Derivation and Systematic Validation of a Refined All-Atom Force Field for Phosphatidylcholine Lipids

Joakim P. M. Jämbeck* and Alexander P. Lyubartsev*

Division of Physical Chemistry, Arrhenius Laboratory, Stockholm University, Stockholm, SE-10691, Sweden

Supporting Information

ABSTRACT: An all-atomistic force field (FF) has been developed for fully saturated phospholipids. The parametrization has been largely based on high-level ab initio calculations in order to keep the empirical input to a minimum. Parameters for the lipid chains have been developed based on knowledge about bulk alkane liquids, for which thermodynamic and dynamic data are excellently reproduced. The FFs ability to simulate lipid bilayers in the liquid crystalline phase in a tensionless ensemble was tested in simulations of three lipids: 1,2-diauroyl-sn-glycero-3-phosphocholine (DLPC), 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC), and 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC). Computed areas and volumes per lipid, and three different kinds of bilayer thicknesses, have been investigated. Most importantly NMR order parameters and scattering form factors agree in an excellent manner with experimental data under a range of temperatures. Further, the compatibility with the AMBER FF for biomolecules as well as the ability to simulate bilayers in gel phase was demonstrated. Overall, the FF presented here provides the important balance about bulk alkane liquids, for which thermodynamic and dynamic data are excellently reproduced. The FFs ability to simulate lipid bilayers in the liquid crystalline phase in a tensionless ensemble was tested in simulations of three lipids: 1,2-diauroyl-sn-glycero-3-phosphocholine (DLPC), 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC), and 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC). Computed areas and volumes per lipid, and three different kinds of bilayer thicknesses, have been investigated. Most importantly NMR order parameters and scattering form factors agree in an excellent manner with experimental data under a range of temperatures. Further, the compatibility with the AMBER FF for biomolecules as well as the ability to simulate bilayers in gel phase was demonstrated. Overall, the FF presented here provides the important balance between the hydrophilic and hydrophobic forces present in lipid bilayers and therefore can be used for more complicated studies of realistic biological membranes with protein insertions.

INTRODUCTION

Biological membranes are crucial components of the cell. They divide the inner part of the cell from the outer environment as well as dictate the intercellular transport and cell fusion/division. Membranes also provide the matrix in which important proteins (transmembrane pumps, channels, receptors etc.) reside. In nature, biological membranes are very complex systems, consisting of a mixture of different lipids, sterols, and proteins. These constituting parts are involved in a number of processes making the cell membrane a very dynamic structure. It has been proposed that protein–lipid interactions regulate the functionality of proteins, making the lipid bilayer more than a passive hydrophobic slab. Most studies are, however, performed on pure phospholipid bilayers since phospholipids are a major part of biological membranes. Before one can study biological membranes in their true essence, the properties of simpler membranes must be fully understood. Experimental techniques used to study lipid bilayers span from X-ray and neutron scattering to IR/Raman and NMR.

Lipid bilayers have a very pronounced temperature-dependent phase behavior. The biologically most relevant phase for membranes is the liquid-crystalline phase (Lc-phase), which means that the aliphatic chains of the lipid have a certain degree of freedom which result in a disordered structure. Although a large number of experimental studies have been performed on membrane systems and already shed some light on the dynamical and structural properties, a more detailed understanding of these systems is desirable. Due to the difficulties of studying these systems in the correct phase, computer simulation methods are an ideal complement to experiments. Molecular dynamics (MD) computer simulations have become an essential tool for studying membrane on an atomistic level during the past two decades.

If a sufficiently large system is chosen to be simulated, the quality of the simulations is limited by two factors: (1) sampling or time scale and (2) the empirical potential energy function or force field (FF) employed. The sampling issue is always present and it has been found that fluctuations in simulations of pure lipid bilayers happen on a picosecond time scale. As a consequence of this, long simulations (>100 ns) are needed in order to study equilibrium properties of lipid bilayers. For systems with proteins embedded, much longer simulations may be needed. The other issue, i.e., the ability of the potential energy function to provide realistic dynamics and thermodynamic properties of the system, is extremely crucial. Two major types of FFs are available for lipid bilayers: united atom (UA) and all-atomic (AA). In UA FFs all nonpolar hydrogens have been included in the heavier atoms, creating pseudoatoms. The popular Berger lipids, the G53A6, FF, the 43A1-S3 FF, and the potential developed by Ulmschneider and Ulmschneider and Kukol are examples of these FFs. The AA FFs have explicit hydrogens included and the most common ones for lipid simulations are CHARMM and the general AMBER FF (GAFF).
Until now, these have been the most widely used AA FFs, although several problems have been reported regarding older versions of CHARMM and GAFF FFs. Other FFs are nonadditive (polarizable) or coarse-grained. In some cases, an AA description is preferred and recent studies have shown that there exists a unique designated hydrogen bond network around the choline head group. In order to describe and study this in full detail, an accurate AA model is crucial. AA description may also be crucial for describing order parameters of the glycerol group and in the description of double bonds in unsaturated lipids. An inaccurate FF often captures this effect. The final comparisons show that the FFs presented here reproduce the experimental findings very well and therefore is highly suited for simulations of lipid bilayers in the correct phase.

**METHODS AND MODELS**

**Parameterization Strategy.** The main focus of the parameterization procedure has been on the hydrophobic tails of the lipids, since these parameters were previously shown to have strong influence on simulated properties of bilayers, and the partial atomic charges for the whole lipid. As model compounds for the hydrophobic tails, a range of \( \text{n-alkanes} \) has been chosen (hexane to hexadecane). The assumption that the hydrophobic core behaves similar to a bulk alkane liquid has proven to be valid a number of times. If the fitting is done simultaneously for a series of alkanes, the results are generally better. Furthermore, studies on NMR relaxation rates have shown that the hydrophobic part of the lipid bilayer behaves similar to bulk alkane solutions.

