Introduction

Molecular dynamics (MD) simulations are increasingly being applied to explore the liquid phase behavior of lipid-based formulations and the gastrointestinal tract lumen. In order for the results from these simulations to be of value, the manner in which molecules interact with both aqueous and oil phases present needs to be as correct as possible. An existing molecular dynamics force field, GROMOS 53a6, was demonstrated to poorly reproduce the partitioning of straight-chain alcohol and short-chain polyethylene glycol (PEG) molecules between octanol and water phase (logP), with the molecules too hydrophobic. Force field parameters for Lennard-Jones interactions between CH2 and CH3 with water oxygen were adjusted to reproduce the experimental octanol logP, with all other Lennard-Jones and force field parameters left untouched. This parameter set, called 53a6DBW, was subsequently used to recalculate straight-chain alcohol and short-chain PEG molecules, with significant improvement in the values obtained. Simulations of a nonionic surfactant in water, octaethylene glycol monocaprylate, were also performed to observe the aggregation behavior. 53a6DBW demonstrated significant improvements in water interactions with the PEG chains, well hydrating the PEG groups, and allowing the formation of micelles. Further improvements and evaluation of the improved parameter set are ongoing.
electrostatic interactions using Coulomb's law. Of interest, of the many different force fields that are available, are. GROMOS 53a6 force field\textsuperscript{12} (reparameterized the polar functional groups from GROMOS 45a3\textsuperscript{13}), which was parameterized based on the enthalpies of hydration and enthalpies of solvation in cyclohexane for a range of compounds, and GROMOS 53a6\textsubscript{OXY,14} which was developed to improve 53a6 for oxygen-containing functional groups based on thermodynamics of pure liquid and solvation free energy of amino acid analogues in cyclohexane and water. Both of these force fields have been developed to be appropriate to use for the molecular simulation of liquid and biomolecular systems.

Using the 53a6 force field as a starting point, the relative interaction strength between the aqueous and oil phases of a series of molecules was checked using the octanol-water partition coefficient (logP). LogP was chosen as the parameter to check simulation accuracy as these values are available for an extensive range of molecules, including lipid formulation excipients and drug molecules. This is not a method to be used for the prediction of logP values for molecules, there are more robust existing techniques available.\textsuperscript{15} Using octanol as model molecule to perform the logP fitting, a limited set of force field parameters were adjusted to better approximate the experimental logP, and then the logP of a series of short-chain polyethylene glycol (PEG) molecules were calculated. The adjusted force field parameters were selected to use existing force fields and modify them as little as possible to better reproduce the property of interest. The literature logP values for straight-chain alkane, alcohol, and short-chain PEG molecules used to fit or check the force field parameters can be found in Supplementary Information Table S1.

After adjusted force field parameters were obtained to better reproduce literature logP values, simulations were performed of a nonionic surfactant in water, octaethylene glycol monocuprylate, with structure shown in Figure 1. This molecule was chosen as it is representative of PEG containing nonionic surfactants that are used in lipid-based formulations.

**Methods**

**Force Field Parameters**

Force fields studied include 53a6,\textsuperscript{12} 53a6\textsubscript{OXY,14} and 53a6\textsubscript{DBW}\textsuperscript{(this study)}, with the Supplementary Information Table S2 containing the force field parameters were selected to use existing force fields and modify them as little as possible to better reproduce the property of interest. The literature logP values for straight-chain alkane, alcohol, and short-chain PEG molecules used to fit or check the force field parameters can be found in Supplementary Information Table S1.

After adjusted force field parameters were obtained to better reproduce literature logP values, simulations were performed of a nonionic surfactant in water, octaethylene glycol monocuprylate, with structure shown in Figure 1. This molecule was chosen as it is representative of PEG containing nonionic surfactants that are used in lipid-based formulations.

