Predicting Partition Coefficients of Neutral and Charged Solutes in The Mixed SLES-fatty acids Micellar System

Mattia Turchi, †, ‡, *Abhishek A. Kognole,* & *Anmol Kumar,* & *Qiong Cai,* ‡ *Guoping Lian,* †, ‡, * and *Alexander D. MacKerell Jr.*&, *

† Unilever Research Colworth, Colworth Park, Sharnbrook, Bedfordshire MK44 1LQ, UK,

‡Department of Chemical and Process Engineering, University of Surrey, Guildford GU27XH, UK

&University of Maryland Computer-Aided Drug Design Center, Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, MD 21201, USA.

*Corresponding authors: Guoping.Lian@unilever.com and alex@outerbanks.umaryland.edu

Abstract

Sodium laureth sulfate (SLES) and fatty acids are common ingredients in many cosmetic products. Understanding how neutral and charged fatty acid compounds partition between micellar and water phases is crucial to achieve the optimal design of the product formulation. In this paper, we first study the formation of mixed SLES and fatty acids micelles using molecular dynamics simulations. Micelle/water partition coefficients of neutral and charged fatty acids are then calculated using COSMOmic as well as a molecular dynamics (MD) approach based on Potential of Mean Force (PMF) calculations performed using Umbrella Sampling (US). The combined US/PMF approach was performed with both the additive, non-polarizable CHARMM General Force Field (CGenFF) and the classical Drude polarizable force field. The partition coefficients for the neutral solutes are shown to be accurately calculated with the COSMOmic and additive CGenFF US/PMF approaches while only the US/PMF approach with the Drude polarizable force field accurately calculated the

experimental partition coefficient of the charged solute. These results indicate the utility of the Drude polarizable force field as a tool for the rational development of mixed micelles.

1 Introduction

Pharmaceutical and cosmetic products often have complex formulations involving cationic or anionic surfactants which form complex microstructures such as mixed micelles. Understanding how active solutes partition in such microstructures helps in achieving their optimal delivery and maximum efficacy of functional benefits such as health, nutrition, hygiene and wellbeing. The solute partition coefficient in multiphase materials is a thermodynamic property that is related to the free energy change associated with the transfer of a solute molecule from one phase to another. It is defined as the ratio of the solute concentrations between the two phases at equilibrium. Experimentally, partition coefficients can be measured in different ways. Direct measurements are usually performed by using High-Pressure Liquid Chromatography[1,2] (HPLC), Micellar Electrokinetic Chromatography[3,4] (MEKC) and microemulsion electrokinetic chromatography[5] (MEEKC). While solute partitioning in complex formulations can be measured by experimental methods, these are often quite expensive and time consuming. For this reason, development of *in-silico* methodologies for accurate prediction of partition coefficients would make product development more cost effective.

Early *in-silico* models have been reported for predicting solutes partitioning in model biphasic systems[6]. Group contributions (GC) and Quantitative structure-activity relationship (QSAR) methods have been extensively used and can often predict partition coefficients with good accuracy[6]. These methods require an extensive amount of experimental data for the estimation of model parameters and therefore their applicability is limited to molecules groups for which

a large number of experimental data are available. An alternative method for the prediction of thermodynamic properties of mixed solvents is the "**co**nductor-like screening model for real solvents", the so-called COSMO-RS[7,8] theory that is implemented in the COSMOtherm software[6]. The COSMO-RS theory predicts partition coefficients using quantum mechanical (QM) calculations as a basis and its applicability is much wider than that of GC and QSAR methods. COSMOmic[9] method is an extension of the COSMO-RS theory for inhomogeneous systems such as micelles and other molecular assemblies. In the COSMOmic approach, the molecular structure of assemblies is explicitly considered at the atomistic scale where the resolution of the molecular assembly is usually achieved by performing molecular dynamics (MD) simulations. This is also referred as the MD/COSMOmic approach. Several publications showed that the MD/COSMOmic approach accurately predicts partition coefficients for neutral solutes not only in homogeneous fluids such as octanol-water[10] but also in structured fluids such as micelles containing anionic, cationic, zwitterionic surfactants mixtures[11–16] and microemulsions[17]. However it fails in accurately predicting the partition coefficients of charged solutes in micellar systems[15].

Alternatively, a molecular mechanics[18] (MM) modelling approach, that employs MD simulations and Umbrella Sampling[19,20] (US), from which potentials of mean force (PMF) are obtained, can be employed for the *in-silico* prediction of partition coefficients. This approach is hereafter referred to as the US/PMF approach. The US/PMF approach was used by Yordanova *et al.*[15] for predicting partition coefficients of neutral and charged solutes in micellar systems. The US/PMF approach showed good accuracy in predicting partition coefficients of neutral solutes but severe inaccuracy in predicting partition coefficients for charged solutes. Yordanova *et al.*[15] suggested that the inaccuracy of the US/PMF approach in predicting partition coefficients for charged solutes was due to the employment of a non-polarizable force field that cannot accurately model the electrostatic interactions.

In this paper we apply both the combined MD/COSMOmic and the US/PMF approaches to predict the partition coefficient of neutral capric acid and charged capric acid anion (caprate) in the mixed micelle of sodium laureth sulfate and capric acid (mixed SLES/CA micelle). Calculations were also performed for the partition coefficient of neutral palmitic acid in the mixed micelle of SLES and palmitic acid (mixed SLES/PA micelle). The combination of SLES and fatty acids (capric and palmitic acid) is ubiquitous in hair shampoos and liquid soap formulations[21,22]. The interactions between fatty acids and SLES anionic surfactants have been extensively studied by Tzocheva *et al.*[23] by experimentally measuring the free energies of transfer of fatty acid molecules from water to the mixed SLES/fatty acid (SLES/FA) micelles. In the current work, MD simulations are initially performed to predict the self-assembly of mixed SLES/capric acid (SLES/CA) and SLES/palmitic acid (SLES/PA), at the experimental concentration of fatty acids in SLES[23], using the CGenFF[24] (Charmm General Force Field) non-polarizable force field. COSMOmic is subsequently used for predicting the partition coefficients based on the structures from the MD simulations. The comparison of the predicted partition coefficients with the experimental data[23] shows that the combined MD/COSMOmic approach is accurate for predicting the micelle/water partition coefficients of the mixed SLES/PA micelles for the neutral solutes (capric and palmitic acid) whereas it lacks accuracy for the charged solute (caprate). Therefore, the non-polarizable force field (CGenFF) and the CHARMM classical Drude polarizable force field are used for performing the US/PMF calculations to predict solute partitioning in the mixed SLES/FA micellar systems. To the best of our knowledge, this is the first time that a polarizable force field, that can accurately model the effect of the polarizability of anionic surfactant molecules near the water/micelle interface, is employed for predicting the micelle/water partition coefficients of a charged solute. The comparison of the predicted values with the experimental data shows that the use of the

polarizable force field in the US/PMF approach is accurate and robust for predicting the partition coefficient of both neutral and charged solutes in the SLES/FA micellar solution.

2 Experimental dataset

The reported experimental value of critical micelle concentration of the SLES surfactant, in pure water at 25 °C, is 0.003 M[25]. Values for the solubility limits of capric acid and palmitic acid in the respective mixed SLES/CA and SLES/PA micellar solutions were from Tzocheva *et al.*[23] who reported saturation molar fractions of 0.301 and 0.0909 for capric acid and palmitic acid, respectively using light absorbance measurements. The free energies of transfer of capric acid and caprate ion from water to the mixed SLES/CA micelle phase and that of palmitic acid from water to the mixed SLES/PA micelle were also collected from Tzocheva *et al.*[23] These values along with the relative values of the micelle/water partition coefficients are listed in Table 1. Capric acid ion values were derived from monomer concentration in water and in the micelle phase (calculation of $K_{mic,A}$ in supplementary information). In order to derive the free energies of transfer of capric acid, caprate and palmitic acid from water to the respective mixed micellar phase, the authors used a semi-empirical approach[26,27], based on the measured solubilities of fatty acids in water and in the SLES micelle.

Table 1: Micelle/water partition coefficients, $K_{mic,A}$, and free energy of transfer, $\Delta G_{transfer}$, for fatty acids at 25 °C in the mixed SLES/FA micelles.

Compound	$\log K_{mic,A}$(mole/mol)*	ΔG_{transf}(kJ/mol)*
Capric acid	3.37	-19.24
Palmitic acid	6.36	-36.20
Capric acid anion	1.03	-5.88

*Data collected from Tzocheva *et al.*[23]

3 Methods

3.1 CHARMM classical Drude polarizable force field parametrization

The fixed partial atomic charges used in the non-polarizable force field do not take into account the induced polarization arising from the perturbation of the electronic structure of molecules in response to the external electric field[28]. On the contrary, the CHARMM classical Drude polarizable force field explicitly models the effect of polarization by attaching a charged particle (Drude oscillator) to each polarizable atom through a harmonic spring. As a consequence, a finite induced dipole is created, and the atomic dipole varies by changing the spatial relationship between the atomic nucleus and the Drude particles. The CHARMM Drude force field was first proposed by Lamoureux *et al.*[29,30] for water and it was further developed for a range of small molecules and biomolecules[31]. Li *et al.* [32], Chowdhary *et al.* [33] and Harder *et al.* [65] showed how the use of the Drude force field in simulating lipid membranes leads to significantly different profiles for the electrostatic potential compared to the additive CHARMM36 field. The authors reported that the most significant difference between the polarizable and the non-polarizable force fields appears in the lipids/water interface region. In that region the effect of the induced polarization between water and lipids headgroups is a particularly important feature that cannot be captured by an additive force field. Similarly, in the case of the mixed SLES/CA micellar systems, the interaction of the headgroups of the SLES and CA with water leads to significant degree of induced polarization. The effect of the polarizability of anionic surfactant molecules near the water/micelle interface is particularly important when predicting the micelle/water partition coefficients of the charged solutes, whereas it does not affect predictions for neutral solutes, as shown by Yordanova *et al.*[15] Accordingly, the polarizable Drude force field was extended in this study to SLES, capric acid and caprate. Details of the parameter optimization procedure are presented in the parametrization section of supplementary information.

3.2 MD simulations of SLES/FA mixed micelles

For simulations using the non-polarizable force field all MD simulations were performed with GROMACS 5.5.1[34]. Non-polarizable force field parameters for SLES and fatty acids were obtained from the CGenFF[24] program and the TIP3P[35,36] force field was used for water molecules. The Verlet cut-off scheme was employed and both the short-range electrostatic cut-off and the short-range van der Waals (vdW) cut-off were set at 1.2 nm. The vdW interactions were smoothed over 1.0 to 1.2 nm using the forced switch method[37] while Particle Mesh Ewald (PME) method was used for long-range electrostatic interactions[38]. The Nosé-Hoover[39] thermostat with a coupling constant of $\tau_t = 1$ ps was used for maintaining a constant temperature at $T = 298.15$ K and the Parrinello-Rahman[40] barostat with a coupling constant of $\tau_p = 1\ ps$ was used for maintaining the pressure at a constant value of $P = 1$ bar.

All simulations performed with the polarizable force field were carried out using the OpenMM software[41] (http://openmm.org/). The Drude polarizable force field was employed for the SLES and fatty acid molecules along with the SWM4-NDP[42] model for the water molecules. The PME method[38] was used to calculate electrostatic interactions with a real-space cutoff of 1.2 nm. The van der Waals potential was smoothed to zero from 1.0 to 1.2 nm using a potential switch function. Covalent bonds to hydrogen atoms were constrained and the Drude particle to atom nucleus separation were limited to 0.2 Å by using the hard-wall constraint. The thermostat was set to a reference temperature of 298.15 K and maintained with a friction coefficient of 5 ps^{-1}. The Drude oscillator thermostat was set to 1 K with a friction coefficient of 20 ps^{-1}. The pressure was maintained at 1 bar using the Monte-Carlo barostat in OpenMM.

3.2.1 Micelles self-assembly

MD simulations were performed to simulate the self-assembly of the mixed SLES/CA and the SLES/PA micelles under the experimental conditions reported by Tzocheva *et al.* [23], using the CGenFF non-polarizable force field. For the mixed SLES/CA system, 216 SLES molecules were randomly placed in a cubic simulation box of 8 nm × 8 nm × 8 nm, by means of the insert-molecules command in GROMACS. Subsequently, 95 capric acid molecules were added in order to match the experimental value of the saturation molar fraction of 0.301[23]. The system was then solvated with 22,216 water molecules of which 216 water molecules were replaced by 216 sodium ions to achieve electroneutrality. The resulting SLES concentration in water was $\hat{c}_{SLES}$ = 0.47 M. For the mixed SLES/PA system, the procedure for generating the MD simulations system is identical to that for the SLES/CA, except that 20 palmitic acid molecules, rather than 95 of capric acid, were added to match the experimental value for the saturation molar fraction of 0.0909. The internal energy of the system was minimized by means of the steepest descent algorithm. After the energy minimization, a short equilibration simulation of 600 ps was carried out in the isothermal-isochoric ensemble (NVT) with a timestep of 2 fs. The production simulation was run in the isothermal-isobaric ensemble (NPT) for 45 ns with a timestep of 2 fs. Snapshots were saved at intervals of 100 ps. By the end of the MD simulation, several micelles of different sizes had formed. The criterion to identify the micelle to which each SLES and fatty acid molecule belongs followed the method originally proposed by Sammalkorpi[43] for pure SDS micelles and further developed by Koneva *et al.*[44] for mixed micelles. For the SLES/CA system three sets of distances between the selected atoms are computed for all pairs of SLES and capric acid molecules (Figure 1 and Table 2). Similarly, for the SLES/PA system three sets of distances are computed for all pairs of SLES and palmitic acid molecules (Figure 1 and Table 2). SLES and fatty acid molecules are considered to be in the same micelle if at least one of the computed distances in either set 1, set 2 or set 3 is shorter

than $r1_{cutoff}$ = 0.55, or if two distances from two different sets are shorter than $r2_{cutoff}$ = 0.68 or if all the three distances from the three different sets are shorter than $r3_{cutoff}$ = 0.70. Values of the cut-offs were chosen in accordance with the work conducted by Storm *et al.*[45] on similar mixed surfactant systems.