The following potential energy function was chosen for this work

\[
U = \sum_{\text{bonds}} k_r (r - r_0)^2 + \sum_{\text{angles}} k_\theta (\theta - \theta_0)^2 + \sum_{\text{nonbonded}} k_q (\sigma_q - r_q)^6 - \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}
\]

which is a standard form for FFs like AMBER, CHARMM, and GROMOS. As a starting point for our parameterization, we took parameters of the CHARMM FF. Parameters for all covalent bonds and angles as well as LJ and torsional parameters for the lipid head group were taken from the FF described by Klauda et al. We then derived new parameters for the lipid head group from the parameters used for proteins to describe the same type of atoms. One possible explanation for this is the anisotropic nature of these systems. Lipid bilayers are very peculiar systems with a polar head group in contact with the aqueous environment and a strong attraction of the tails driven by the hydrophobic effect and van der Waals (vdW) forces. A complexity like this poses a challenge for the modeling community. Especially, the fact that in order to study complexes including both lipids and proteins we need FFs that work well together implies further complications. Previous work has shown that the interactions between amino acids and lipid molecules need to be carefully evaluated before embarking on larger simulations.
these configurations. Because of the known difficulties of reproducing the vdW dispersion interaction by ab initio methods, experimental heats of vaporization and densities were used during the fitting of the LJ parameters. First, the new charges were used with the original parameters from C36 (LJ and torsional) and the LJ parameters were then altered until satisfactory agreement between simulations and experiments was obtained. After this, the torsional potentials were fitted from ab initio computations for the model compound. After one round in the parametrization scheme, it was necessary to refit the LJ parameters again and the torsional potentials until self-consistency was obtained. The Lorentz–Berthelot combination rules were used for the vdW interactions.51 The introduction of scaling factors for the 1–4 interactions with the CHARMM FFs for lipids together with a new set of charges dependent but previous FFs have been parametrized from well-known that partial atomic charges are conformation chosen as a model compound for computing the charges. It is polarizable continuum with different dielectric constant (see solvent effects were modeled by placing the molecule in a quantum mechanical calculations were performed with the Kollman scheme60 by single-point calculations and fitted during electrostatic potential was sampled with the Merz correlation functional55 (RESP) with the DFT method using the B3LYP exchange–correlation functional53–58 and the cc-pVTZ basis set.59 The electrostatic potential was sampled with the Merz–Singh–Kollman scheme60 by single-point calculations and fitted during the two-stage procedure developed by Cornell et al.61 All solvent effects were modeled by placing the molecule in a polarizable continuum with different dielectric constant (see above) with the IEFPCM continuum solvent model.62,63 The quantum mechanical calculations were performed with the Gaussian09 software package64 and the RESP calculations were performed with the Red software.65 In subsequent molecular dynamics simulations, Coulombic 1–4 interactions were scaled by a factor of 0.8333.

The way the atomic charges have been calculated and used in MD simulations makes them compatible with the AMBER03 FF53 and since the charges in all AMBER FFs are derived from the RESP the lipid FF presented here is compatible with most members of the AMBER FF family. This is of importance since there is a growing interest in simulating membrane proteins in their native environment19,66 and also peptide partitioning in biological membranes.67,68 Ongoing work aims to clarify which AMBER biomolecular FFs that work sufficiently well together with the current parameters. A preliminary test of the compatibility of the lipid parameters and the AMBER03 FF is presented further down.

Boltzmann averaging over charges introduces temperature dependency on the charges and in order to see the impact of temperature, simulations with different temperatures (298, 303, 310, 318, and 325 K) were performed with hexadecane using the methodology described above. No explicit temperature dependence could be found over this range of temperatures, making the charges reliable and robust with respect to temperature, at least within the interval tested here (data not shown).

**Lennard-Jones Parameters.** By reproducing experimental densities correct σ-values for the LJ potential can be assigned. Heats of vaporization are important since they describe the interactions between the molecules which is primarily determined by the well-depth (ε) of the LJ potential. The average density is straightforward to obtain from MD simulations and the heat of vaporization was computed according to

\[ \Delta H_{\text{vap}} = \langle E \rangle_{\text{gas}} - \langle E \rangle_{\text{liquid}} + RT \]

where \( \langle E \rangle_{\text{gas}} \) was obtained as an average of eight separate 50 ns long, in vacuo simulations with different starting conformations taken from a condensed-phase simulation. A time step of 0.25 fs was used for the in vacuo simulations. Test simulations in condensed phase lasted for 5 ns and production simulations for 100 ns at 298.15 K with the partial atomic charges obtained in the previous step. The simulation length of the condensed phase is longer than usually considered but it is necessary since a variation of parameters during the fitting procedure can freeze the system and this transition can occur on a relatively long time scale.25 The LJ 1–4 interactions between the hydrogens of the lipid carbon chains were scaled by a factor of 0.5 and for the remaining interactions special 1–4 interactions were specified. In order to make sure that the alkanes intermolecular interactions are of the right magnitude and that the molecules have the correct size, the LJ parameters ε and σ were fitted simultaneously.

**Torsion Potential.** The potential energy surface (PES) for octane was scanned in order to obtain an accurate torsion potential for the hydrophobic tails of the lipids. The torsion angles of interest were CH3–CH2–CH2–CH3 and CH3–CH2–CH3–CH4. The relative energies between different conformations were obtained by employing the hybrid method for interaction energies (HM-IE).69 Applying this method makes it possible to estimate intermolecular and intramolecular interactions with the precision of very high level ab initio methods, namely CCSD(T)70 with a large basis set (LBS). Assuming that the effect of using a larger basis set is additive,
the following expression can be used to estimate the intermolecular energies

\[
E[\text{CCSD(T)/LBS}] = E[\text{CCSD(T)/SBS}] + E[\text{CCSD(T)/LBS}] - E[\text{CCSD(T)/SBS}] \\
\approx E[\text{CCSD(T)/SBS}] + E[\text{MP2/LBS}] - E[\text{MP2/SBS}] \\
\equiv E[\text{MP2: CC}]
\]

(3)

where SBS is the abbreviation for the small basis set. This energy was estimated at a CCSD(T)/cc-pVQZ level of theory; i.e., LBS and SBS are equivalent to cc-pVQZ and cc-pVDZ, respectively. The geometry of the molecules was optimized with the second order of Möller–Plesset perturbation theory (MP2)\(^a\) with the cc-pVDZ basis set and then single-point calculations were used to estimate the relative energies with the HM-IE scheme. It has been proven that MP2/cc-pVDZ gives reasonable geometries whereas it is necessary to use more exact methods for the relative energies.\(^b\)

Since longer carbon chains give lower rotational barriers\(^c\) and due to the fact that the number of medium-range 1,3-interactions increases the stability of the compounds,\(^d\) a slightly larger model compound (octane) was used than that used in parametrization of CHARMM.\(^e\) Different conformations were used during the scanning of the PES and used in the fitting procedure, resulting in a total number of roughly 200 conformations estimated at a CCSD(T)/cc-pVQZ level of theory. More detailed information regarding the fitting procedure can be found in the Supporting Information along with all FF parameters derived in this work.