**Methods**

**Gibbs Energy Simulations**

Simulations were performed using the GROMACS 5.1.1 software package\textsuperscript{16,19} on super computer facilities provided by Multi-modal Australian ScienceS Imaging and Visualisation Environment. Simulations were performed using the following force fields; GROMOS 53a6,\textsuperscript{12} GROMOS 53a6\textsubscript{OXY,14} and GROMOS 53a6\textsubscript{DBW} (this study), using

![Figure 1. Structure of octaethylene glycol monocuprylate, also known as PEG8 monocuprylate.](Image 49x625 to 287x716)

![Figure 2. Gibbs energy change thermodynamic cycle for the transfer of a solute from octanol to water.](Image 318x628 to 556x716)

**Table 1**

<p>| Lennard-Jones Interaction Parameters for CH2 and CH3 Atom Types Interacting With the Water Oxygen Atom Type |
|-------------------------------------------------|-------------------------------------------------|-------------------------------------------------|</p>
<table>
<thead>
<tr>
<th>Lennard-Jones Parameters for OW to United Carbon Atom Types</th>
<th>Lennard-Jones Parameters for OW to United Carbon Atom Types</th>
<th>Lennard-Jones Parameters for OW to United Carbon Atom Types</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Force Field</strong></td>
<td><strong>GROMOS 53a6</strong>\textsuperscript{9,10}</td>
<td><strong>GROMOS 53a6</strong>\textsuperscript{9,10}</td>
</tr>
<tr>
<td><strong>Atom Type</strong></td>
<td><strong>C6</strong> (kJ mol\textsuperscript{-1} nm\textsuperscript{6})</td>
<td><strong>C12</strong> (kJ mol\textsuperscript{-1} nm\textsuperscript{12})</td>
</tr>
<tr>
<td>CH2</td>
<td>0.0044212472</td>
<td>9.458844 \times 10\textsuperscript{-6}</td>
</tr>
<tr>
<td>CH3</td>
<td>0.005016238</td>
<td>8.377926 \times 10\textsuperscript{-6}</td>
</tr>
</tbody>
</table>
SPC water, and a 2 fs time step. These force fields are subsequently referred to by the force field numbers that is 53a6. Geometry of water molecules was constrained using SETTLE, with the remaining bond constrained with LINCS. Periodic boundary conditions were used on the cubic simulation cell. For the Gibbs energy calculations, electrostatic interactions were modeled using reaction field ($\varepsilon_r = 62$) and a cutoff distance of 1.4 nm, Lennard-Jones nonbonding interactions had a 1.4 nm cutoff. Soft-core potentials were used for Lennard-Jones and electrostatic interactions were used to avoid singularities when particles transition to and from dummy (noninteracting) atoms during the Gibbs energy calculations. The stochastic dynamics algorithm was used with constant volume to perform the production runs ($T_{\text{ref}} = 298$ K, $\gamma = 0.5$ ps). Molecules were inserted into prebuilt solvent boxes, consisting of either 4142 water molecules in a 5.00 nm cube ($d_{\text{initial}} = 0.991$ g cm$^{-3}$) or 400 octanol plus 134 water molecules in a 4.80 nm cube ($4.4\% w/w$, 0.251 mole fraction, $d_{\text{initial}} = 0.818$ g cm$^{-3}$) in the ratio used by DeBolt and Kollman. Wet octanol was used as water saturated octanol contains 0.255 mol fraction (4.5% w/w) water at 298K. The following protocol was used to take the initial simulation cell through to the completion of the production runs:

1. Steepest decent energy minimization ($\Delta t = 2$ fs, 500 steps), was used to remove bad van der Waals contacts between atoms and allow the forces between atoms to reach a reasonable value.
2. Velocity rescaling temperature coupling ($\tau_T = 0.5$ ps) applied after initial velocities generated to give a starting temperature of 298K, with a 2 fs time step for 100 ps.
3. Berendsen isotropic pressure coupling ($\tau_P = 2.0$ ps with reference pressure of 1 bar and $4.5 \times 10^{-5}$ compressibility) was turned off for 100 ps.
4. Stochastic dynamics was then used, with temperature coupling turned off; isotropic pressure coupling was then changed to Parrinello-Rahman ($\tau_P = 2$ ps) and run for 100 ps.
5. Pressure coupling was then turned off and run for a further 100 ps, and
6. Production runs for each $\lambda$ value/state were performed for a total of 2,500,000 steps, 5 ns simulation time. Soft-core potentials were used for Lennard-Jones interactions using $\alpha = 0.5$ (soft-core parameter), $p = 1.0$ (soft-core power), and $s = 0.3$ (soft-core limiting distance).

logP Calculations

The octanol-water partition coefficient, logP, can be calculated from simulations using the Gibbs energy change of solvation in octanol and the Gibbs energy change of solvation in water, see Equation 3. Gibbs energy is a function of state; therefore, it can be calculated using a thermodynamic cycle between the 2 states of interest, which may include nonphysical transformations. The thermodynamic cycle that can be used with simulations to calculate the Gibbs energy of solvation is shown in Figure 2. Therefore, the Gibbs energy of solvation from octanol to water can be calculated via the Gibbs energy of solvation in octanol and Gibbs energy

![Figure 3](image-url)  
**Figure 3.** Experimental and calculated (using 53a6) logP for the (a) straight-chain C$_{2n}$H$_{2n+2}$ series of alkanes and (b) straight-chain C$_{2n+1}$OH series of alcohols.
of solvation in water, see Equation 2. Gibbs energy changes are calculated from stochastic dynamics simulations by performing separate simulations between an initial, fully interacting state to an uncoupled, dummy final state, as shown in Figure 2. Electrostatic interactions are uncoupled first and separately to Lennard-Jones interactions, with the coupling parameter, \( \lambda \), indicating the degree of transition between the 2 states. Simulations were run with \( \lambda \) values of 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1.0 for both uncoupling of the electrostatic and Lennard-Jones interactions, with a total of 42 simulations performed for each molecule.

Once the simulations have been completed for each \( \lambda \) state, the Gibbs energy change between the 2 states can be calculated using either a thermodynamic integration or Bennett’s acceptance ratio method. The Bennett’s acceptance ratio method was selected, due to being more robust and accurate for calculating Gibbs energy changes, and calculated using the `gmx bar` script included with the GROMACS package.

\[
\Delta G_{\text{octanol to water (solute)}} = \Delta G_{\text{octanol}} + \Delta G_{\text{octanol to water (dummy)}} - \Delta G_{\text{water}} \\
= \Delta G_{\text{octanol}} - \Delta G_{\text{water}} = -2.303RT \log P
\]

(2)

\[
\log P = \frac{\Delta G_{\text{octanol to water (solute)}} - \Delta G_{\text{water}}}{2.303RT}
\]

(3)

where \( \log P \) = octanol-water partition coefficient; \( \Delta G_{\text{oct}} \) = Gibbs energy change from a vacuum to octanol (kJ mol\(^{-1}\)); \( \Delta G_{\text{water}} \) = Gibbs energy change from a vacuum to water (kJ mol\(^{-1}\)); \( R \) = gas constant (8.314 x 10\(^{-3}\) kJ mol\(^{-1}\) K\(^{-1}\)); \( T \) = temperature (K).