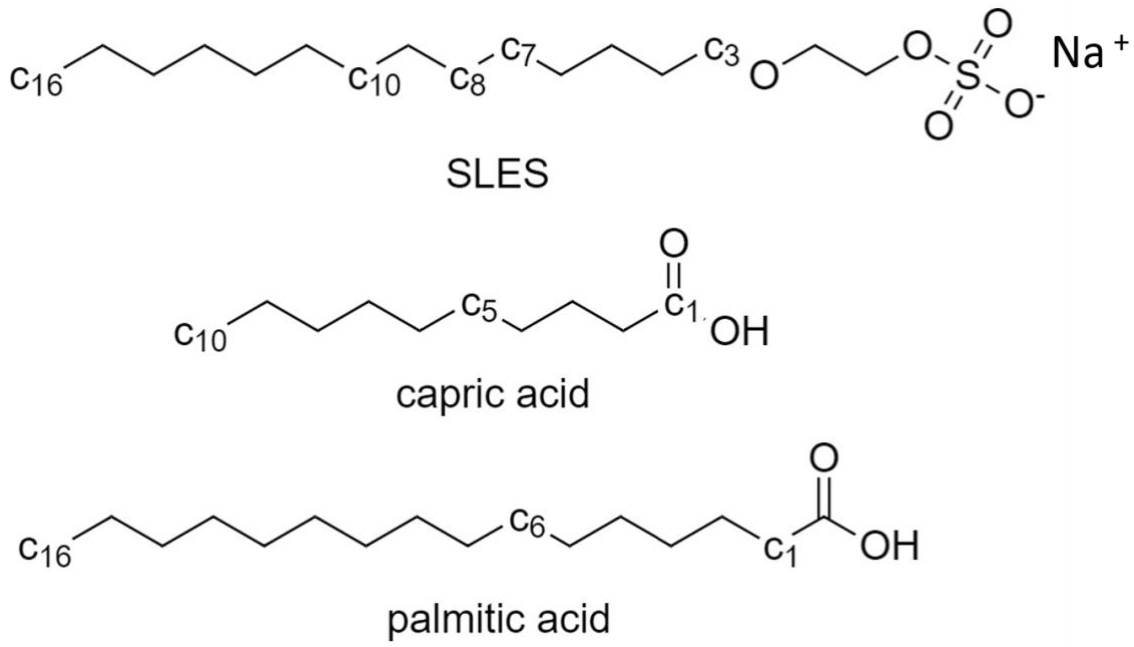


Figure 1: Reference atoms for SLES, capric acid and palmitic acid.

Table 2: Distances between carbon atoms and cutoffs used for the definition of micelles.

Molecules	Distances between atoms			r_1	r_2	r_3
SLES and capric acid	C3_SLES/ C3_SLES or C3_SLES/C1_Capric or C1_Capric/C1_Capric	C7_SLES/ C7_SLES or C7_SLES/C5_Capric or C5_Capric/C5_Capric	C10_SLES/ C10_SLES or C10_SLES/C10_Capric or C10_Capric/C10_Capric	0.55	0.68	0.70

SLES and palmitic acid	C3_SLES/ C3_SLES or C3_SLES/C1_Palmitic or C1_Palmitic/C1_Palmitic	C8_SLES/ C8_SLES or C8_SLES/C6_Palmitic or C6_Palmitic/C6_Palmitic	C16_SLES/ C16_SLES or C16_SLES/C16_Palmitic or C16_Palmitic/C16_Palmitic	0.55	0.68	0.70

3.2.2 Partition coefficient predictions from the US/PMF approach

The largest mixed SLES/CA and SLES/PA micelles were extracted from the final configuration of the MD simulations and each of the micelles was transferred to a water box of 8 nm × 8 nm × 8 nm for setting up the steered molecular dynamic (SMD) simulations (Figure 2). Altogether, 5 SMD simulations were performed to generate the configurations for the three solutes of capric acid, palmitic acid and caprate in the two mixed SLES/CA and SLES/PA micelles. 3 US/PMF simulations were performed using the non-polarizable force field for the capric acid and caprate solutes in the mixed SLES/CA micelle and for the palmitic acid in the mixed SLES/PA micelle. 2 US/PMF simulations were performed using the polarizable force field for the capric acid and caprate solutes in the mixed SLES/CA micelle. In each of the SMD simulations, the solute molecule was placed at a distance of 3.5 nm from the micelle centre of mass (COM). For the non-polarizable force field simulation, the system was relaxed in order to minimize the internal energy with the steepest descent algorithm for 1000 steps and equilibrated in the NVT ensemble for 600 ps. For the polarizable force field simulation, the internal energy was minimized by allowing the Drude particles to move while the positions of all the atoms were kept fixed. Subsequently all atoms and Drude particles were relaxed with an energy minimization of 1000 steps of steepest descent followed 1000 steps of the adopted basis Newton-Raphson algorithm. Minimization was followed by 50 ns of equilibration in the NPT ensemble with a timestep of 1 fs. At the end of the equilibration, the COM of each micelle was constrained to the COM of the respective solute molecule by means of a harmonic potential of 3000 kJ/mol/nm$_2$. PLUMED[46] was used to apply the COM distance restrains in the simulations performed with the polarizable force field. US/PMF calculations were performed

from the starting COM of each solute at 3.5 nm over 36 US windows to the COM of the micelle. The configuration for each US window was obtained through the SMD simulations by pulling each solute molecule towards the respective micelle COM by applying a velocity of 10 nm/ns when the non-polarizable force field was employed (Figure 3). For polarizable systems this velocity was reduced to 2 nm/ns. For each of the 5 US/PMF simulations, the respective US simulations were run for each configuration for 10 ns in the NPT ensemble. The force constant for the umbrella potential was set to 3000 kJ/mol/nm$_2$. Subsequently, calculation of the PMF to account for the US biasing harmonic potential was performed using the weighted histogram analysis method (WHAM)[47] in the WHAM software[48], version 2.0.9 (http://membrane.urmc.rochester.edu/page_id=94) in which the first 5 ns of sampling in each 10 ns window was considered as equilibration and discarded with the PMFs calculated over the final 5 ns. This yielded the unbiased free energy profile (i.e. PMF) for the transfer of each solute from the water to the micellar phase. The convergence of the PMFs was tested by running WHAM for different time portions of the 36 windows (i.e. 0-2.5 ns, 0-5ns, 0-7.5ns, 0-10ns). Plots of PMF showing the achieved convergence for the caprate solute in both the polarizable and the non-polarizable systems are shown in Figure S1 of the supplementary information. As the US from the COM of the micelle to the water phase involves sampling in discretized shells of increasing volumes, the Jacobian correction[49,50] was implemented in the free energy calculations in order to take into account the effect of transforming the Cartesian coordinates into the distance reaction coordinate, using the following equation proposed by Ciccotti *et al.*[51] Error estimates for the PMFs were calculated as the averaged absolute differences between values at 10 ns and 8 ns and between values at 10 ns and 6 ns.

$$\Delta G(r_i) = \Delta G^{WHAM}(r_i) + 2k_B T \ln \frac{r_{i+1}}{r_i} \quad (1)$$

where i is the index that runs over the bins of the discretised shell. $\Delta G^{WHAM}(r_i)$ is the unbiased free energy computed by WHAM and $2k_B T \ln \frac{r_{i+1}}{r_i}$ is the applied Jacobian correction. Subsequently, the partition coefficient of each solute is calculated from the corrected free energy profiles as follows:

$$\text{Log } K_{mic/w} = \frac{\Delta G_{transf}}{-2.303RT} \quad (2)$$

where ΔG_{transf} is the difference between the free energy in the water phase (0) and the minimum free energy value in the PMF profile that corresponds to the most populated state of the solute in the micelle.

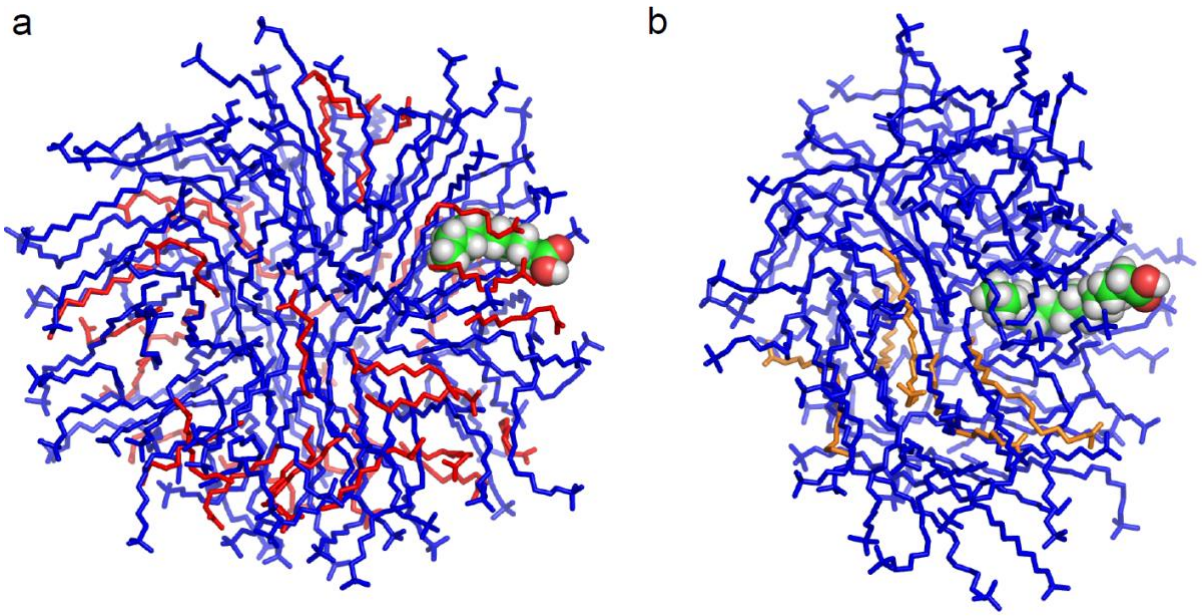


Figure 2: Extracted micelle configurations used for the steered molecular dynamic simulation for capric acid (a), palmitic acid (b) in the respective SLES/FA mixed micelles. SLES molecules in blue, capric acid in red, palmitic acid in orange; water molecules are excluded. Solute molecules are highlighted by representing oxygen atoms in red, carbon atoms in green and hydrogen atoms in white as spheres.

3.2.3 Partition coefficient predictions from the MD/COSMOmic approach

The largest micelle from each of the two simulated SLES/FA micellar systems, i.e. SLES/CA and SLES/PA, with an adjacent shell of water molecules, as shown in Figure 3, was extracted and fed into COSMOmic for predicting the micelle/water partition coefficient of capric acid and caprate in the SLES/CA micelle and that of palmitic acid in the SLES/PA micelle. COSMOmic[9] is an extension of the COSMO-RS theory for inhomogeneous systems in which the chemical potential of a solute within its surrounding solvent is computed from the screening charge density (σ) on the surface of molecules[7,8,52]. Screening charge densities are calculated by quantum mechanics (QM). The σ values for capric acid, palmitic acid, caprate and SLES were computed using density functional theory (DFT) with the Becke-Perdew[53,54] (BP) functional, the triple-zeta valence polarization[55,56] (TZVP) basis set and the resolution of identity[57] (RI) approximation. QM calculations were carried out using Turbomole 7.3[58] package (http://www.turbomole.com). Details about parameters needed for performing DFT QM calculations can be found in the Turbomole user manual (http://www.cosmologic.de/files/downloads/manuals/TURBOMOLE-Users-Manual_70.pdf). COSMOmic 19.0 (http://www.cosmologic.de) is used in this work (http://www.cosmologic.de/files/downloads/manuals/COSMOmic_Manual.pdf), along with the COSMO-RS parameter file BP_TZVP_19.ctd. In COSMOmic, each of the two micelles was discretized into 30 layers along the radius with the last layer exclusively consisting of water molecules. The micelle/water partition coefficients were computed by the default equations implemented in COSMOmic 19.0.

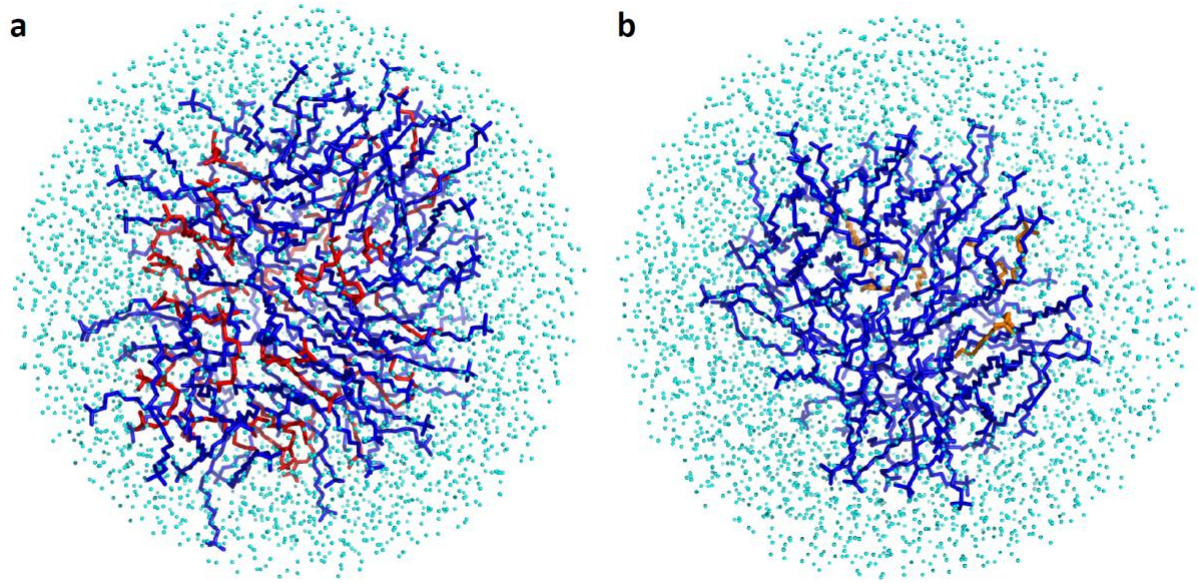


Figure 3: Self assembled mixed micelles of SLES/FA predicted by MD simulation: (a) SLES/CA and (b) SLES/PA surrounded by a water shell, used as input for COSMOmic calculations. Water molecules are colored in cyan, SLES molecules in blue, capric acid in red, and palmitic acid in orange.

4 Results and discussion

4.1 Micelle structure

4.1.1 Self-Assembly of the mixed SLES/FA micelles

The evolutions of the maximum and mean aggregation numbers along with the numbers of micelles over the course of the simulated self-assembly are plotted in Figures 4a and 4b for the SLES/CA and the SLES/PA systems, respectively. The numbers of molecules in the largest micelles reached the maximum values of 120 for SLES/CA and 74 for SLES/PA by 5ns and remained constant afterwards. The number of micelles decreases as coalescence leads to the formation of bigger micelles characterized by a higher value of the average aggregation number. At the end of the simulation, the SLES/CA and the SLES/PA systems reached average aggregation numbers of 63 and 48, respectively. Five micelles were distinguishable for both the SLES/CA and the SLES/PA systems.