**Testing of Newly Derived Parameters.** A number of properties were computed for the bulk alkane systems and several lipid bilayers (DLPC, DMPC, and DPPC) in order to test the quality of the newly derived parameters. For bulk alkane liquids, uncorrected \(D_{\text{rho}}\) and corrected \(D\) self-diffusion coefficients\(^f\) were determined as well as \(^1^3\)C longitudinal relaxation times \(T_2\) and the number of gauche bonds per molecule. The ability of the newly derived FF to reproduce experimental data for lipid bilayers was verified by computing a series of structural and dynamic properties: \(A_0\), volume per lipid \(V_0\), isothermal area compressibility modulus \(K_0\), thermal area expansivity \(\alpha_T \), and lateral diffusion coefficient. Three types of bilayer thicknesses were computed and compared to experiments: head-to-head distance \(D_{\text{HH}}\), the Luzzati thickness \(D_{\text{LB}}\), the hydrophobic core thickness \(2D_{\text{C}}\), and the two corresponding thermal contractivities, \(\alpha_{DB}^l\) and \(\alpha_{DC}^l\). Electron density profiles along the bilayer normal were obtained from simulations with the contributions from different groups (see Figure 1). From the total electron density, the scattering form factors, \(F(q)\), were obtained via Fourier transformation. Deuterium order parameters \(I_{\text{CD}}(q)\) were also computed and compared to experimental findings. More information regarding the computed properties for both alkane liquids and lipid bilayers can be found in the Supporting Information.

**Simulations Details.** All simulations were performed in the isothermal–isobaric ensemble, NPT, at a pressure of 1 atm. The pressure was held constant by using the Parrinello–Rahman barostat\(^g\) with a coupling constant of 10.0 ps with an isothermal compressibility of \(4.5 \times 10^{-5}\) bar\(^{-1}\). For the bulk liquids an isotropic pressure coupling was used and for the bilayer simulations a semi-isotropic pressure coupling scheme was used. The temperature was kept constant by the Nosé–Hoover thermostat\(^h\) with a coupling constant of 0.5 ps. The lipid bilayer and water were coupled separately to the thermostat. Long-range electrostatic interactions were treated by a particle-mesh Ewald scheme\(^i\) with a real-space cutoff at 1.4 nm with a Fourier spacing of 0.10 nm and a fourth-order interpolation to the Ewald mesh. Single-atom charge groups were used. van der Waals interactions were truncated at 1.5 nm and treated with a switch function from 1.4 nm. Long-range corrections for the potential and pressure were added.\(^j\) The inclusion of long-range corrections should eliminate the LJ cutoff dependency in the simulations. Due to the fact that lipid bilayers are inhomogeneous systems the method introduced by Lagüe et al.\(^k\) to add long-range corrections could be applied instead. Periodic boundary conditions were imposed in every dimension. A time step of 2 fs was used with a Leap-Frog integrator. The LINCS algorithm\(^l\) was used to freeze all covalent bonds in the lipid, and the analytical SETTLE\(^m\) method was used to hold the bonds and angle in water constant. The TIP3P water model\(^n\) was the water model of choice. The choice of water model can be explained by the fact that TIP3P is the default water model in major FFs such as AMBER and CHARMM and since one of the aims of the work presented here was to create a lipid FF compatible with AMBER this was a natural choice. Further, earlier work of Högborg et al.\(^o\) has shown that there is flexibility in the choice of water model for AA simulations of lipid bilayers. Atomic coordinates were saved every 1 ps and the neighbor list was updated every 10th step.

Bulk liquids were simulated with a simulation box consisting of 128 molecules for the larger alkanes and 256 for the smaller alkanes (hexane and heptane) at a temperature of 298.15 K. The lipid bilayer systems were prepared using the CHARMM-GUI\(^p\)\(^q\) with 128 lipids in total, 64 in each leaflet. In order to achieve proper hydration, 30 TIP3P water molecules were added per lipid. Three different lipid types were simulated, DLPC (12:0/12:0), DMPC (14:0/14:0), and DPPC (16:0/16:0). These system were investigated under a range of temperatures; see Table 1 for an overview of all simulations performed. All lipid bilayer systems were equilibrated for 40 ns before production runs were initiated which lasted for 300–500 ns. All MD simulations were performed with the Gromacs\(^q\)\(^8\)

| lipid | temp (K) | time (ns) |
|-------|----------|-----------|
| DLPC  | 303      | 400       |
|       | 323      | 400       |
|       | 333      | 400       |
| DMPC  | 303      | 400       |
|       | 323      | 400       |
|       | 333      | 400       |
| DPPC  | 293      | 300       |
|       | 323      | 500       |
|       | 333      | 500       |
|       | 353      | 500       |

\(^{\text{All simulations were performed with 128 lipids and 30 water molecules per lipid. Total simulation time: 4.2 \mu s.}}\)
The Journal of Physical Chemistry B

RESULTS AND DISCUSSION

Bulk Alkane Liquids. Our initial charge calculations implied that the major part of the alkanes have almost zero (less than 0.005) charge of carbons and hydrogens with small but noticeable dipole moments at the end of the chains. Therefore, charges were set to zero for the inner methylene groups (both for carbons and hydrogens). As a result of this, the performance gap between AA and UA FFs can be reduced significantly. The total charge on the outermost methylene groups was set to +0.033 which is balanced by the charge of −0.033 on the methyl groups. The partial atomic charges used during the simulations of bulk alkane solutions were the same as presented for the lipid tail in Figure 1. The major parts of the

![Figure 1. Atomic charges calculated via the RESP method for the whole lipid. Charges within parentheses are the charges on the hydrogen(s) attached to the heavier atom. The second column corresponds to the groups used in Figure 7. α, α, and α are torsional angles that are discussed in the Results section.](image-url)

The final LJ parameters for methylene and methyl groups are shown in Table 2 and the final torsional parameters in Table 3.