**Molecular Dynamics Simulations**

MD simulations were performed using the same GROMACS 5.1.1 software, super computer facilities (Multi-modal Australian...
Science Imaging and Visualisation Environment), and the 3 force fields; GROMOS 53a6,12 GROMOS 53a6OXY,14 and GROMOS 53a6DBW (this study). SPC water20 was the solvent, with a large 5 fs time step attained by increase in mass of polar hydrogens to 4 Da (with a corresponding decrease in attached heavy atom mass to conserve mass).38 Water molecule geometry was constrained using SETTLE,21 and remaining bonds were constrained by LINCS.22 Periodic boundary conditions were used on the cubic simulation cell. Electrostatic interactions were treated with a short range cutoff distance of 1.4 nm, and the particle-mesh Ewald for long range \((\text{Fourier grid spacing} = 0.12, \text{interpolation order} = 4, \text{relative strength of Ewald-shifted direct potential} = 1 \times 10^{-5})\).19 Lennard-Jones nonbonding interactions had a 1.4 nm cutoff14 with long-range dispersion corrections for energy and pressure applied. The isothermal-isobaric ensemble (NPT) was used, with the velocity rescaling temperature coupling \((T_{\text{ref}} = 310 \text{ K}, \tau_T = 0.1 \text{ ps})\) and Parrinello-Rahman pressure coupling\(^{30,31}\) (isotropic, \(P_{\text{ref}} = 1 \text{ atm}, \tau_P = 2 \text{ ps}\)). To avoid bias to the structure of the aggregates formed within the simulations, initial molecule geometries were generated by random placement of 300 octaethylene glycol monocaprylate with 103,731 water molecules in a 15 ns cubic simulations box \((d = 0.993 \text{ g cm}^{-3})\), giving a total of 321,693 atoms. Once filled, the simulation cell was processed using the following protocol:

1. Steepest decent energy minimization\(^{27}\) was used to remove bad van der Waals contacts between atoms and allow the forces between atoms to reach a reasonable value.
2. Berendsen temperature coupling\(^{29}\) \((\tau_T = 0.01 \text{ ps with reference temperature of } 310 \text{ K})\) was turned on, no pressure coupling, and run for 10 ps with a 2 fs time step.
3. Temperature coupling was then switched to velocity rescaling\(^{28}\) \((\tau_T = 0.1 \text{ ps})\) and isotropic Berendsen pressure coupling\(^{29}\) \((\tau_P = 2.0 \text{ ps with reference pressure of } 1 \text{ bar and } 4.5 \times 10^{-5} \text{ compressibility})\), and run for 20 ps with a 2 fs time step.
4. Isotropic pressure coupling was then changed to Parrinello-Rahman\(^{30,31}\) \((\tau_P = 2 \text{ ps})\) and run for 40 ps with a 2 fs time step.
5. Time step was increased to 4 fs and run for 40 ps.
6. Finally, the time step was increased to 5 fs and run for 50 ps, and
7. Production runs were performed for a total of 40,000,000 steps, 100 ns.

**Results and Discussion**

**logP With 53a6**

The logP for the series of straight-chain alkanes \((\text{C}_n\text{H}_{2n+2})\) from \(\text{C}_1\) to \(\text{C}_{12}\) (excluding \(\text{C}_{11}\), no logP value available) was calculated using the 53a6 force field and are presented in Figure 3a). The force
field manages to fit the logP relatively well until the chain length reaches heptane, with at least a 1 log unit difference between the experimental and calculated values starting from this point on. This increases significantly more to a difference of 2.2 units for dodecane. This overestimation of the logP value indicates that the existing parameters for long-chain alkanes allow insufficient water interactions. An equivalent series of straight-chain alcohols (C₆H₂₃₋₁O-H) from C₁ to C₁₂ (excluding C₁₁, no logP value available) were also calculated, with the results shown in Fig. 3b. The difference in the logP between the calculated and experimental values is stark, with the logP being from 1.0 to 1.5 log units too high, with larger deviations found as the chain length increases. The simulated alcohols are significantly more hydrophobic than determined experimentally.