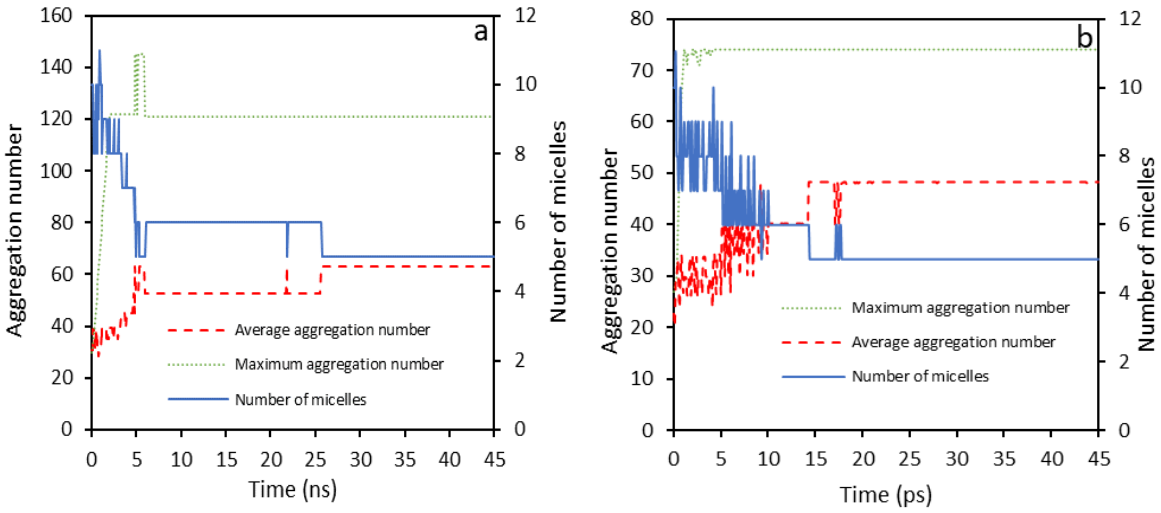


Figure 4: MD simulation of self-assembly of the mixed SLES/CA micelles (a) and the mixed SLES/PA micelles (b), as characterized by the aggregation number (left axis) and number of formed micelles (right axis).

From the MD simulation, the probability distributions of aggregation numbers are obtained for the mixed SLES/CA and the mixed SLES/PA systems. The aggregation numbers that occurred more often are 21, 24, 50, 91 and 121 for SLES/CA and 19, 34, 45, 69 and 74 for SLES/PA, as shown in Figure 5. The aggregation numbers for pure SLES micelles were experimentally measured to be 43 by Aoudia *et al.*[59] and between 67-79 by Anachkov *et al*[60]. Predicted aggregation numbers for both the SLES/CA and the SLES/PA systems are in these ranges. The aggregation numbers for the mixed SLES/PA system are closer to the experimental values of SLES pure micelles compared to the SLES/CA system; this is due to the lower solubility of palmitic acid compared to capric acid in SLES micelles.

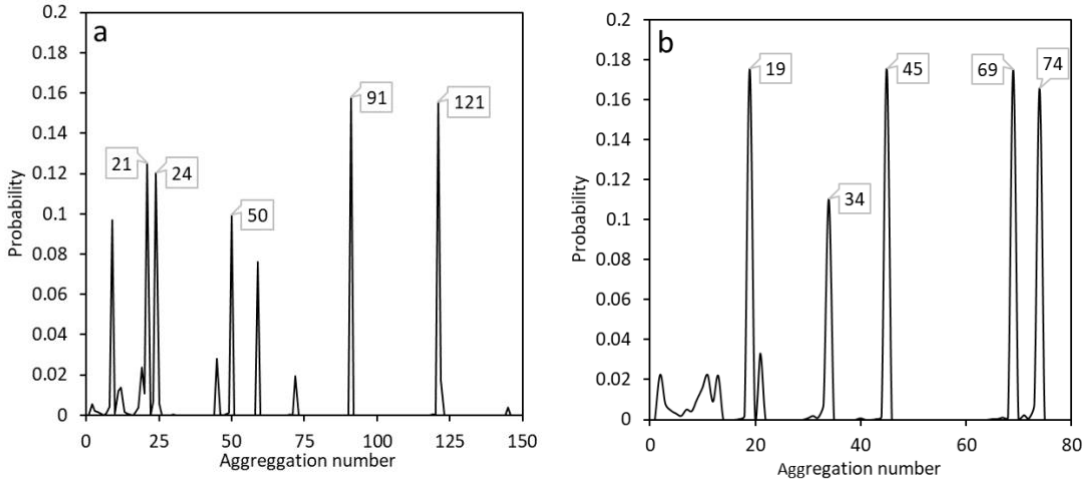


Figure 5: Probability distribution for aggregation numbers for the SLES/CA system (a) and the SLES/PA system (b).

4.1.2 Density profile and probability distribution of terminal atoms

The free energy profiles of fatty acids in the mixed SLES/FA micelles depend on how fatty acid and SLES molecules assemble. In order to have a better understanding of the predicted free energy, we calculated the density profiles of SLES, fatty acids, and water, as well as the probability distribution of hydrophobic and hydrophilic groups of SLES and fatty acids within the selected micelles. The density profiles of capric acid, SLES and water in the largest SLES/CA micelle are shown in Figure 6a and those of palmitic acid, SLES and water in the largest SLES/PA micelle are shown in Figure 6b. In the mixed SLES/CA micelle, the water profile first intersects SLES at r = 2.17 nm and then with capric acid at $r_{micelle}$ = 2.31 nm. The latter distance is therefore chosen as the radius of the micelle. For the SLES/CA micelle the density of capric acid is higher than that of the SLES at the micelle/water interface. The water profile intersects first with palmitic acid at r = 1.74 nm and then with SLES at $r_{micelle}$ = 2.07 nm. This is due to the concentration of palmitic acid in the micelle being much lower than that of capric acid. The density of palmitic acid is higher at the centre of the micelle, whereas for capric acid the maximum value occurs at 1.5 nm, due to palmitic acid having a longer carbon

chain than capric acid. Palmitic acid and SLES have the same number of carbon atoms (C16), therefore the carbon chains of palmitic acid molecules reach the centre of mass of the micelle. The radius of gyration for the two extracted micelles was computed using the following equation suggested by Bogusz *et al.*[61]:

$$R_{SLES/FA} = \sqrt{5/3}\, R_{gyration} \quad (3)$$

The calculated mean radii for the mixed SLES/CA and the SLES/PA micelles were $R_{SLES/CA}$ = 2.21 nm and $R_{SLES/PA}$ = 1.97 nm, respectively, which are in good agreement with the radii of the micelles selected from the density profiles, i.e. 2.31 nm for SLES/CA and 2.07 nm for SLES/PA.

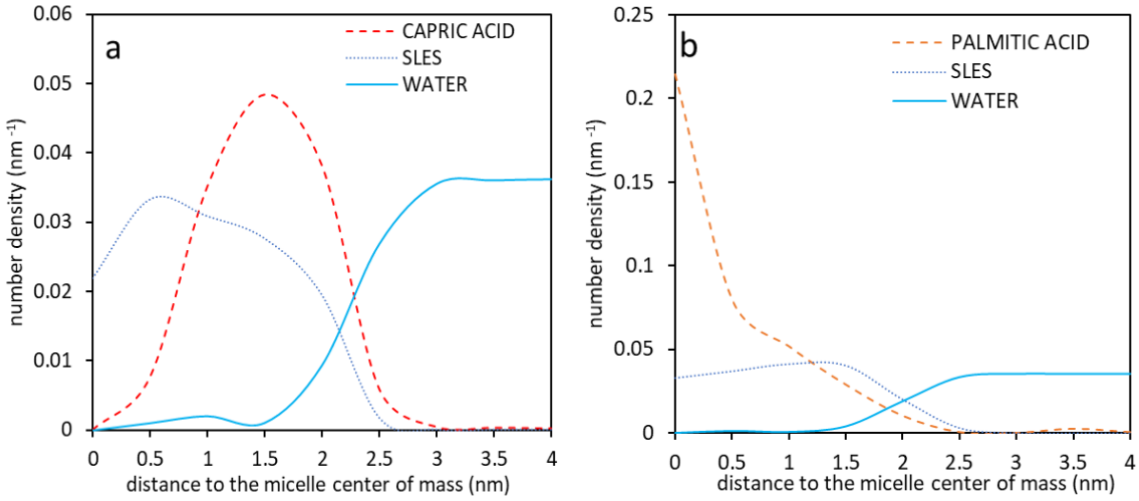


Figure 6: Density profiles for SLES, capric acid and water (a) and SLES, palmitic acid and water (b) in the MD simulated mixed SLES/FA micelle assemblies.

Figure 7 shows the probability distributions of the hydrophobic groups of the terminal carbon atoms of SLES and fatty acids, and hydrophilic groups of the sulfate group of the SLES molecule and the hydroxyl group of fatty acids in the two self-assembled micelles. For the mixed SLES/PA micelle, the probability distributions of SLES_C16 and PALMITIC_C16

overlapped due to that the SLES and palmitic acid have the same carbon chain lengths. This is not the case for the SLES/CA micelle where the mismatch in the molecular carbon chains (10 carbons for capric acid and 16 for SLES) resulted in the shifting of the CAPRIC_C10 distribution towards the micelle/water interface. The probability distribution of the hydrophobic groups provides a preliminary estimation of where the fatty acid solute will preferentially distribute in the micelle. Capric acid atoms preferentially distribute between the peak of the distribution for the CAPRIC _C10 atoms at 1.58 nm and the peak of the distribution for the CAPRIC _OH atoms at 2.08 nm from the COM of the micelle. In the case of palmitic acid, the hydrophobic atoms distribute between the peak of the distribution for the PALMITIC _C16 atoms at 1.15 nm and the peak of the distribution for the PALMITIC _OH atom and 1.83 nm.

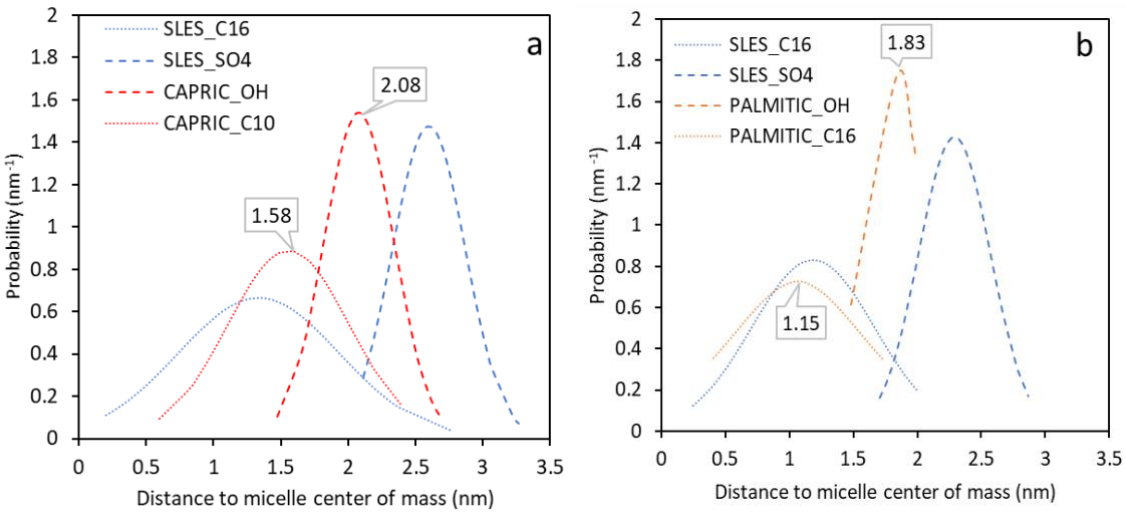


Figure 7: Probability distributions of hydrophobic and hydrophilic reference groups for (a) the SLES/CA mixed micelle and (b) the SLES/PA mixed micelle.

4.2 Free energy profiles and micelle/water partition coefficients

4.2.1 Neutral solutes

Free energy profiles for the transfer of capric acid and palmitic acid solutes from water to the respective mixed SLES/PA micelle are presented in Figure 8a (capric acid) and 8b (palmitic acid). The free energy profile for the capric acid solute in the mixed SLES/CA micelle was calculated by means of COSMOmic and US/PMF; for the latter both the non-polarizable (US_CGenFF) and the polarizable (Drude) force fields were used. When using US/PMF, free energy profiles were calculated for the intervals of 5-6 ns, 5-8 ns and 5-10 ns in order to further estimate the extent of convergence. In Figures 8a and 8b, free energy profiles obtained with US/PMF using the non-polarizable and the polarizable force fields for the interval 5-10 ns are shown as solid lines whereas error bars represent the differences of the free energies between the values computed in the intervals 5-6 ns and 5-8 ns and the one computed in the interval 5-10 ns. The energy profile predicted by the US/PMF with the polarizable force field (US/PMF-P) is about ~20 kJ/mole higher at the centre of the micelle than those predicted by the US/PMF with the non-polarizable force field (US/PMF-N) and COSMOmic. The free energy predicted by the US/PMF-N and COSMOmic showed a rapid decrease in the first 0.5 nm and 0.7 nm and remained high. The free energy remains constant in the plateau region where only the SLES carbon chain is present; this is also evident from the probability distribution of terminal carbon atoms (Figure 7a). The end of the plateau corresponds to the beginning of the capric acid molecules carbon chain and therefore a subsequent drop in the free energy occurs. The plateau region is not observed in the US/PMF-P as the free energy continuously decreased to ~15 kJ/mol from 0.5 nm to 1 nm along the reaction coordinate. This steady decrease indicates the higher electrostatic potential of the induced dipole. This dipole effect cannot be captured by the non-polarizable force field [32,62,63]. From 1 to 3.5 nm there are no major differences between the free energy profiles of US/PMF-P and that of US/PMF-N. The minima of all three profiles

occurred between 1.5 nm and 2 nm from the centre of the micelle, in agreement with the results from the probability distributions. The minima of the US/PMF-N and COSMOmic profiles are close to 2 nm, whereas the minimum of the polarizable force field is close to 1.5 nm. With regards to the mixed SLES/PA micelle, the minima of US/PMF-N and COSMOmic are located just after 1.5 nm from the centre of the micelle, in agreement with the results from the probability distributions. The minima for palmitic acid are closer to the centre of the micelle, in agreement with the results from density profiles which showed a maximum for the density of palmitic acid at the centre of the micelle.