### Table 2. Final Values for the Optimized Lennard-Jones Parameters in This Work

| atom type | σ (nm) | ε (kJ mol⁻¹) |
|-----------|--------|-------------|
| C (CH₃)   | 0.358  | 0.228       |
| H (CH₃)   | 0.240  | 0.112       |
| C (CH₂)   | 0.350  | 0.340       |
| H (CH₂)   | 0.220  | 0.095       |

### Table 3. Values for the Optimized Torsional Potentials

| torsion | force const, k_b (kJ mol⁻¹) | mult., n | phase shift, δ (deg) |
|---------|-----------------------------|---------|----------------------|
| CH₃−CH₂−CH₂−CH₃ | 0.0603 2 0 | | |
| CH₂−CH₂−CH₂−CH₃ | −0.0835 2 0 | | |

The same number of terms in the Fourier sum with the same multiplicity and phase shift as used in C36 gave the best result. The results of computations with the final optimized parameter set for heat of vaporization, density, and diffusion are summarized in Table 4. The linear trend in ΔH_vap is reproduced perfectly by the simulations with only a slight overestimation for the longest alkanes resulting in a rms error of only 0.9%. Further, the rms error for the density is 0.8% which means that the correct balance between size and interactions in the FF is obtained. During the parametrization procedure the alkane system can easily freeze due to incorrect parameters in a multiananosecond simulation. A too high value of ε combined with too small atomic sizes (σ) gave rise to a stacked lamellae phase which was also observed in the present investigation. This behavior will affect the FFs ability to describe a lipid bilayer in the L₂-phase at the correct temperature in a tensionless ensemble. The freezing sometimes occurred after only a couple of nanoseconds, making it straightforward to avoid this problem. In some cases, longer simulations were needed just as it was for Ulmschneider and Ulmschneider. Due to the possibilities of these phase transitions, all simulations were run for 100 ns.

As the thermodynamic properties for the alkanes converged and resulted in good agreement with experimental data, the FFs ability to reproduce dynamic properties such as diffusion and NMR relaxation times was investigated. Since the FF only has been parametrized against thermodynamic data, it is of utmost interest to test the dynamics of the alkanes. The uncorrected diffusion coefficients, D_{PBC}, in Table 4 are all underestimated due to the finite box size in the simulations. Once corrections for the hydrodynamics and long-range effects were added, the rms error decreases from 20% to 8.6%. Similar trends have been reported by Klauda et al. and Davis et al. Figure 2 shows the ¹³C NMR T₂ relaxation times obtained from simulations compared to experiments (in the Supporting Information more details on how to obtain ¹³C NMR T₂ from simulations are discussed). The dynamics of the chain is important since it will affect how the inner part of the bilayer behaves when
The statistics of different conformations of alkane chains were also investigated by computing the fraction of gauche conformations in condensed phase. In Table 5 the average number of gauche and trans bonds per alkane molecule are presented. The experimental estimation of 35% gauche conformations per molecule\(^99\) is reproduced; however, this comparison should be done with some care since the interpretation of the experimental data is not trivial. Still, it is reassuring that a high enough number of gauche conformations occur since in some previous FFs the trans conformations have been too pronounced, resulting in lipid bilayers that are too contracted when simulated under zero surface tension.\(^{31}\)

Transfer free energies from a polar environment to an apolar environment have been used earlier to investigate the compatibility of lipid–protein FF combinations.\(^{92,100}\) It has been shown that this scale, which is referred to as the Radzicka scale,\(^{101}\) correlates with studies where the same amino acid analogues have been transferred from aqueous solution to an explicit membrane environment.\(^{102–104}\) Similar approaches have also been used in the parametrization of the GROMOS FF.\(^{105}\) The computed free energy of transfer, \(\Delta \Delta G\), together with free energy of solvation in water and cyclohexane are presented in Figure 3. The free energies of hydration indicate that the AMBER03 FF (used to describe amino acids) is not very well compatible with the FFs used to describe solutes and proteins.\(^{92,106,107}\) This is a general feature of most major FFs developed for proteins.\(^{92}\)

The rms values in \(\Delta G_{\text{hyd}}\) are 4.86 kJ mol\(^{-1}\) which does not differ much from the rms values obtained by Shirts et al.\(^{108}\) and Sapay and Tieleman.\(^{92}\) The largest errors are for HIE (7.77 kJ mol\(^{-1}\)), ASN (6.61 kJ mol\(^{-1}\)), and MET (6.93 kJ mol\(^{-1}\)) which also correlates with earlier work.\(^{92,106,108}\) Computed free energies of solvation of the amino acid analogues in cyclohexane are in better agreement with experimental data\(^{101}\) and have a rms error of 2.76 kJ mol\(^{-1}\). Largest errors in \(\Delta G_{\text{cyc}}\) are obtained for TRP (5.68 kJ mol\(^{-1}\)), CYS (4.86 kJ mol\(^{-1}\)), and HIE (4.34 kJ mol\(^{-1}\)).
The transfer free energies also are in good agreement with experimental findings with a rms error of 3.40 kJ mol\(^{-1}\). Some error cancellation also occurs, e.g. for MET, HIE and TRP. The residues that contributes mostly to the error are GLN (6.75 kJ mol\(^{-1}\)), ASN (6.26 kJ mol\(^{-1}\)), and PHE (5.14 kJ mol\(^{-1}\)). Overall, the free energies agree well with experiments and the partitioning of amino acid analogues between the polar and apolar solvents suggests that the LJ parameters presented here are compatible with the AMBER03 FF. Since there are no intermolecular interactions of Coulombic nature present in the cyclohexane simulations, it is likely that other AMBER FFs will work together with the FF presented here. The data presented here shows that despite the fact that a different cutoff scheme was applied than the one used in the derivation of AMBER03, the two parameter sets are very compatible. For future work, we will test explore this compatibility further by performing simulations of membrane-bound proteins in their native environment. However, one should bear in mind that this hydrophobicity scale is missing some crucial points as protein backbone contributions etc. and therefore is only a crude measurement.