One method to change the interactions of a functional group with water is to alter the functional group polarity, by adjusting the partial charges (used in the parameterisation of the 53a6OXY force field). An attempt to improve the simulated logP of octanol using this method was attempted by Birru, using linear scaling of the 3 atom types making up the alcohol functional group (CH₂-O-H), maintaining a net neutral charge for the functional group. However, this was insufficient to get the logP of octanol close to the experimental value, see Figure 4. Additional evidence of the inability of partial charge adjustment to successfully match the calculated logP values is shown below for the logP of PEG chains using 52a6OXY (PEG logP section, Fig. 4). Adjusting the partial charges of the PEG functional group by Birru was also unable to reproduce the experimental values. It is proposed that the failure of these force fields to sufficient model interactions with water may be due to the force fields not being polarizable. To overcome this limitation and improve the interactions between hydrocarbon groups, it was decided to adjust the Lennard-Jones interactions between CH₂ and CH₃ atom types and water oxygen, while leaving all other Lennard-Jones interactions and force field parameters untouched. Adjusting these specific interactions minimizes changes to the force field parameters, which are a complex, interrelated set of values. The increase in the deviation between simulated and experimental logP with increasing chain length, see Figure 3, indicates that it is the chain interactions, not the alcohol functional group, that are the cause of the discrepancy. Changing the CH₂ and CH₃ Lennard-Jones interactions with water oxygen will directly improve the affinity of these atom types for water and improve the partitioning of the molecules between water and oil.

**Parameterization Using Octanol logP**

Parameterization of the CH₂ and CH₃ atom type interactions with water oxygen type was achieved by adjusting the C₆ and C₁₂ Lennard-Jones parameters for these atom type pairs, with the values for 53a6 in Supplementary Information Table S4. All other Lennard-Jones interaction parameters and force field parameters were left untouched. An initial scaling of the C₆ and C₁₂ parameters, independent of each other, was performed by calculating the logP of octanol, with the results presented in Figure 5a. Increasing C₆ (line from A to B) and decreasing C₁₂ (line from D to C) both decreased the calculated octanol logP value, a direct result of decreasing the minimum potential value and minimum radius of the Lennard-Jones potential for these atom type pairs. A 2-dimensional scan of scaling C₆ from 0 to +10% and C₁₂ from 0 to −10% was performed to identify the best combination of Lennard-Jones parameter values to reproduce the experimental octanol logP value, with the results presented in the contour diagram in Figure 5b. Subsequently, a finer scale scan was performed around the values that were close to reproducing the experimental octanol logP of 3.0771 (see Figure 5c). From this contour diagram, the shortest distance line from the 53a6 parameters to the best approximation to the experimental octanol logP was taken, corresponding to a 5% increase in the C₆ and 5% decrease in the C₁₂ parameters. The value of these parameters is presented in Table 1.

**Alcohol logP with 53a6DBW**

Taking the previously obtained values (5% increase in C₆ and 5% decrease in C₁₂ parameters for both CH₂ and CH₃ atom type Lennard-Jones interactions with water oxygen atom type), the logP of the straight-chain alkanes and alcohols (C₆H₂₃₋₁O-H) series from C₁ to C₁₂ (excluding C₁₁) were recalculated. The results of these
calculations are presented in Figure 6, showing the significant improvement in the calculated logP compared to the experimental values.

Once the chain length get past C9, nonanol, a discrepancy between the calculated and experimental values for the alcohols appears. A possible reason for this is the C-C-C-C proper dihedral rotation barrier is too high, restricting the longer molecules from being able to wrap back on themselves. This would stop a decrease in the surface area exposed to water allowed by curling up, thereby making the longer chains more hydrophobic. This possibility was investigated by scaling down the C-C-C-C proper dihedrals for dodecanol and recalculating the logP. Scaling the potential energy barrier down to 10% of the original value decreased the calculated logP of dodecanol from 6.20 to 6.03, insufficient to even approach the experimental value of 5.13.40 The reason for this observed discrepancy is yet to be identified and study is ongoing.