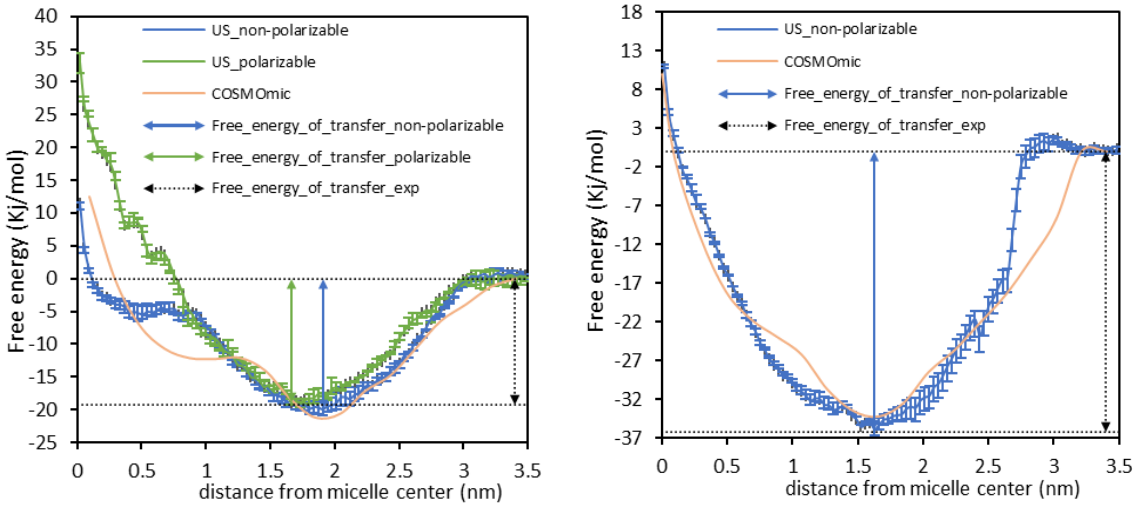


Figure 8: Free energy profiles for the transfer of fatty acids from water to the mixed SLES/PA micelle, (a) for capric acid and (b) for palmitic. Error bars in the US profiles represent the differences of the free energies between the values computed in the intervals 5-6 ns and 5-8 ns and the one computed in the interval 5-10 ns.

The free energies of transfer of fatty acids were used for calculating the micelle/water partition coefficients as given in Table 3. The predicted values for the free energies of transfer of capric acid and palmitic acid are very close to the experimental values for both the non-polarizable and polarizable force field US/PMF simulations. From Table 3 we can see that COSMOmic also performed well in predicting the partition coefficients of the neutral fatty acids, capric and palmitic acids, in both the mixed SLES/CA and SLES/PA micelles. The partition coefficient

for palmitic acid is slightly underpredicted by COSMOmic since COSMOmic tends to underpredict the partition coefficients of highly hydrophobic compounds[64]. The prediction errors of US/PMF using the non-polarizable force field are comparable with COSMOmic, although the predicted value for palmitic acid is slightly closer to the experimental data than COSMOmic. The US/PMF-P showed a good prediction accuracy for the partition coefficient of capric acid, in line with values predicted by COSMOmic and the non-polarizable force field.

Table 3: Predicted free energies of transfer and micelle/water partition coefficients for capric acid and palmitic acid, by US/PMF-P, US/PMF-P and COSMOmic methods. The averaged absolute differences, calculated as the average between the values of the intervals 5-10 ns and 5-8 ns and those of the intervals 5-10 ns and 5-6 ns, are reported in parenthesis next to reported values computed at 10 ns.

Capric acid						
ΔG_{transf}				ΔG_{transf} Difference vs. Exp.		
Non-polarizable force field	Polarizable force field		exp	US_ CGenFF	Polarizable force field	
-19.91 (0.47)	-18.89 (0.57)		-19.24	0.67	-0.65	
Log $K_{mic/w}$				Log $K_{mic/w}$ Difference vs. Exp.		
Non-polarizable force field	Polarizable force field	COSMOmic	exp	Non-polarizable force field	Polarizable force field	COSMOmic
3.49	3.31	3.52	3.37	0.12	-0.06	0.15
Palmitic Acid						
ΔG_{transf}				ΔG_{transf} Difference vs. Exp.		
Non-polarizable force field	Polarizable force field		exp	Non-polarizable force field	Polarizable force field	
-35.57 (0.35)	N.E.		-36.2	-0.63	N.E.	
Log $K_{mic/w}$				Log $K_{mic/w}$ Difference vs. Exp.		
Non-polarizable force field	Polarizable force field	COSMOmic	exp	Non-polarizable force field	Polarizable force field	COSMOmic
6.23	N.E.	5.99	6.36	-0.13	N.E.	-0.37

*Data collected from Tzocheva *et al.*[23]

4.2.2 Caprate anion

The free energy profiles for the transfer of the charged fatty acid caprate from water to the mixed SLES/CA micelle are shown in Figure 9. As with the neutral solutes, the predicted free

energy profiles of the US/PMF with both force fields in the interval 5-10 ns are shown as solid lines with the error bars representing the differences of the free energies computed in the intervals 5-6 ns and 5-8 ns and the one of the interval 5-10 ns.

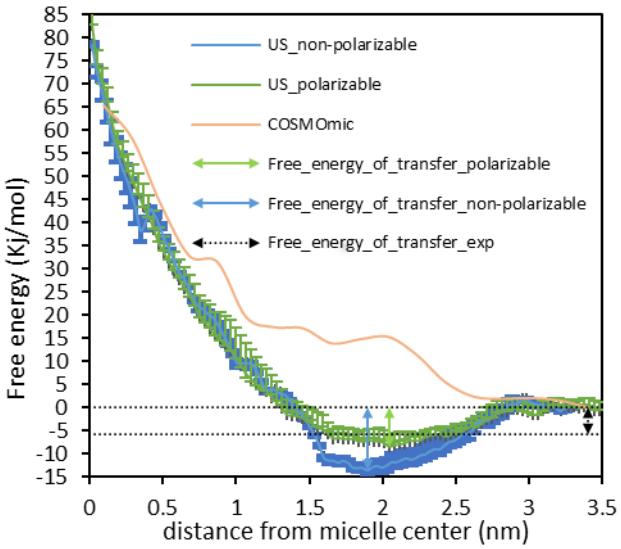


Figure 9: Free energy profiles for the transfer of caprate anion from water to the mixed SLES/CA micelle. Error bars represent the differences of the free energies computed in the intervals 5-6 ns and 5-8 ns and the one computed in the interval 5-10 ns.

The free energy profiles for the transfer of caprate predicted by the US/PMF-P and the US/PMF-P are similar in shape. The free energy of caprate at the centre of the micelle predicted by the US/PMF-P is higher than the one predicted by the US/PMF-P but the difference is relatively small. Towards the micelle/water interface the energy profile of caprate from the US/PMF-P shows a less favourable minimum compared to that of the non-polarizable US/PMF. In the interfacial region, where the anionic oxygen of caprate interacts with the sulfate anion of SLES, the induced polarization effect is an important feature that can be captured by the polarizable force field but not by the non-polarizable force field. Values for the free energies of transfer for caprate, predicted by the US/PMF with both force fields are compared with the experimental data in Table 4. The predicted value obtained from the US/PMF-P is much closer to the experimental value compared to the non-polarizable US/PMF.

The minimum in the free energy profile as predicted by US/PMF-P results in a higher value of the free energy of transfer and consequently a considerably improved prediction of the logarithm of the micelle-water partition coefficient compared to the non-polarizable US/PMF. The COSMOmic profile shows a minimum in correspondence with the water phase and the free energy profile does not assume negative values along the micelle radius. This results in a positive value for the free energy of transfer and, consequently, in a negative value for the logarithm of the micelle-water partition coefficient (Log $K_{mic/w}$).

In Table 4 both COSMOmic and US/PMF-P showed substantial disagreements for the micelle-water partition coefficient. US/PMF-P improves considerably the prediction of the caprate partition coefficient, yielding an absolute prediction difference of 0.43 that is almost more than one logarithmic unit smaller than the ones of COSMOmic (-1.83) and US/PMF-N (1.31). Thus, the use of a polarizable force field is necessary for improving the prediction accuracy of the micelle-water partition coefficients of charged solutes in anionic micelles.

Table 4: Free energies of transfer and micelle/water partition coefficients for caprate solute in the mixed SLES/CA micelle, predicted by US/PMF-P and US/PMF-P and COSMOmic. The averaged absolute errors are calculated between the values of the intervals 5-10 ns and 5-8 ns and those between 5-10 ns and 5-6 ns are reported in parenthesis next to reported values of the interval 5-10 ns.

ΔG_{transf}				ΔG_{transf} Difference vs Exp.		
Non-polarizable force field	Polarizable force field		exp	Non-polarizable force field	Polarizable force field	
-13.14 (0.64)	-8.41 (0.57)		-5.88	-7.53	-2.53	
Log $K_{mic/w}$				Log $K_{mic/w}$ Difference vs. Exp.		
Non-polarizable force field	Polarizable force field	COSMOmic	exp	Non-polarizable force field	Polarizable force field	COSMOmic
2.35	1.47	-0.8	1.03	1.31	0.43	-1.83

*Data collected from Tzocheva *et al*[23].

Figure 10 shows the hydrophilic group of caprate solute at the minimum of the profiles for the two force fields. With the polarizable force field, the hydrophilic group is more embedded in the water phase (Figure 10b). In contrast, the ionic group is more shifted towards the micellar

phase in the case of the non-polarizable force field prediction (Figure 10a). This difference is consistent with the higher value of the free energy of transfer predicted by the polarizable force field, which indicates a higher affinity of the solute for the water phase. Such high affinity is consistent with the presence of the electronic polarization, leading to the more favourable interactions of the charged head group with the aqueous environment.

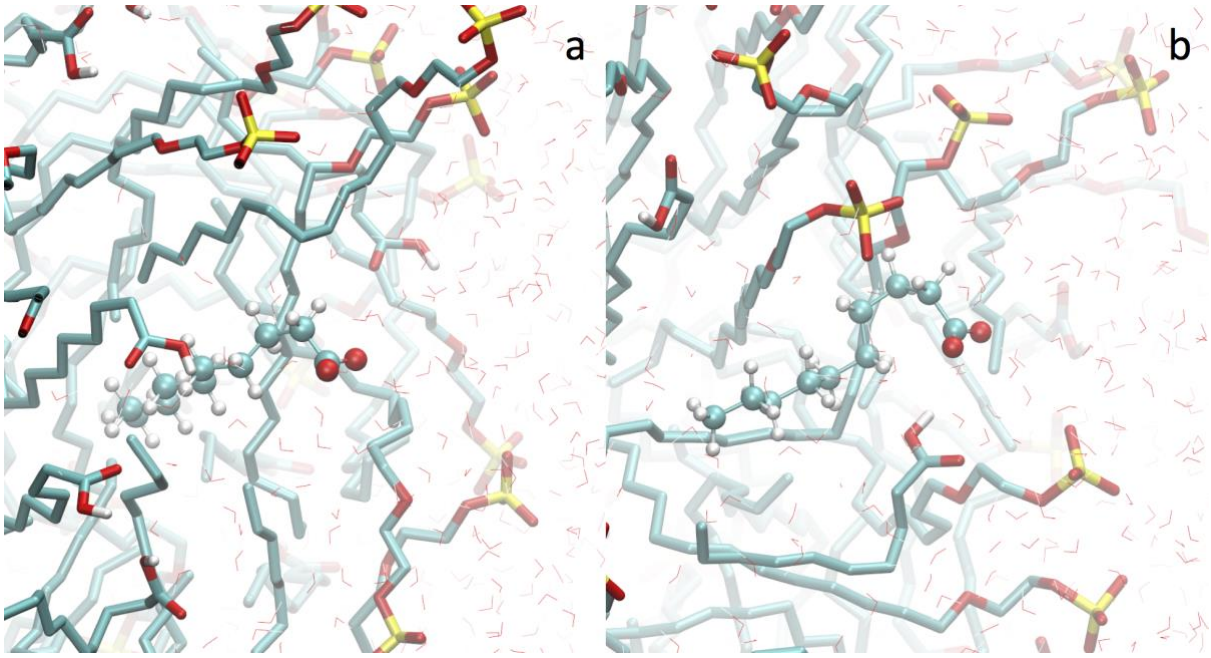


Figure 10: Capric acid anion (caprate), at the minima of the free energy, interacting with capric acid and SLES molecules predicted by the non-polarizable force field (a) and polarizable force field (b). Capric acid and SLES atoms are represented as pipes (hydrogen atoms are excluded for clarity), water atoms as lines and caprate atoms as spheres. Carbon atoms are colored in cyan, hydrogen atoms in white, oxygen atoms in red, sulfur atoms in yellow; and water molecules are colored in red.

Finally, the behaviour of sodium ions when an ionic species, such as caprate solute, is absorbed in the micelle is analyzed. In figure 11 the probability distributions of sodium ions, carboxyl oxygen atoms of capric acid and sulphur atoms of SLES in the mixed SLES/capric acid micelle are shown for both the non-polarizable and the polarizable systems. As is evident, both the peaks of probability distributions for carboxyl oxygens (2.30 and 2.31 nm for the polarizable and the non-polarizable systems, respectively) and sulphurs (2.51 and 2.55 nm, respectively) are close to each other for the polarizable and the non-polarizable systems. On the contrary, the peaks of the probability distributions of sodium ions are at 2.75 nm and 2.97 nm from the

micelle center of mass for the polarizable and the non-polarizable systems, respectively. In both models the sodium condensed in the vicinity of the negatively charged micelle, as expected based on simple electrostatic consideration. However, with polarizable force field the sodium ions cluster at a shorter distance from the micelle center of mass, suggesting that attractive interactions between the negatively charged micelle and sodium cations, as modelled by the polarizable force field, are stronger as compared to the additive force field.

