phys. Chem. B* **2017**, *121*, 3443–3457.

(7) Zgorski, A.; Pastor, R. W.; Lyman, E. Surface Shear Viscosity and Interleaflet Friction from Nonequilibrium Simulations of Lipid Bilayers. *J. Chem. Theory Comput.* **2019**, *15*, 6471–6481.