**PEG logP**

The logP for the series of short-chain PEG molecules from mono- to pentaethylene glycol were then calculated for 53a6, 53a6OXY, and 53a6DBW, see Figure 7. The functional group partial charge increase that occurred with the reparameterization of 53a6 to 53a6OXY is shown here to improve the interactions with water, causing a significant decrease in the calculated logP values. However, again, as shown previously in Figure 4, this change in partial charges is insufficient to achieve the correct logP values. Only when the CH2 and CH3 atom type Lennard-Jones interactions with water oxygen atom types is adjusted, as done with 53a6DBW, the more substantial water interactions required are obtained.

**Micellar Structure**

To provide a larger scale validation of the manner in which water interacts with molecules containing the PEG functional group, 3 simulations (using all 3 force fields, 53a6, 53a6OXY and 53a6DBW, with equilibrated system densities of 0.9781, 0.9788 and 0.9817 ± 0.0008 g cm\(^{-3}\), respectively) of octaethylene glycol monocaprylate in water were performed to observe aggregation behavior. Octaethylene glycol monocaprylate is representative of the types of excipients found within lipid formulations. The final aggregates formed at the completion of these simulations are presented in Figure 8. 53a6 shows strong aggregation behavior, Figure 8a, forming 2 large aggregates with very poor hydration of the PEG chains, which are lying flat on the surface. Formation of large aggregates in this manner within MD simulations typically represents phase separation.\(^{1-3,5}\) This aggregating behavior is consistent with the high logP found for PEG chains (Fig. 7). In the case of 53a6OXY, 2 large aggregates are formed as well, Figure 8b. However, an improvement in the PEG chain hydration is now apparent, with some PEG chains protruding out into the bulk water phase. Finally, when 53a6DBW is used, Figure 8c, small micelles are formed, and it is no longer phase separated. The formation of the micelles is caused by the increase in degree of hydration of the PEG chains, allowing the surfactant to form a more curved oil/water interface. This trend of improving degree of hydration of the PEG chains across the force field series 53a6, 53a6OXY to 53a6DBW is...
quantified via the radial distribution functions of selected PEG oxygen atoms to water oxygen, see Figure 9.

Conclusions

Existing GROMOS force fields 53a6 and 53a6OXY have been demonstrated to poorly model the required relative aqueous and oil phase interactions to correctly reproduce correct logP values, with the interactions with water too unfavorable, as indicated by the simulated logP being too high. Therefore, these force fields would be poor candidates for performing MD simulation of the gastrointestinal tract, lipid formulations, and the complex liquid phases that form when lipid formulations are dispersed and digested in the gastrointestinal tract. Using the logP of octanol as a fitting parameter, the Lennard-Jones interactions between CH2 and CH3 atom types with water oxygen atoms were scaled until a logP value was obtained close to the experimental, with all other Lennard-Jones and force field parameters left untouched. These parameters were called the 53a6DBW force field, which improves the calculation of logP values for straight-chain alcohols and short-chain PEG molecules. The improvement in the water interactions was also illustrated by simulation of a lipid formulation excipient, octaethylene glycol monocaprylate, in water to observe its aggregation behavior. Although 53a6OXY is an improvement over 53a6, it does not allow sufficient hydration of the PEG chains. Meanwhile, 53a6DBW is able to more accurately model micellar formation by a nonionic surfactant and further evaluation is forthcoming of this force field (Guruge et al., unpublished results).41

Acknowledgments

Financial support from Lonza Pharma Sciences (formerly Capsugel) is gratefully acknowledged. This work was supported by CPU time from Melbourne Bioinformatics (formerly Victorian Life Sciences Computation Initiative, VLSCI) and the Multi-modal Australian ScienceS Imaging and Visualisation Environment (MASSIVE).

Supporting Information

Supporting Information is available; Tables S1 (literature values for alkanes, alcohols, and ethylene glycols), S2 (force field atomic partial charges), and S3 (dihedral parameters for PEG groups), and Figure S1 (Lennard-Jones energy potentials).

References