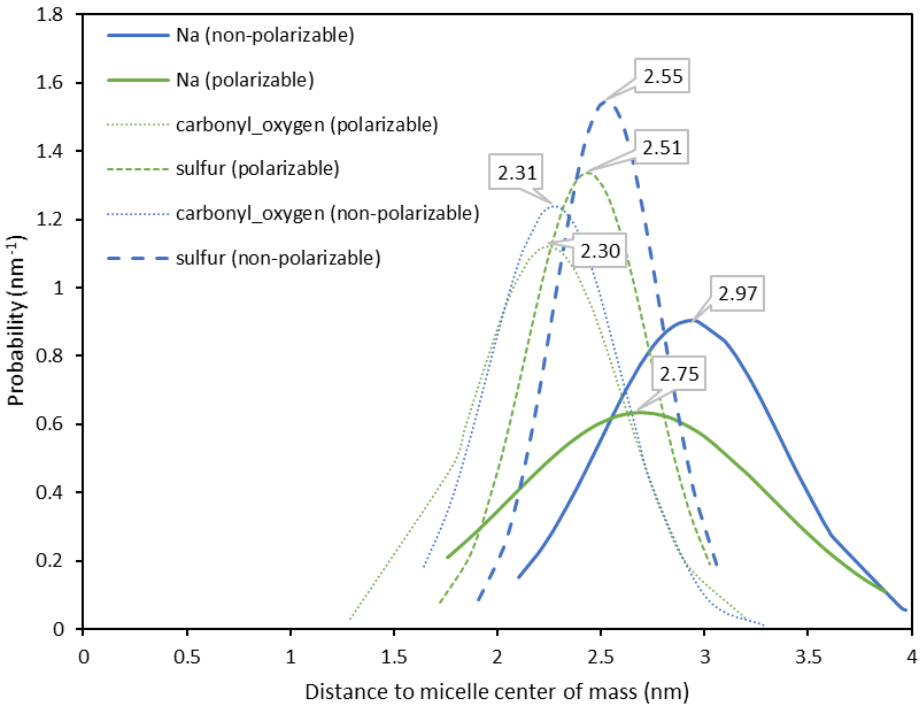


Figure 11: Probability distributions of sodium ions around the mixed SLES/capric acid micelle for the non-polarizable and the polarizable systems when the caprate anion is at its minimum location in the PMFs.

5 Conclusions

In this paper, the performance of the MD/COSMOmic and the US/PMF approaches for predicting solute partition coefficients of fatty acids in the mixed SLES/FA micelles have been evaluated. MD simulations have been performed to obtain structures from the self-assembly of the SLES/CA and the SLES/PA mixed micelles. A number of micelles were formed for both simulated systems and the largest micelle of each system was selected to predict solute partition

coefficients of capric acid, palmitic acid and caprate anion in the mixed micelles. The selected micelles of the two systems were first entered into COSMOmic. The predicted micelle/water partition coefficients showed good agreement with the experimental data for the neutral solutes of capric acid and palmitic acid, but a severe underprediction occurred for the micelle-water partition coefficient of the capric acid anion (caprate). Subsequently, the US/PMF approach has been explored for predicting partition coefficients using both polarizable and non-polarizable force fields. The results of the non-polarizable force field produced similar results to COSMOmic for neutral solutes but not for the charged solute, showing that the use of a non-polarizable force field is limited to good prediction accuracy for neutral solutes. Good prediction accuracy for the micelle-water partition coefficient of the charged caprate solute was achieved by using the Drude polarizable force field. Thus, for simulating charged solutes in anionic surfactant micelles, the use of an accurate polarizable force field is crucial to model the electrostatic interaction. The use of the polarizable force field improved the ability to accurately model the dipole potential as previously reported for the lipid bilayer[65]. The present results indicate that with the rapid progress in the development of polarizable force fields, such as the CHARMM Drude polarizable force field, *in-silico* prediction of partition coefficients for charged solutes will be a more robust low-cost alternative to laboratory experiments.

Conflict of Interest

ADM Jr. is co-founder and CSO of SilcsBio LLC.

ASSOCIATED CONTENT

Supporting Information

Derivation of experimental data for free energy of transfer and water/micelle partition coefficient of caprate. Convergence analysis for PMF of caprate for the polarizable and the non-polarizable systems. Parametrization of SLES, capric acid and caprate ion with the polarizable Drude FF.

Acknowledgements

We thank the National Institutes of Health [GM131710] for financial support for this work and the Computer-Aided Drug Design Center at the University of Maryland Baltimore for computing time. We appreciate the financial support provided by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 675251. We thank Prof. Svetoslav Anachkov and Dr. Gergana S. Georgieva, from University of Sofia, for the fruitful discussion about experimental data relative to mixed SLES/fatty acid micelles.

References

(1) Mirrlees, M. S.; Moulton, S. J.; Murphy, C. T.; Taylor, P. J. Direct Measurement of Octanol-Water Partition Coefficients by High-Pressure Liquid Chromatography. *J. Med. Chem.* **1976**, *19* (5), 615–619. https://doi.org/10.1021/jm00227a008.

(2) Valkó, K. Application of High-Performance Liquid Chromatography Based Measurements of Lipophilicity to Model Biological Distribution. *J. Chromatogr. A* **2004**, *1037* (1–2), 299–310. https://doi.org/10.1016/j.chroma.2003.10.084.

(3) Kelly, K. A.; Burns, S. T.; Khaledi, M. G. Prediction of Retention in Micellar Electrokinetic Chromatography from Solute Structure. 1. Sodium Dodecyl Sulfate Micelles. *Anal. Chem.* **2001**, *73* (24), 6057–6062. https://doi.org/10.1021/ac0105944.

(4) Burns, S. T.; Khaledi, M. G. Predictions of Micelle-Water Partition Coefficients and Retention in Micellar Electrokinetic Chromatography from Solute Structure. 2. Fragmental Constant Approach. *Anal. Chem.* **2004**, *76* (18), 5451–5458. https://doi.org/10.1021/ac0498718.

(5) Ishihama, Y.; Oda, Y.; Uchikawa, K.; Asakawa, N. Evaluation of Solute Hydrophobicity by Microemulsion Electrokinetic Chromatography. *Anal. Chem.* **1995**, *67* (9), 1588–1595. https://doi.org/10.1021/ac00105a018.

(6) Klamt, A.; Eckert, F.; Diedenhofen, M. Prediction or Partition Coefficients and Activity Coefficients of Two Branched Compounds Using COSMOtherm. *Fluid Phase Equilib.* **2009**, *285* (1–2), 15–18. https://doi.org/10.1016/j.fluid.2009.05.010.

(7) Klamt, A. Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. *J. Phys. Chem.* **1995**, *99* (7), 2224–2235. https://doi.org/10.1021/j100007a062.

(8) Eckert, F.; Klamt, A. Fast Solvent Screening via Quantum Chemistry: COSMO-RS Approach. *AIChE J.* **2002**, *48* (2), 369–385. https://doi.org/10.1002/aic.690480220.

(9) Klamt, A.; Huniar, U.; Spycher, S.; Keldenich, J. COSMOmic: A Mechanistic Approach to the Calculation of Membrane−Water Partition Coefficients and Internal Distributions within Membranes and Micelles. *J. Phys. Chem. B* **2008**, *112* (38), 12148–12157. https://doi.org/10.1021/jp801736k.

(10) Turchi, M.; Cai, Q.; Lian, G. An Evaluation of In-Silico Methods for Predicting Solute Partition in Multiphase Complex Fluids – A Case Study of Octanol/Water Partition Coefficient. *Chem. Eng. Sci.* **2019**, 150–158. https://doi.org/10.1016/j.ces.2018.12.003.

(11) Storm, S.; Jakobtorweihen, S.; Smirnova, I.; Panagiotopoulos, A. Z. Molecular Dynamics Simulation of SDS and CTAB Micellization and Prediction of Partition Equilibria with COSMOmic. *Langmuir* **2013**, *29* (37), 11582–11592. https://doi.org/10.1021/la402415b.

(12) Ingram, T.; Storm, S.; Kloss, L.; Mehling, T.; Jakobtorweihen, S.; Smirnova, I. Prediction of Micelle/Water and Liposome/Water Partition Coefficients Based on Molecular Dynamics Simulations, COSMO-RS, and COSMOmic. *Langmuir* **2013**, *29* (11), 3527–3537. https://doi.org/10.1021/la305035b.

(13) Turchi, M.; Lian, G.; Cai, Q.; Wood, I.; Rabone, J.; Noro, M. Multi-Scale Modelling of Solute Partition Equilibria of Micelle-Water and Microemulsion-Water Systems Using Molecular Dynamics and COSMOtherm. *Computer Aided Chemical Engineering*. 2017, pp 2773–2778. https://doi.org/10.1016/B978-0-444-63965-3.50464-5.

(14) Jakobtorweihen, S.; Yordanova, D.; Smirnova, I. Predicting Critical Micelle Concentrations with Molecular Dynamics Simulations and COSMOmic. *Chemie-Ingenieur-Technik* **2017**, *89* (10), 1288–1296. https://doi.org/10.1002/cite.201700061.

(15) Yordanova, D.; Ritter, E.; Gerlach, T.; Jensen, J. H.; Smirnova, I.; Jakobtorweihen, S. Solute Partitioning in Micelles: Combining Molecular Dynamics Simulations, COSMOmic, and Experiments. *J. Phys. Chem. B* **2017**, *121* (23), 5794–5809. https://doi.org/10.1021/acs.jpcb.7b03147.

(16) Yordanova, D.; Ritter, E.; Smirnova, I.; Jakobtorweihen, S. Micellization and Partition Equilibria in Mixed Nonionic/Ionic Micellar Systems: Predictions with Molecular Models. *Langmuir* **2017**, *33* (43), 12306–12316. https://doi.org/10.1021/acs.langmuir.7b02813.

(17) Turchi, M.; Cai, Q.; Lian, G. In Silico Prediction of the Thermodynamic Equilibrium of Solute Partition in Multiphase Complex Fluids: A Case Study of Oil-Water Microemulsion. *Langmuir* **2019**. https://doi.org/10.1021/acs.langmuir.9b01513.

(18) Engler, E. M.; Andose, J. D.; Schleyer, P. V. R. Critical Evaluation of Molecular Mechanics. *J. Am. Chem. Soc.* **1973**, *95* (24), 8005–8025.

(19) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23* (2), 187–199. https://doi.org/10.1016/0021-9991(77)90121-8.

(20) Virnau, P.; Müler, M. Calculation of Free Energy through Successive Umbrella Sampling. *J. Chem. Phys.* **2004**, *120* (23), 10925–10930. https://doi.org/10.1063/1.1739216.

(21) Hawkins, G. R.; Kowcz, O. A. Peroxide-Containing Conditioning Shampoo. Google Patents April 1987.

(22) Smith, W. P.; Dunn, L. J. Hair Conditioning Shampoo. Google Patents July 1985.

(23) Tzocheva, S. S.; Kralchevsky, P. A.; Danov, K. D.; Georgieva, G. S.; Post, A. J.; Ananthapadmanabhan, K. P. Solubility Limits and Phase Diagrams for Fatty Acids in Anionic (SLES) and Zwitterionic (CAPB) Micellar Surfactant Solutions. *J. Colloid Interface Sci.* **2012**, *369* (1), 274–286. https://doi.org/https://doi.org/10.1016/j.jcis.2011.12.036.

(24) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; et al. CHARMM General Force Field: A Force Field for Drug-like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **2010**, *31* (4), 671–690. https://doi.org/10.1002/jcc.21367.

(25) Vollhardt, D.; Czichocki, G.; Rudert, R. Effect of the Molecular Structure on the Adsorption of Alkyl Ether Sulphates and Alkane Ether Sulphonates at the Air–Water Interface. *Colloids Surfaces A Physicochem. Eng. Asp.* **1998**, *142* (2), 315–322. https://doi.org/https://doi.org/10.1016/S0927-7757(98)00356-2.

(26) Mysels, K. J.; Otter, R. J. Thermodynamic Aspects of Mixed Micelles-Application to an Empirically Established Equilibrium. *J. Colloid Sci.* **1961**, *16* (5), 474–480. https://doi.org/10.1016/0095-8522(61)90024-1.

(27) Bourrel, M.; Schechter, R. S. No Title. *Microemulsions Relat. Syst.* **1988**.

(28) Anisimov, V. M.; Lamoureux, G.; Vorobyov, I. V; Huang, N.; Roux, B.; MacKerell Jr., A. D. Determination of Electrostatic Parameters for a Polarizable Force Field Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* **2005**, *1* (1), 153–168. https://doi.org/10.1021/ct049930p.

(29) Lamoureux, G.; Roux, B. Modeling Induced Polarization with Classical Drude Oscillators: Theory and Molecular Dynamics Simulation Algorithm. *J. Chem. Phys.* **2003**, *119* (6), 3025–3039. https://doi.org/10.1063/1.1589749.

(30) Lamoureux, G.; MacKerell, A. D.; Roux, B. A Simple Polarizable Model of Water Based on Classical Drude Oscillators. *J. Chem. Phys.* **2003**, *119* (10), 5185–5197.

https://doi.org/10.1063/1.1598191.

(31) Lemkul, J. A.; Huang, J.; Roux, B.; MacKerell, A. D. An Empirical Polarizable Force Field Based on the Classical Drude Oscillator Model: Development History and Recent Applications. *Chem. Rev.* **2016**, *116* (9), 4983–5013. https://doi.org/10.1021/acs.chemrev.5b00505.

(32) Li, H.; Chowdhary, J.; Huang, L.; He, X.; MacKerell, A. D.; Roux, B. Drude Polarizable Force Field for Molecular Dynamics Simulations of Saturated and Unsaturated Zwitterionic Lipids. *J. Chem. Theory Comput.* **2017**, *13* (9), 4535–4552. https://doi.org/10.1021/acs.jctc.7b00262.

(33) Chowdhary, J.; Harder, E.; Lopes, P. E. M.; Huang, L.; MacKerell, A. D.; Roux, B. A Polarizable Force Field of Dipalmitoylphosphatidylcholine Based on the Classical Drude Model for Molecular Dynamics Simulations of Lipids. *J. Phys. Chem. B* **2013**, *117* (31), 9142–9160. https://doi.org/10.1021/jp402860e.

(34) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. https://doi.org/https://doi.org/10.1016/j.softx.2015.06.001.

(35) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926–935. https://doi.org/10.1063/1.445869.

(36) Durell, S. R.; Brooks, B. R.; Ben-Naim, A. Solvent-Induced Forces between Two Hydrophilic Groups. *J. Phys. Chem.* **1994**, *98* (8), 2198–2202. https://doi.org/10.1021/j100059a038.

(37) Steinbach, P. J.; Brooks, B. R. New Spherical-Cutoff Methods for Long-Range Forces in Macromolecular Simulation. *J. Comput. Chem.* **1994**, *15* (7), 667–683. https://doi.org/10.1002/jcc.540150702.