**Lipid Bilayer Simulations.** Atomic Charges for the Lipid Head Group. Partial atomic charges for the head group of the lipid were Boltzmann-averaged over a number of conformations generated from a 20 ns MD simulation and are presented in Figure 1. The sum of the individual atoms in the groups agree very well with the work of Högborg et al.\(^{31}\) and Sonne et al.\(^{30}\) which both describe AA FFs. The charge distribution in the head group alter the dipole moment. When using the popular Berger FF.\(^{23}\)\(^{5}\) the charges presented in Figure 1 (\(\alpha_0\), \(\alpha_1\), and \(\alpha_3\)) are all in gauche conformations which agrees with the experimental findings of Akutsu and Nagamori.\(^{111}\) This means that the intramolecular interactions in the lipid head group are described well and that its conformation is not fully extended. This might be important when studying membrane-bound proteins and the interactions between small solutes and lipid bilayers since the conformations of the head group alter the dipole moment. When using the popular Berger lipids, the head group is in a more extended conformation which is in disagreement with experiments.\(^{111,116}\)

Other parts of the head group that are important are the charges describing the electron distribution of the glycerol/carbonyl region. Previous studies have shown that the dipole moments of the ester groups correlate with the area per lipid, where a stronger dipole results in a higher area per lipid.\(^{23}\)\(^{47}\) A more polar ester group results in a higher water density in this region which in turn increases the area per lipid. In the C36 FF the polarity of this group was increased and it has been proposed that this polarity is one of the reasons for the success of the popular Berger FF.\(^{23}\)\(^{5}\) The charges presented in Figure 1 are similar to previously suggested charges.\(^{30}\)\(^{31}\) Interestingly, the ester oxygen bears almost the same charge in the present FF as the sn-1 ester oxygen in the charge set proposed by Sonne et al.\(^{30}\)

**Area and Volume Per Lipid and Isothermal Area Compressibility Modulus.** Area per lipid for DLPC, DMPC, and DPPC at various temperatures is presented in Table 6. The
simulations using the FF described here give values that are in excellent agreement with experiments for these lipids. It should be mentioned that experimental data have been reported in a wide range, see, e.g., Table 1 in ref 47. In Table 6 the simulated 
A\textsubscript{T} are compared to the most recent experimental data\cite{117,118}.

Generally, area per lipid is one of most computed properties from simulated lipid bilayers but the difficulties in obtaining accurate experimental data\cite{10} make these comparisons not conclusive. It has been proposed that the area per lipid alone cannot be an indicator of the quality of a FF\cite{18} and the summary of published experimental and simulation values needs to be used for FF validation.\cite{10,47} Therefore, other properties than just area per lipid must be computed and verified against experiments and only a comparison with an ensemble of different experimental values can really verify the FF and the methodology used. Below, such a comparison for a range of temperatures for the simulated lipid bilayers is presented. Despite the above arguments, area per lipid can still be used to indicate whether or not the lipid bilayer ends up in the correct phase. Incorrect parametrization may lead to underestimated values indicating that the lipid bilayer is in the L\textsubscript{\beta}-phase.\cite{23,30,31}

This is not the case in the present investigation, which can be seen from Table 6 showing area per lipid corresponding to L\textsubscript{\alpha}\textsuperscript{'}-phase for the range of temperatures, in good agreement with experimental results.

Volume per lipid is a property which can be more rigorously defined, and the interpretation of these properties from experiments is more straightforward compared to area per lipid. In Table 8 these values are compared to experimental data. Values obtained from simulations are systematically underestimated when compared to experiments. It is likely that this lack of consensus is due to the parameters for the head group since the agreement between simulations and experiments for bulk alkanes in Table 4 is very good. This indicates future improvements of the parameter set. The isostructural area compressibility is a property that can be difficult to obtain accurate values for in MD simulations. Often the variance in area per lipid is underestimated resulting in a too high value of K\textsubscript{A}\cite{14,45}. As can be seen from Table 8, the agreement between simulations and experiments is very good. Since the fluctuations in area per lipid occur on a long time scale,\cite{18,47} relatively long simulations are needed and therefore the simulations were run for 400–500 ns. As a consequence of this, fluctuations on several time scales can occur and contribute to K\textsubscript{A}. This results in computed values of K\textsubscript{A} that agree well with experiments. However, in a small system thermal fluctuations could be suppressed resulting in a size dependency\cite{119} since in a larger system undulations make the bilayer more compressible.\cite{45}

Poger and Mark studied the impact of system size by using lipid bilayers consisting of 128 and 722 lipids.\cite{47} It was found that K\textsubscript{A} for both systems converged to the same value, but roughly double the sampling time was required for the smaller system to reach the value of the larger. Possibly, longer simulations could have made the difference between the computed K\textsubscript{A} and the experimental data even smaller. Based on these conclusions, the system-size dependencies of the systems studied here were not explored.

Computed thermal area expansivities are shown in Table 7 and compared to experimental findings. The values obtained from simulations are systematically underestimating this property, which means that the lipid bilayer does not have the same tendency to expand as the temperature is increased. However, the disagreement is not significant and the relative values are reproduced.

**Bilayer Thickness.** Three types of membrane thicknesses were analyzed: the distance between the phosphate peaks in the electron density, the Luzzati thickness\cite{8} (2D\textsubscript{L}), and the hydrophobic thickness (2D\textsubscript{H}). D\textsubscript{H} is calculated according to\cite{2,120}

\[
D_H = d_z - \int_{-d_z/2}^{d_z/2} \rho_w(z) \, dz
\]

(4)

where d\textsubscript{z} is the repeat z-spacing along the the bilayer normal of the simulation box, i.e., the z-dimension of the simulation box. \(\rho_w(z)\) is the probability distribution of water in the z-direction.
and is calculated from the time-averaged histograms of the water distribution along the z-axis with a thickness of $d$: 

$$\rho_w(z) = \frac{n_w V_w}{dV}$$  \hspace{1cm} (5)

where $n_w$ is the time average of water molecules per slice and $dV$ is the time-averaged volume of each slice. $2D_C$ is computed in a similar way to $D_B$:

$$2D_C = \int_{-d_z}^{d_z} \rho_{CH}(z) \, dz$$  \hspace{1cm} (6)

where $\rho_{CH}(z) = \rho_{CHL} + \rho_{CHC}$.