(38) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98* (12), 10089–10092. https://doi.org/10.1063/1.464397.

(39) Nosé, S.; Klein, M. L. Constant Pressure Molecular Dynamics for Molecular Systems. *Mol. Phys.* **1983**, *50* (5), 1055–1076. https://doi.org/10.1080/00268978300102851.

(40) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52* (12), 7182–7190. https://doi.org/10.1063/1.328693.

(41) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D. OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. *PLoS Comput. Biol.* **2017**, *13* (7), e1005659.

(42) Lamoureux, G.; Harder, E.; Vorobyov, I. V; Roux, B.; MacKerell Jr., A. D. A Polarizable Model of Water for Molecular Dynamics Simulations of Biomolecules. *Chem. Phys. Lett.* **2006**, *418* (1–3), 245–249. https://doi.org/10.1016/j.cplett.2005.10.135.

(43) Sammalkorpi, M.; Karttunen, M.; Haataja, M. Structural Properties of Ionic Detergent Aggregates: A Large-Scale Molecular Dynamics Study of Sodium Dodecyl Sulfate. *J. Phys. Chem. B* **2007**, *111* (40), 11722–11733. https://doi.org/10.1021/jp072587a.

(44) Koneva, A. S.; Ritter, E.; Anufrikov, Y. A.; Lezov, A. A.; Klestova, A. O.; Smirnova, N. A.; Safonova, E. A.; Smirnova, I. Mixed Aqueous Solutions of Nonionic Surfactants Brij 35/Triton X-100: Micellar Properties, Solutes' Partitioning from Micellar Liquid Chromatography and Modelling with COSMOmic. *Colloids Surfaces A Physicochem. Eng. Asp.* **2018**, *538*, 45–55. https://doi.org/https://doi.org/10.1016/j.colsurfa.2017.10.044.

(45) Storm, S.; Jakobtorweihen, S.; Smirnova, I. Solubilization in Mixed Micelles Studied by Molecular Dynamics Simulations and COSMOmic. *J. Phys. Chem. B* **2014**, *118* (13), 3593–3604. https://doi.org/10.1021/jp410636w.

(46) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New Feathers for an Old Bird. *Comput. Phys. Commun.* **2014**, *185* (2), 604–613. https://doi.org/https://doi.org/10.1016/j.cpc.2013.09.018.

(47) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. THE Weighted Histogram Analysis Method for Free-energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13* (8), 1011–1021. https://doi.org/10.1002/jcc.540130812.

(48) Grossfield, A. WHAM: The Weighted Histogram Analysis Method, Version 2.0. 9. *Available Membr. urmc. rochester. edu/content/wham. Accessed Novemb.* **2013**, *15*, 2013.

(49) Trzesniak, D.; Kunz, A. E.; van Gunsteren, W. F. A Comparison of Methods to Compute the Potential of Mean Force. *ChemPhysChem* **2007**, *8* (1), 162–169.

(50) Khavrutskii, I. V; Dzubiella, J.; McCammon, J. A. Computing Accurate Potentials of Mean Force in Electrolyte Solutions with the Generalized Gradient-Augmented Harmonic Fourier Beads Method. *J. Chem. Phys.* **2008**, *128* (4), 44106. https://doi.org/10.1063/1.2825620.

(51) Paci, E.; Ciccotti, G.; Ferrario, M.; Kapral, R. Activation Energies by Molecular Dynamics with Constraints. *Chem. Phys. Lett.* **1991**, *176* (6), 581–587.

(52) Klamt, A.; Jonas, V.; Bürger, T.; Lohrenz, J. C. W. Refinement and Parametrization of COSMO-RS. *J. Phys. Chem. A* **1998**, *102* (26), 5074–5085. https://doi.org/10.1021/jp980017s.

(53) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38* (6), 3098–3100. https://doi.org/10.1103/PhysRevA.38.3098.

(54) Perdew, J. P. Density-Functional Approximation for the Correlation Energy of the Inhomogeneous Electron Gas. *Phys. Rev. B* **1986**, *33* (12), 8822–8824. https://doi.org/10.1103/PhysRevB.33.8822.

(55) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. Auxiliary Basis Sets to Approximate Coulomb Potentials (Chem. Phys. Letters 240 (1995) 283) (PII:0009-2614(95)00621-4). *Chem. Phys. Lett.* **1995**, *242* (6), 652–660. https://doi.org/10.1016/0009-2614(95)00838-U.

(56) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. Auxiliary Basis Sets for Main Row Atoms and Transition Metals and Their Use to Approximate Coulomb Potentials. *Theor. Chem. Acc.* **1997**, *97* (1–4), 119–124. https://doi.org/10.1007/s002140050244.

(57) Schäfer, A.; Huber, C.; Ahlrichs, R. Fully Optimized Contracted Gaussian Basis Sets of Triple Zeta Valence Quality for Atoms Li to Kr. *J. Chem. Phys.* **1994**, *100* (8), 5829–5835. https://doi.org/10.1063/1.467146.

(58) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. Electronic Structure Calculations on Workstation Computers: The Program System Turbomole. *Chem. Phys. Lett.* **1989**, *162* (3), 165–169. https://doi.org/10.1016/0009-2614(89)85118-8.

(59) Aoudia, M.; Al-Haddabi, B.; Al-Harthi, Z.; Al-Rubkhi, A. Sodium Lauryl Ether Sulfate Micellization and Water Solubility Enhancement towards Naphthalene and Pyrene: Effect of the Degree of Ethoxylation. *J. Surfactants Deterg.* **2010**, *13* (1), 103–111. https://doi.org/10.1007/s11743-009-1131-9.

(60) Anachkov, S. E.; Danov, K. D.; Basheva, E. S.; Kralchevsky, P. A.; Ananthapadmanabhan, K. P. Determination of the Aggregation Number and Charge of Ionic Surfactant Micelles from the

Stepwise Thinning of Foam Films. *Advances in Colloid and Interface Science*. 2012, pp 55–67. https://doi.org/10.1016/j.cis.2012.08.003.

(61) Bogusz, S.; Venable, R. M.; Pastor, R. W. Molecular Dynamics Simulations of Octyl Glucoside Micelles: Structural Properties. *J. Phys. Chem. B* **2000**, *104* (23), 5462–5470. https://doi.org/10.1021/jp000159y.

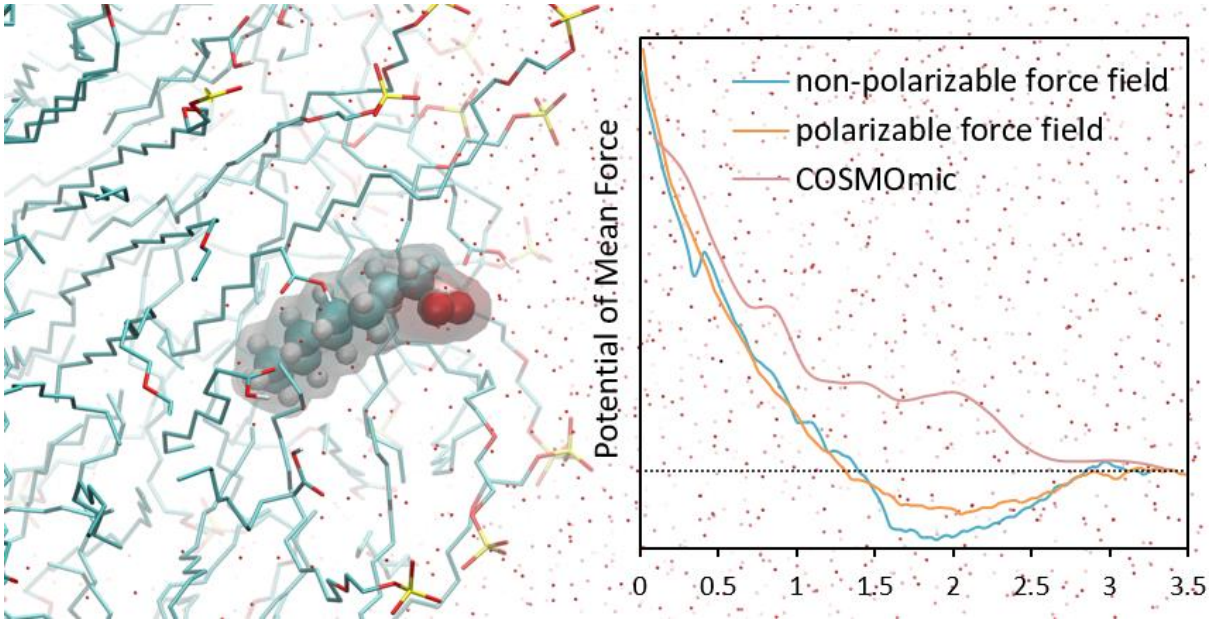
(62) Patel, S. A.; Brooks III, C. L. Revisiting the Hexane-Water Interface via Molecular Dynamics Simulations Using Nonadditive Alkane-Water Potentials. *J. Chem. Phys.* **2006**, *124* (20). https://doi.org/10.1063/1.2198538.

(63) Vorobyov, I. V; Anisimov, V. M.; MacKerell Jr., A. D. Polarizable Empirical Force Field for Alkanes Based on the Classical Drude Oscillator Model. *J. Phys. Chem. B* **2005**, *109* (40), 18988–18999. https://doi.org/10.1021/jp053182y.

(64) Jakobtorweihen, S.; Zuniga, A. C.; Ingram, T.; Gerlach, T.; Keil, F. J.; Smirnova, I. Predicting Solute Partitioning in Lipid Bilayers: Free Energies and Partition Coefficients from Molecular Dynamics Simulations and COSMOmic. *J. Chem. Phys.* **2014**, *141* (4), 45102. https://doi.org/10.1063/1.4890877.

(65) Harder, E.; MacKerell Jr., A. D.; Roux, B. Many-Body Polarization Effects and the Membrane Dipole Potential. *J. Am. Chem. Soc.* **2009**, *131* (8), 2760–2761. https://doi.org/10.1021/ja806825g.

Table of contents graphic



Supplementary Information

for

Predicting Partition Coefficients Of Neutral And Charged Solutes In The Mixed SLES-fatty Acids In Micellar System

Mattia Turchi,[†, ‡], *Abhishek A. Kognole,*[&] *Anmol Kumar,*[&] *Qiong Cai,*[‡], *Guoping Lian,*[†, ‡,]* and *Alexander D. MacKerell Jr*[&,]*

[†] Unilever Research Colworth, Colworth Park, Sharnbrook, Bedfordshire MK44 1LQ, UK,

[‡]Department of Chemical and Process Engineering, University of Surrey, Guildford GU27XH, UK

[&]University of Maryland Computer-Aided Drug Design Center, Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, MD 21201, USA.

Calculation of micelle/water partition coefficients, $K_{mic,A}$

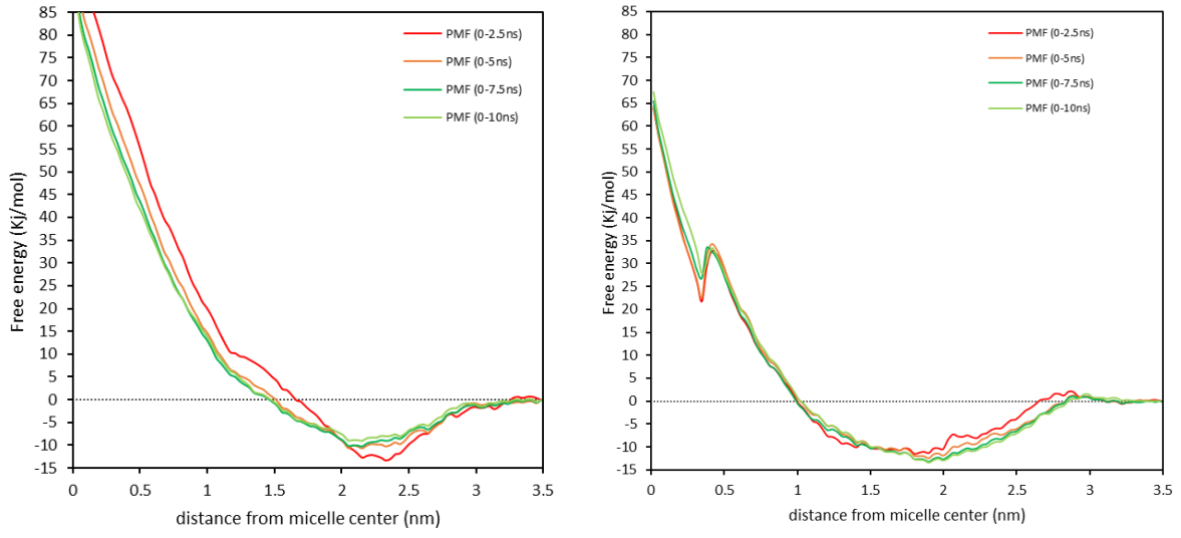
Log $K_{mic,A}$(mole/mol) for caprate ion was derived using data relative to the concentration of the monomer in water and in the micelle phase from Table A.1 of Tzocheva *et al*[1] using equation 26 of the same paper:

$$K_{mic/w} = \frac{\gamma_z y_z}{c_z} \qquad \text{Eq. 1}$$

In which γ_z and y_z are the activity coefficient and the molar fraction of caprate in the micellar phase and c_z is the concentration of caprate in the water phase. Free energy value was then calculated as: $\Delta G_{transf} = -2.303RT \text{ Log } K_{mic/w}$.

Convergence analysis for caprate PMF with the polarizable and the non-polarizable force fields

Figure S1 shows the convergence of the PMF for caprate at four different time portions of the windows (i.e. 0-2.5 ns, 0-5 ns, 0-7.5 ns and 0-10ns) when the non-polarizable force field (a) and the polarizable force field (b) are employed. As can be seen, in both cases the free energy surfaces in the portions 0-5 ns, 0-7.5 ns and 0-10 ns are very similar indicating that the free energy surfaces in the last 5ns of the PMF profiles have converged.



(a) **(b)**

Figure S1: Convergence of PMF for caprate at different time portions of the 36 windows for the **a)** non-polarizable force field and **b)** polarizable force field.

Parametrization of SLES, capric acid and caprate ion with the CHARMM Drude FF

Initial residue topologies and parameters for SLES, capric acid and caprate were assigned based on analogy with those already available in the Drude force field. In the case of SLES, a new atom-type was introduced for the sulfur atom in sulfate wherein the Lennard-Jones parameters were taken from Drude force field parameter of phosphorus atom of phosphate[2]. Since the missing parameters were mainly associated with the head-group of SLES, a prototype structure was prepared by attaching the head-group of SLES to a neutral methyl moiety. This procedure allows patching of the head-group parameters with the rest of the aliphatic chain of the concerned molecule. The head-group parameters were refined to improve the agreement with QM water-interactions, molecular polarizability and the intramolecular stretching, bending and twisting patterns along with the dihedral potential energy scan. The QM target data were generated by using the Gaussian03 software[3]. The structure of methylsulfate was initially optimized at MP2/6-31G*[4] level of theory. The optimized structure was then subjected to single point calculation at MP2/cc-pVQZ[5] level of theory to obtain the molecular polarizability. Target data for partial charges were based on water-methylsulfate dimer interaction (Figure S2) optimized at the MP2/aug-cc-pVDZ[6] model chemistry. The interaction energies between water and methylsulfate were calculated by subtracting their monomer energies obtained from single-point calculation at MP2/aug-cc-pVDZ level of theory from the total energy of water-methylsulfate complexes. A scan of the potential energy surface (PES) along the C-O-S-O dihedral was performed at MP2/aug-cc-pVDZ model chemistry.

The Drude-force field parameters of the head-group were parametrized using CHARMM-c42b1[7] targeting the above QM data. The bond and angle parameters were adjusted such that molecular mechanics (MM) optimized geometry matches the QM geometry. Values for QM and MM bonds, angles and dihedrals and their differences for the optimized geometries are reported in Table S1. The dihedral parameter of the C-O-S-O dihedral was modified to reproduce the QM PES. The plot of MM and QM PES against the C-O-S-O dihedral is given in Figure S3. The alpha and Thole parameters were adjusted to match QM molecular polarizability values and are reported in Table S2. The MM interaction energies of methylsulfate with water are reported in Table S3 in comparison to the corresponding QM values. The MM values are systematically less favourable than the QM values. However, this difference was not corrected based on a calculation of the free energy of aqueous solvation of methylsulfate, yielding a value of -84.05 kcal/mol, which is in good agreement with -80.12 kcal/mol as predicted by COSMOtherm program. Details of the free energy of aqueous solvation calculation are shown along with final residue topologies and parameters for not currently available in the Drude force fields are presented in Tables S4, S5 and S6.

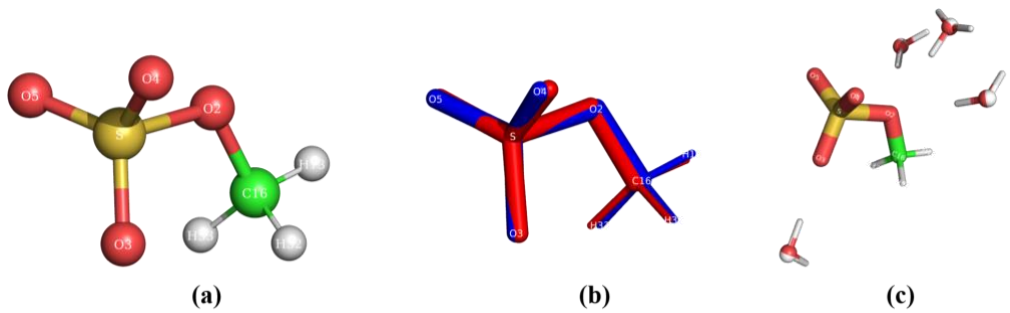


Figure S2: a) Molecular structure of minimized methylsulfate, **b)** comparison of 3D coordinates obtained from QM and MM after parametrization, and **c)** four different orientations of water interacting with methylsulfate used for fitting the MM charges combined in a single image.

Table S1: Internal coordinates of methylsulfate (C(H1)(H2)(H3)-O1-S(O2)(O3)(O4)-) obtained from optimized QM (MP2/6-31G*) and the Drude polarizable MM calculations.

#IC_LIST	QM_V	MM_V	ΔV(MM-QM)
C-O1	1.408	1.474	0.067
C-H1	1.107	1.111	0.004

C-H2	1.107	1.113	0.006
C-H3	1.107	1.113	0.006
S-O1	1.751	1.774	0.023
S-O2	1.493	1.499	0.006
S-O3	1.485	1.493	0.007
S-O4	1.485	1.493	0.007
H1-C-O1	106.498	109.631	3.132
H1-C-H2	108.620	108.396	-0.224
H1-C-H3	108.620	108.396	-0.224
O1-C-H2	112.990	110.591	-2.399
O1-C-H3	112.990	110.591	-2.399
H2-C-H3	106.995	109.178	2.183
C-O1-S	114.386	114.935	0.549
O1-S-O2	103.445	107.775	4.329
O1-S-O3	102.122	99.099	-3.023
O1-S-O4	102.122	99.099	-3.023
O2-S-O3	115.357	115.718	0.361
O2-S-O4	115.357	115.718	0.361
O3-S-O4	115.503	115.920	0.417
H1-C-O1-S	180.000	-180.000	0.000
H2-C-O1-S	-60.831	-60.533	-0.299
H3-C-O1-S	60.831	60.533	-0.298
C-O1-S-O2	0.003	0.000	-0.003
C-O1-S-O3	120.117	120.854	0.737
C-O1-S-O4	-120.111	-120.854	0.743

QM_V and MM_V are distances, angles and dihedral values obtained from QM and MM calculations, respectively. ΔV(MM-QM) is the difference between MM and QM values. Distance values are in Å, and angles and dihedral values are in degrees.

Table S2: Values of components of molecular polarizability (Pxx, Pyy and Pzz) and total molecular polarizability (P_{tot}) of methylsulfate obtained from QM (MP2/cc-pVQZ) and Drude polarizable MM calculations.

Polarizability	**QM_Pol**	**MM_Pol**	**ΔPol(MM-QM)**
$\mathbf{P}_{xx}$	5.931	5.671	-0.260
$\mathbf{P}_{yy}$	4.552	5.672	1.120
$\mathbf{P}_{zz}$	4.615	5.703	1.088
$\mathbf{P}_{tot}$	5.033	5.682	0.649

QM_Pol and MM_Pol are polarizability values (xx, yy and zz components) obtained from QM and MM calculations, respectively. ΔPol(MM-QM) is the difference between MM and QM values. Polarizability values are represented in $Å_3$.

Table S3: Interaction energies and corresponding distances between water and methylsulfate obtained from QM (MP2/aug-cc-pvDZ) and the Drude polarizable MM calculations. Various entries correspond to different orientation of water molecules placed to represent maximum interaction with different oxygen atoms in methylsulfate.

Complex	QM_IE	MM_IE	ΔIE(MM-QM)	QM_D	MM_D	ΔD(MM-QM)
SULF_O2_WAT	-10.980	-4.539	6.441	1.96	2.14	0.18
SULF_O3_WAT	-12.568	-10.108	2.460	1.83	1.98	0.15
SULF_O4_WAT	-12.805	-9.959	2.846	1.83	1.98	0.15
SULF_O5_WAT	-12.465	-10.051	2.414	1.83	1.98	0.15

QM_IE and MM_IE are interaction energies obtained from QM and MM calculations, respectively; QM_D and MM_D are distances between concerned oxygen atom with closest hydrogen atom of the water molecule in water in QM and MM calculations, respectively. Energy values are in kcal/mol. Distance values are in Å.

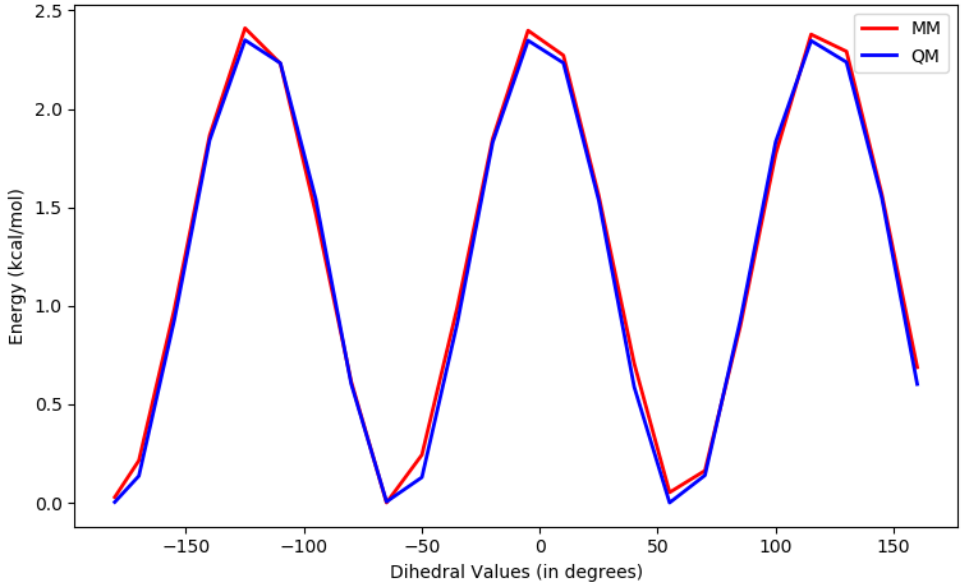


Figure S3: Relaxed potential energy scan of C-O1-S-O2 dihedral of methylsulfate obtained from QM (MP2/aug-cc-pvDZ) and the Drude polarizable MM calculations.

Free energy of hydration

Free energy of hydration was calculated via free energy perturbation method using the staged protocol developed by Deng and Roux.[2] In this procedure, free energy is divided into nonpolar (LJ potential) and electrostatic contributions. The nonpolar contribution is further divided into repulsive and dispersive (attractive) part using Weeks, Chandler and Andersen (WCA) scheme. Following equations show the calculation of free energy of hydration and the obtained energies for methylsulfate. We note that there is

no experimental estimate of the free energy of hydration of methylsulfate available such that the results are included only for informational purposes.

$$\Delta G^{hyd} = \Delta G^{aq} - \Delta G^{vac} + zF\Phi + corr + lrc = -84.05\ kcal/mol \quad \text{Eq. 2}$$

$$\Delta G^{aq} = \Delta G^{aq}_{nonp} + \Delta G^{aq}_{elec} = 14.23 + (-24.08) = -9.86 \quad \text{Eq. 2a}$$

$$\Delta G^{aq}_{nonp} = \Delta G^{aq}_{rep} + \Delta G^{aq}_{dis} = 31.19 + (-16.97) = 14.23;\ \Delta G^{aq}_{elec} = -24.08 \quad \text{Eq. 2b}$$

$$\Delta G^{vac} = \Delta G^{vac}_{nonp} + \Delta G^{vac}_{elec} = 14.05 + 74.62 = 88.67 \quad \text{Eq. 2c}$$

$$\Delta G^{vac}_{nonp} = \Delta G^{vac}_{rep} + \Delta G^{vac}_{dis} = 15.04 + (-0.990943) = 14.0512;\ \Delta G^{vac}_{elec} = 74.62 \quad \text{Eq. 2d}$$

$$z = -1;\ F = 23.06\ ;\ \Phi = -0.540\ ;\ corr = 0.0\ ;\ lrc = -0.12 \quad \text{Eq. 2e}$$

In Eq. 1, ΔG^{hyd} , ΔG^{aq} , $\Delta G^{vac}, z, F, \Phi, corr, lrc$ are free energy of hydration of methylsulfate, free energy of methylsulfate in water, free energy of methylsulfate in gas, total charge, Faraday constant, electrostatic Galvani potential at the liquid vacuum interface, entropy related contributions and long range correction computed using particle mesh Ewald summation, respectively. The terms ΔG^{aq}_{nonp} and ΔG^{aq}_{elec}, are nonpolar (LJ potential) and electrostatic contributions to aqueous free energy. LJ potential is further divided into ΔG^{aq}_{rep} and ΔG^{aq}_{dis} terms representing repulsive and dispersive terms. Similar notation are used for free energy in vacuum as $\Delta G^{vac}_{nonp}, \Delta G^{vac}_{rep}, \Delta G^{vac}_{dis}$ and ΔG^{vac}_{elec}.