The probability distributions are calculated according to

$$\rho_{CHL} = \frac{n_{CHL} V_{CHL}}{dV}$$  \hspace{1cm} (7)

As can be seen from Table 8, the agreement between simulations and experiments is excellent for the phosphate−phosphate distance for all three lipids. Consistent with previous work, the Luzzati thickness is slightly underestimated,\textsuperscript{22,120} The probability distributions used are reproducible. As a result of this, the Luzzati thickness is slightly underestimated whereas for DLPC it is underestimated to a small degree. The agreement for the bilayer thickness contractivity in Table 7 is presented.

| lipid | $T$ (K) | $V_L$ (nm$^3$) | $K_s$ (mN m$^{-1}$) | $D_{HH}$ (nm) | $D_B$ (nm) | $2D_C$ (nm) |
|------|--------|----------------|-------------------|--------------|-----------|-------------|
| DLPC | 303    | 0.951          | 268 ± 24          | 3.01         | 3.08\textsuperscript{122} | 3.04        | 3.14\textsuperscript{122} | 2.06 | 2.00\textsuperscript{122} | 2.10\textsuperscript{117} |
|      | 323    | 0.974          | 223 ± 31          | 2.95         | 3.00      | 3.10\textsuperscript{117} | 2.00 | 2.11\textsuperscript{117} |
| DMPC | 303    | 0.974          | 226 ± 33          | 2.92         | 2.99      | 3.07\textsuperscript{117} | 1.98 | 2.06\textsuperscript{117} |
|      | 323    | 1.110          | 255 ± 32          | 3.35         | 3.41      | 3.52\textsuperscript{117} | 2.39 | 2.54\textsuperscript{122} |
| DPPC | 303    | 1.120          | 230 ± 20          | 3.29         | 3.31      | 3.42\textsuperscript{117} | 2.32 | 2.56\textsuperscript{117} |
|      | 323    | 1.201          | 238 ± 35          | 3.77         | 3.80\textsuperscript{120} | 3.83        | 3.90\textsuperscript{117,120} | 2.80 | 2.84\textsuperscript{117,120} |
|      | 333    | 1.209          | 234 ± 42          | 3.71         | 3.70      | 3.81\textsuperscript{117} | 2.74 | 2.85\textsuperscript{117} |
|      | 353    | 1.223          | 211 ± 22          | 3.63         | 3.59      | 3.61         | 2.61 | 2.60\textsuperscript{117} |

The decrease in bilayer thickness with increasing temperature can be explained by the increase in chain distortion.\textsuperscript{117} The differences observed. As the chain disorder increases, the lipid bilayer becomes more compact in the direction of the membrane normal (the z-direction). In an expansion of the lipid bilayer in the $xy$-plane, the area is increased and the thickness has to be decreased. In Table 9 the number of gauche bonds for the DPPC lipid simulated in a bilayer at various temperatures are presented.

| lipid | $T$ (K) | $V_L$ (nm$^3$) | $K_s$ (mN m$^{-1}$) | $D_{HH}$ (nm) | $D_B$ (nm) | $2D_C$ (nm) |
|------|--------|----------------|-------------------|--------------|-----------|-------------|
| DLPC | 293    | 0.8            | 3.9               | 4.0          | 4.2       |
|      | 323    | 3.6−4.2\textsuperscript{130} | 3.7\textsuperscript{131} | 3.8\textsuperscript{126} |
| DPPC | 333    | <1.0 (300 K)\textsuperscript{133} | 3.9               | 4.0          | 4.2       |
|      | 353    | 3.6−4.2\textsuperscript{130} | 3.7\textsuperscript{131} | 3.8\textsuperscript{126} |
properties of a set of FF parameters for which to get accurate values. Computed deuterium order parameters ($|S_{CD}|$) together with experimental data for the aliphatic tails of DLPC, DMPC, and DPPC are presented in Figure 5. Values obtained from simulations agree very well with experiments. The splitting of order parameter for the second carbon in the $sn$-2 tail observed in experiments is also present in the simulations for each lipid type, however, not to the same extent. Experimental studies have shown that the upper part of the two lipid tails are aligned differently with the $sn$-1 tail in a more perpendicular position relative the bilayer normal than the $sn$-2 tail. As can be seen in Figure 5, this difference is also present in the simulations for all lipids. The agreement in order parameters indicates that the simulated lipid bilayers have the correct internal (disordered) structure and that the bilayers are all in the $L_α$-phase. A large discrepancy between experimental findings and simulations where $|S_{CD}|$ would have been overestimated by the latter would indicate a lipid bilayer in a too ordered phase, i.e., the $L_β'$-phase.

The number of gauche bonds in the aliphatic chains of the lipids can be computed in order to further test if the lipid bilayer is in the correct phase during the simulations. In Table 9 the number of gauche bonds from simulations of a DPPC lipid bilayer is compared to experiments. Simulations performed at 323 K agree well with the experimental estimates, and this together with the calculated deuterium order parameter and area per lipid indicates that the lipid bilayer is in the correct phase. A higher percentage of trans conformations (and low number of gauche bonds) induces order in the system resulting in a gel phase. In Figure 6 three simulation snapshots are shown at different temperatures where the increase in gauche bonds per lipid molecule and general disorder with increasing temperature are well illustrated.

**Electron Density Profiles and Scattering Form Factors.** In MD simulations it is straightforward to obtain electron density profiles and it is possible to divide the total electron density into the constituting groups of the lipid molecule as defined in Figure 1. The overall electron densities for the studied lipids are presented in Figure 7 together with the contributions from each group. From these plots the head-to-head distance ($D_{HH}$) can be obtained (see previous sections). Other valuable information such as the fact that water molecules penetrate the lipid bilayer down to the carbonyls of the aliphatic tails can also be extracted, which is in agreement with experiments. Furthermore, it is evident that water does not penetrate the hydrophobic core of the lipid bilayer, in agreement with experimental findings. For all lipids, the leaflets are interdigitated and the peaks corresponding to the head groups are broad, indicating that they are all in the $L_α$-phase since in gel phase the leaflets are less interdigitated and the head group’s electron density peaks are divided into two sharp peaks instead of one broad peak.

One inherent problem with experimental electron density curves is that they are all based on certain assumptions and models which makes direct comparison between simulations...
Figure 7. Electron density profiles from simulations of DLPC (top), DMPC (middle), and DPPC (bottom) lipid bilayers. Contributions from individual components are shown with colors: N, choline groups, dark blue; P, phosphate groups, cyan; CO, ester groups, red; CH₂, methylene groups, light gray; CH₃, methyl groups, dark gray.

and experiments difficult. For this reason, the densities in Figure 7 were not compared to experiments. Electron density profiles are obtained from an inverse Fourier transformation of the scattering form factors. A better way of comparing simulations and experiments is to compute the form factors from the total electron density profiles obtained from the simulations. These form factors can then be directly compared with the experimental data. This comparison together with the order parameters should serve as the major part of the testing procedure when the quality of a FF is established. F(q) and SCD (at least for the lipid tail) can be obtained from experiments in a unambiguous manner and are reproducible and therefore should receive the most focus. In Figure 8 form factors obtained from MD simulations are compared to experimental form factors obtained from simulations compared to experiments. The SDP model introduced by Kucera et al. was used to compute the FFs. As can be seen, the agreement is good and the points where the form factors have their minima are reproduced by the simulations. These findings extend the evidence that the FF presented here is highly suitable for simulating lipid bilayers.

Lateral Diffusion. In membranes the lipids are free to diffuse in the plane that is perpendicular to the bilayer normal. To test the parameter set proposed here, lateral diffusion coefficients, Dₜ, were therefore computed. The Einstein relation was used to calculate Dₜ via the mean-squared displacement (MSD).

\[ \text{MSD}(t) = \frac{4Dₜ}{3}t \]

In order to correct for the artificial center of mass motion of each monolayer, this contribution was subtracted during the course of the MSD calculations. The diffusion of lipids in a bilayer occurs on two different time scales: on a short time scale (<5 ns) where the center of mass of the lipid is in motion due to the conformational freedom of the acyl chains and the head group is relatively fixed in space, and on the longer time scale the whole lipids move in the monolayer more freely. In the present investigation, only the long time diffusion was investigated. The linear fitting was performed from 50 to 350/450 ns. The final results are shown in Table 10. For DLPC, DMPC, and DPPC Dₜ is underestimated with respect to the experimental data. The values obtained for DMPC at 303 K are close to other values obtained from MD simulations. The drastic decrease in Dₜ observed for DPPC when the temperature is lowered is expected due to the phase transition. As in the latter case the lipids are in a gel phase, the configurational freedom for the acyl tails is very limited, resulting in a very low diffusion coefficient. A large drop in lateral diffusion (about 2 orders of magnitude) when going from the Lβ'-phase to the Lα'-phase has been reported earlier by Marrink and co-workers in coarse-grained simulations. The increase in the lateral mobility of the lipids with an increase in temperature can be observed which is related to the area per lipid. As the temperature is increased, the area per lipid is increased, making it easier for the lipid to move in the two-dimensional liquid. For DMPC, the lateral diffusion coefficients obtained from simulations when the temperature is increased are about half of the experimental values as can be seen in the fifth and sixth columns of Table 10. In order to enhance Dₜ accelerated MD simulations could be applied where the sampling of the conformational space is improved. This approach has been used earlier in lipid bilayer simulations. One possible explanation for the systematic underestimation of the lateral diffusion could be the system size. In a lipid bilayer patch of 128 lipids, no efficient collective movement can occur that could result in a decrease in Dₜ. This is similar to the case of bulk alkanes when the diffusion
coefficients computed in a box of limited size were underestimated, but after correction on collective motion they occur in very good agreement with experimental results (Table 4). The agreement between simulations and experiments for the normalized values (with respect to the highest temperature) in Table 10 supports this, indicating that there could be a systematic underestimation which could originate from this effect.

**Fully Hydrated Gel Phase.** To further test the FFs ability to describe the nature of a lipid bilayer under various thermal conditions, simulations of a DPPC bilayer in gel phase were conducted. The phase transition from the \( \text{L}_\alpha \)-phase to the \( \text{L}_\beta' \)-phase occurs at 314 K for DPPC, so the simulations were performed at a temperature of 293 K in order to ensure that the lipid bilayer entered the correct phase. In order to prepare the gel phase, the size of the carbon atoms for the special 1-4 LJ interactions was increased so that the bilayers entered the gel phase without any major problems of metastable configurations. For the production run, the 1-4 LJ interactions were set to the original values with an equilibration time of 40 ns and a production simulation of 300 ns with the same procedure as for the bilayers in the \( \text{L}_\alpha \)-phase. This approach to force a lipid bilayer into the \( \text{L}_\beta' \)-phase is similar to the methodology described by Schubert et al. Simulations where the temperature was slowly decreased to 293 K without the biased 1-4 interactions resulted in a lipid bilayer in the rippled phase (\( \text{P}_\beta' \)). As this is a metastable state at 293 K the system would eventually reach the more stable \( \text{L}_\beta' \)-phase. Since the dynamics of the system is slowed down by roughly 2 orders of magnitude, biased intramolecular interactions were introduced merely to force the lipid bilayer into the correct phase on a shorter time scale. Thus, we can conclude that the correct phase behavior is to be expected. We believe that if the parameter set would have given rise to a more fluid phase than \( \text{L}_\beta' \) at 293 K this would have been observed during the 300 ns production simulation since melting is a process that generally occurs on a smaller time scale than freezing. In Figure 6 the obtained gel phase, \( \text{L}_\beta' \), is illustrated where the lipids in each monolayer are parallel.

In Table 11 the results of the simulation are summarized. It is evident that the bilayer entered the gel phase and that

| lipid | temp (K) | \( D_l \) \( \times 10^{-8} \text{ cm}^2 \text{ s}^{-1} \) | normalized \( D_l \) | sim | expt |
|-------|---------|----------------------------------|--------------------|------|------|
| DLPC  | 303     | 6.66 ± 0.45 \[8.5^{140,142}\] | 0.249              |      |      |
|       | 323     | 14.3 ± 3.50                      | 0.536              |      |      |
|       | 333     | 26.7 ± 4.70                      | 1.000              |      |      |
| DMPC  | 303     | 5.22 ± 0.49 \[5.95^{17}, 9.00^{140,141}\] | 0.332 | 0.291 |
|       | 323     | 11.8 ± 2.10 \[22.3^{151}\] | 0.752 | 0.722 |
|       | 333     | 15.7 ± 3.10 \[30.9^{151}\] | 1.000 | 1.000 |
| DPPC  | 293     | 0.15 ± 0.08                      | 0.005              |      |      |
|       | 323     | 9.84 ± 1.30 \[12.5^{149}, 15.2^{150}\] | 0.316 |      |
|       | 333     | 13.6 ± 0.90 \[9.5^{156}\] | 0.437 |      |
|       | 353     | 31.1 ± 3.50                      | 1.000              |      |      |

Figure 8. Form factors for the simulated lipid bilayers (lines) compared to available experimental data (points) at various temperatures. In the insets, neutron scattering form factors are shown. Experimental \( F(0) \) values are shown as black squares.

Table 10. Lateral Diffusion Coefficients for DLPC, DMPC, and DPPC in a Range of Temperatures Compared to Experimental Findings

298 K.
head-to-head distance also agrees well with comparison between simulations and experiment. The significant decrease in number of experimental findings.159 The significant decrease in number of gauche defect per chain (Table 5) illustrates that the bilayer has entered the gel phase and that the conformations of the chains is of importance in controlling the bilayers fluidity. These findings strengthen the present FFs ability to model lipid bilayers under a range of temperatures and shows that the correct phase behavior can be expected. There is a growing interest in simulating membranes in other phases than the liquid crystalline one and the data presented here makes this FF a good candidate for these simulations.

## CONCLUSION

It is critical for computer simulations of lipid bilayers to have sets of parameters obtained in a clearly defined and physically grounded procedure, in order to produce reliable and meaningful results. Here, the development of new AA lipid FF has resulted in a consistent FF where the hydrophilic and hydrophobic forces balance each other, which is crucial. This balance is what drives the formation of the different structures that can be formed by lipids. As model compounds for the aliphatic chains of saturated lipids, a series of alkanes was chosen and many experimentally accessible properties for these bulk liquids were reproduced in the simulations. Furthermore, a large number of important properties for lipid bilayer are in very good agreement: area and volume per lipid, membrane thicknesses, isothermal area compressibility, and trans/gauche ratio. But, most importantly, the key properties, NMR order parameters and scattering form factors, are excellently reproduced which is a clear proof of the abilities of the current FF. Simulations performed with DPPC in the gel phase show that the FF gives the correct phase behavior as well. To our knowledge, this is the first reported set of parameters that reproduces all these properties consistently under a range of temperatures. The ability of the FF to reproduce a lipid bilayer in a liquid-crystalline phase in a tensionless ensemble and describe the physics of lipid in a membrane accurately from more than just the point of view of one or two properties makes it an excellent choice for more advanced simulations, including several lipid types and important membrane-inserted molecules. Most importantly, \( I_S^{(2)} \) and \( F(q) \), which are properties that can be determined without any empirical input from an experimental point view, are in very good agreement with available experimental data.

As protein simulations in membranes become more popular and since the FF presented here is compatible with the AMBER, FFs for biomolecule simulations can easily be extended to include transmembrane proteins. Ongoing work in our group aims to expand the FF to unsaturated and polyunsaturated lipids, lipids with different head groups, and cholesterol. A FF that has been developed in a consistent manner for a large number of lipids will be important in future modeling of biological membranes.

## ASSOCIATED CONTENT

### Supporting Information

Further details on torsional fittings, how the membrane simulations were analyzed, and data on partial atomic charges for hexadecane before averaging. A zip archive contains topology files and FF parameters for DLPC, DMPC, and DPPC together with coordinates for equilibrated lipid bilayers ready for use with GROMACS (also available on http://web.me.com/jambeck/lipids).

This material is available free of charge via the Internet at http://pubs.acs.org.

## AUTHOR INFORMATION

### Corresponding Author

*E-mail: joakim.jambeck@mmk.su.se, jambeck@me.com (J.P.M.J.); alexander.lyubartsev@mmk.su.se (A.P.L.).

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This research has been supported by the Swedish Research Council, grant No. 621-2010-5005. The simulations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC-003-11-3), at PDC Center for High Performance Computing, High Performance Computing Center North (HPC2N), and National Supercomputer Centre (NSC). We thank A. Rabinovich for fruitful discussions, J. F. Nagle and N. Kučerka for providing experimental data, and L. Malinowsky and J. Vincent at PDC for granting access to machines not open for production yet. The authors also thank the reviewers for contributing with constructive criticism.

## REFERENCES

(1) Tieleman, D. P.; Marrink, S. J.; Berendsen, H. J. C. Biochim. Biophys. Acta 1997, 1331, 235–270.
(2) Cantor, R. S. Biochemistry 1997, 36, 2339–2344.
(3) Lee, A. G. Biochim. Biophys. Acta 2004, 1666, 62–87.
(4) Gullingsryd, J.; Schulten, K. Biochim. Biophys. Acta 2004, 1666, 205–226.
(5) Jensen, M.; Mourtis, O. G. Biochim. Biophys. Acta 2004, 1666, 2344.
(6) Botelho, A. V.; Huber, T.; Sakmar, T. P.; Brown, M. F. Biophys. J. 2006, 91, 4464–4477.
(7) Grossfield, A. V.; Feller, S. E.; Pittman, M. C. Proc. Natl. Acad. Sci. U.S.A. 2006, 103, 4888–4893.
(8) Nagle, J. F.; Tristram-Nagle, S. Biochim. Biophys. Acta 2000, 1469, 159–195.
(9) Vigh, L.; Escrigá, P.; Sonnleitner, A.; Sonnleitner, M.; Pietto, S.; Maresca, B.; Horváth, I.; Harwood, J. L. Prog. Lipid Res. 2005, 44, 303–344.
(10) Nagle, J. F.; Tristram-Nagle, S. Curr. Opin. Struct. Biol. 2000, 10, 474–480.
(11) Marrink, S. J.; Berendsen, H. J. C. J. Phys. Chem. 1994, 98, 4155–4168.
(158) Schubert, T.; Schneck, E.; Tanaka, M. J. Chem. Phys. 2011, 135, 055105.
(159) Tristam-Nagle, S.; Zhang, R.; Suter, R. M.; Worthington, C. R.; Sun, W. J.; Nagle, J. F. Biophys. J. 1993, 64, 1097–1109.
(160) Sun, W.-J.; Tristam-Nagle, S.; Suter, R. M.; Nagle, J. F. Biophys. J. 1996, 71, 885–891.