Table S4. Drude toppar stream file with topology and parameters for SLES molecule.

```
* DRUDE topology and parameter stream file for ! sulfate
*

!requires toppar_drude_master*.str

read rtf card append
* Topology for Drude nucleic acids
*
38

DEFA FIRS NONE LAST NONE
AUTOGENERATE ANGLES DIHEDRALS DRUDE

RESI LES     -1.000
GROUP            ! CHARGE   CH_PENALTY
ATOM C16    CD32C   -0.371 ALPHA -1.678   THOLE 0.862
ATOM O2     OD30B   -0.279 ALPHA -0.670   THOLE 0.181
ATOM S      SD1A     1.930 ALPHA -0.930   THOLE 1.098
ATOM O3     OD2C2B  -0.850 ALPHA -0.990   THOLE 1.083
ATOM O4     OD2C2B  -0.850 ALPHA -0.990   THOLE 1.083
ATOM O5     OD2C2B  -0.850 ALPHA -0.990   THOLE 1.083
ATOM H32    HDA2A    0.135
ATOM H33    HDA2A    0.135
GROUP
ATOM C13    CD32A   -0.120  ALPHA -1.887  THOLE 0.456
ATOM H26    HDA2A    0.060
ATOM H27    HDA2A    0.060
ATOM C14    CD32A   -0.004  ALPHA -1.696  THOLE 0.918
ATOM H28    HDA2A    0.060
```

```
ATOM H29    HDA2A    0.060
ATOM O1     OD30A    0.000  ALPHA -0.705  THOLE 1.312
ATOM LP1A   LPD     -0.116
ATOM LP1B   LPD     -0.116
ATOM C15    CD32A   -0.004  ALPHA -1.798  THOLE 1.074
ATOM H30    HDA2A    0.060
ATOM H31    HDA2A    0.060
GROUP
ATOM C1     CD33A   -0.177  ALPHA -2.051  THOLE 1.3
ATOM H1     HDA3A    0.059
ATOM H2     HDA3A    0.059
ATOM H3     HDA3A    0.059
GROUP
ATOM C2     CD32A   -0.156  ALPHA -1.660  THOLE 1.3
ATOM H4     HDA2A    0.078
ATOM H5     HDA2A    0.078
GROUP
ATOM C3     CD32A   -0.156  ALPHA -1.660  THOLE 1.3
ATOM H6     HDA2A    0.078
ATOM H7     HDA2A    0.078
GROUP
ATOM C4     CD32A   -0.156  ALPHA -1.660  THOLE 1.3
ATOM H8     HDA2A    0.078
ATOM H9     HDA2A    0.078
GROUP
ATOM C5     CD32A   -0.156  ALPHA -1.660  THOLE 1.3
ATOM H10    HDA2A    0.078
ATOM H11    HDA2A    0.078
GROUP
ATOM C6     CD32A   -0.156  ALPHA -1.660  THOLE 1.3
ATOM H12    HDA2A    0.078
ATOM H13    HDA2A    0.078
GROUP
ATOM C7     CD32A   -0.156  ALPHA -1.660  THOLE 1.3
ATOM H14    HDA2A    0.078
ATOM H15    HDA2A    0.078
GROUP
ATOM C8     CD32A   -0.156  ALPHA -1.660  THOLE 1.3
ATOM H16    HDA2A    0.078
ATOM H17    HDA2A    0.078
GROUP
ATOM C9     CD32A   -0.156  ALPHA -1.660  THOLE 1.3
ATOM H18    HDA2A    0.078
ATOM H19    HDA2A    0.078
GROUP
ATOM C10    CD32A   -0.156  ALPHA -1.660  THOLE 1.3
ATOM H20    HDA2A    0.078
ATOM H21    HDA2A    0.078
GROUP
ATOM C11    CD32A   -0.156  ALPHA -1.660  THOLE 1.3
ATOM H22    HDA2A    0.078
ATOM H23    HDA2A    0.078
GROUP
ATOM C12    CD32A   -0.156  ALPHA -1.660  THOLE 1.3
ATOM H24    HDA2A    0.078
ATOM H25    HDA2A    0.078

BOND  C1   C2     C2   C3     C3   C4     C4   C5
BOND  C5   C6     C6   C7     C7   C8     C8   C9   C9   C10
BOND  C10  C11    C11  C12    C12  C13
BOND  C1   H1     C1   H2     C1   H3     C2   H4   C2   H5
BOND  C3   H6     C3   H7     C4   H8     C4   H9
BOND  C5   H10    C5   H11    C6   H12    C6   H13
BOND  C7   H14    C7   H15    C8   H16    C8   H17
BOND  C9   H18    C9   H19    C10  H20   C10   H21  C11  H22
BOND  C11  H23    C12  H24    C12  H25
BOND C13   C14    C14  O1     O1   C15   C13   H27  C13  H26
BOND C14   H29    C14  H28    C15  H31   C15   H30  C16  C15
BOND O1    LP1A   O1   LP1B   O4    S    O3    S    H32   C16
BOND S     O2     S     O5    H33   C16  C16   O2

LONEPAIR bisector LP1A O1 C15 C14 distance 0.35 angle 110.0 dihe  90.0
LONEPAIR bisector LP1B O1 C15 C14 distance 0.35 angle 110.0 dihe 270.0
END

read param card append
* Parameters generated by analogy by
* CHARMM Drude Force Field program version 2.2.0
*

! Penalties lower than 10 indicate the analogy is fair; penalties between 10
```

```
! and 50 mean some basic validation is recommended; penalties higher than
! 50 indicate poor analogy and mandate extensive validation/optimization.

BONDS
OD2C2B    SD1A      525.00      1.493  ! DMP, cmb, 06/09
OD30B     SD1A      240.00      1.701  ! DMP, cmb, 06/09
SD1A      LPD         0.00      0.000  ! DMP
CD32A     OD30A     360.00      1.415
CD32C     OD30B     335.00      1.420

ANGLES
SD1A      OD30B     CD32C      75.00     98.38     ! DMP, cmb, 06/09
OD2C2B    SD1A      OD2C2B     60.00     115.35    ! DMP, csd, EH/IV 2007
OD2C2B    SD1A      OD30B      90.00     98.44     ! DMP, csd, EH/IV 2007**better molvib fit
OD30B     CD32C     HDA2A      60.00    109.50
OD30B     CD32C     CD32A      75.70    110.10
CD32C     CD32A     OD30A      75.70    110.10

DIHEDRALS
CD32C     OD30B     SD1A      OD2C2B      5.056    2   180.00
SD1A      OD30B     CD32C     HDA2A       0.000    3     0.00 ! DMP, cmb, 06/09
OD30B     CD32C     CD32A     OD30A       0.3681   1     0.00
OD30B     CD32C     CD32A     OD30A       1.2036   2     0.00
OD30B     CD32C     CD32A     OD30A       0.1171   3     0.00
OD30B     CD32C     CD32A     OD30A       0.1612   4     0.00
OD30B     CD32C     CD32A     OD30A       0.0340   5     0.00
OD30B     CD32C     CD32A     OD30A       0.0225   6     0.00
OD30B     CD32C     CD32A     HDA2A       0.190    3     0.00
HDA2A     CD32C     CD32A     OD30A       0.190    3     0.00
CD32C     CD32A     OD30A     CD32A       0.570    1     0.00
CD32C     CD32A     OD30A     CD32A       0.290    2     0.00
CD32C     CD32A     OD30A     CD32A       0.430    3     0.00
CD32A     CD32C     OD30B     SD1A        0.203    1    180.00
CD32A     CD32C     OD30B     SD1A        0.182    2    0.00
CD32A     CD32C     OD30B     SD1A        0.123    3    180.00
CD32A     CD32C     OD30B     SD1A        0.089    4    0.00
CD32A     CD32C     OD30B     SD1A        0.143    5    180.00
CD32A     CD32C     OD30B     SD1A        0.093    6    180.00

IMPROPERS

NONBONDED nbxmod  5 atom vatom cdiel vdistance switch vswitch -
cutnb 16.0 ctofnb 12.0 ctonnb 10.0 eps 1.0 e14fac 1.0 wmin 1.5

SD1A      0.0   -0.2700     1.9000 ! DMP, lipids

END
RETURN
```

Table S5. Drude toppar stream file with topology and parameters for CA molecule.

```
* DRUDE topology and parameter stream file
*

!requires toppar_drude_master*.str

!ioformat extended


read rtf card append
* Topology for Drude lipids
*
38

DEFA FIRS NONE LAST NONE
AUTOGENERATE ANGLES DIHEDRALS DRUDE  !note use of DRUD

RESI CAC            0.000 ! param penalty=   0.600 ; charge penalty=   1.952
GROUP            ! CHARGE   CH_PENALTY
ATOM C8     CD32C  -0.208  ALPHA -2.114 THOLE 0.750 ! -0.208
ATOM H15    HDA2A   0.092
ATOM H16    HDA2A   0.092
ATOM C10    CD2O3A  0.858  ALPHA -1.207 THOLE 0.708
ATOM O1     OD2C3A  0.000  ALPHA -0.922 THOLE 1.539
ATOM LPP1   LPD    -0.319
ATOM LPP2   LPD    -0.319
ATOM O2     OD30D   0.000  ALPHA -1.280 THOLE 1.124
ATOM LPP3   LPD    -0.285
ATOM LPP4   LPD    -0.285
```

```
ATOM H20    HDP1A   0.374
GROUP
ATOM C9     CD33A  -0.177  ALPHA -2.051  THOLE 1.3
ATOM H17    HDA3A   0.059
ATOM H18    HDA3A   0.059
ATOM H19    HDA3A   0.059
GROUP
ATOM C2     CD32A  -0.156  ALPHA -1.660  THOLE 1.3
ATOM H3     HDA2A   0.078
ATOM H4     HDA2A   0.078
GROUP
ATOM C3     CD32A  -0.156  ALPHA -1.660  THOLE 1.3
ATOM H5     HDA2A   0.078
ATOM H6     HDA2A   0.078
GROUP
ATOM C4     CD32A  -0.156  ALPHA -1.660  THOLE 1.3
ATOM H7     HDA2A   0.078
ATOM H8     HDA2A   0.078
GROUP
ATOM C5     CD32A  -0.156  ALPHA -1.660  THOLE 1.3
ATOM H9     HDA2A   0.078
ATOM H10    HDA2A   0.078
GROUP
ATOM C6     CD32A  -0.156  ALPHA -1.660  THOLE 1.3
ATOM H11    HDA2A   0.078
ATOM H12    HDA2A   0.078
GROUP
ATOM C7     CD32A  -0.156  ALPHA -1.660  THOLE 1.3
ATOM H13    HDA2A   0.078
ATOM H14    HDA2A   0.078
GROUP
ATOM C1     CD32A  -0.156  ALPHA -1.660  THOLE 1.3
ATOM H1     HDA2A   0.078
ATOM H2     HDA2A   0.078

BOND C1   C2   C2   C4  C3   C5  C4   C6  C5   C7  C6   C8  C7   C9
BOND C8   C10  C1   C3  C10  O1  C10  O2  C1   H1  C1   H2  C2   H3
BOND C2   H4   C3   H5  C3   H6  C4   H7  C4   H8  C5   H9  C5   H10
BOND C6   H11  C6   H12 C7   H13 C7   H14 C8   H15 C8   H16 C9   H17
BOND C9   H18  C9   H19 O2   H20 O1  LPP1 O1  LPP2 O2  LPP3 O2  LPP4
IMPR C10    C8     O1     O2

!standard carbonyl
LONEPAIR relative LPP1 O1 C10 O2 distance 0.30 angle 91.0 dihe   0.0
LONEPAIR relative LPP2 O1 C10 O2 distance 0.30 angle 91.0 dihe 180.0
ANISOTROPY O1 C10 LPP1 LPP2  A11 0.6968  A22 1.2194

!from MeOH
LONEPAIR relative LPP3 O2 C10 H20 distance 0.35 angle 110.9 dihe  91.0
LONEPAIR relative LPP4 O2 C10 H20 distance 0.35 angle 110.9 dihe 269.0
ANISOTROPY O2 C10 LPP3 LPP4  A11 0.8108  A22 1.2162

END
```

Table S6. Drude toppar stream file with topology and parameters for Caprate ion molecule.

```
* DRUDE topology and parameter stream file
*

!requires toppar_drude_master*.str

!ioformat extended

read rtf card append
* Topology for Drude lipids
*
38

DEFA FIRS NONE LAST NONE
AUTOGENERATE ANGLES DIHEDRALS DRUDE  !note use of DRUD

RESI CAP           -1.000 !
GROUP
ATOM C8     CD32C  -0.190  ALPHA -2.528  THOLE 1.414
ATOM H15    HDA2A   0.004
ATOM H16    HDA2A   0.004
ATOM C10    CD2O2A  0.708  ALPHA -1.016  THOLE 0.899
ATOM O1     OD2C2A  0.003  ALPHA -0.699  THOLE 2.399
ATOM LP1A   LPD    -0.383
ATOM LP1B   LPD    -0.383
```

```
ATOM O2     OD2C2A  0.003  ALPHA -0.699  THOLE 2.399
ATOM LP2A   LPD    -0.383
ATOM LP2B   LPD    -0.383
GROUP
ATOM C9     CD33A  -0.177  ALPHA -2.051  THOLE 1.300
ATOM H17    HDA3A   0.059
ATOM H18    HDA3A   0.059
ATOM H19    HDA3A   0.059
GROUP
ATOM C2     CD32A  -0.156  ALPHA -1.660  THOLE 1.300
ATOM H4     HDA2A   0.078
ATOM H3     HDA2A   0.078
GROUP
ATOM C3     CD32A  -0.156  ALPHA -1.660  THOLE 1.300
ATOM H6     HDA2A   0.078
ATOM H5     HDA2A   0.078
GROUP
ATOM C4     CD32A  -0.156  ALPHA -1.660  THOLE 1.300
ATOM H8     HDA2A   0.078
ATOM H7     HDA2A   0.078
GROUP
ATOM C5     CD32A  -0.156  ALPHA -1.660  THOLE 1.300
ATOM H10    HDA2A   0.078
ATOM H9     HDA2A   0.078
GROUP
ATOM C6     CD32A  -0.156  ALPHA -1.660  THOLE 1.300
ATOM H12    HDA2A   0.078
ATOM H11    HDA2A   0.078
GROUP
ATOM C7     CD32A  -0.156  ALPHA -1.660  THOLE 1.300
ATOM H14    HDA2A   0.078
ATOM H13    HDA2A   0.078
GROUP
ATOM C1     CD32A  -0.156  ALPHA -1.660  THOLE 1.300
ATOM H1     HDA2A   0.078
ATOM H2     HDA2A   0.078

BOND C1   C2   C2   C4   C3   C5   C4   C6   C5   C7
BOND C6   C8   C7   C9   C8   C10  C1   C3
BOND C10  O1   C10  O2
BOND C1   H1   C1   H2   C2   H3   C2   H4   C3   H5
BOND C3   H6   C4   H7   C4   H8   C5   H9   C5   H10
BOND C6   H11  C6   H12  C7   H13  C7   H14  C8   H15
BOND C8   H16  C9   H17  C9   H18  C9   H19
BOND O1  LP1A    O1  LP1B    O2  LP2A    O2  LP2B
IMPR C10    C8     O1     O2

LONEPAIR relative LP1A O1 C10 C8  distance 0.35 angle 110.0 dihe   0.0
LONEPAIR relative LP1B O1 C10 C8  distance 0.35 angle 110.0 dihe 180.0
LONEPAIR relative LP2A O2 C10 C8  distance 0.35 angle 110.0 dihe   0.0
LONEPAIR relative LP2B O2 C10 C8  distance 0.35 angle 110.0 dihe 180.0
ANISOTROPY O1 C10 LP1A LP1B A11 0.7229 A22 1.265
ANISOTROPY O2 C10 LP2A LP2B A11 0.7229 A22 1.265

END

read param card append
* Parameters generated by analogy by
* CHARMM Drude Force Field program version 2.2.0
*
ANGLES

CD2O2A   CD32C    CD32A     30.60    120.70                  ! CTER, PEML, Glu


DIHEDRALS
CD2O2A   CD32C    CD32A    CD32A       0.200    3     0.00 !
CD2O2A   CD32C    CD32A    HDA2A       0.200    3     0.00 !
OD2C2A   CD2O2A   CD32C    CD32A       0.200    3     0.00 !

IMPROPERS

CD2O2A   CD32C    OD2C2A   OD2C2A      71.000   0     0.00 ! CTER, Gly


END
RETURN
```

6 References

(1) Tzocheva, S. S.; Kralchevsky, P. A.; Danov, K. D.; Georgieva, G. S.; Post, A. J.; Ananthapadmanabhan, K. P. Solubility Limits and Phase Diagrams for Fatty Acids in Anionic (SLES) and Zwitterionic (CAPB) Micellar Surfactant Solutions. *J. Colloid Interface Sci.* **2012**, *369* (1), 274–286. https://doi.org/https://doi.org/10.1016/j.jcis.2011.12.036.

(2) Li, H.; Chowdhary, J.; Huang, L.; He, X.; MacKerell, A. D.; Roux, B. Drude Polarizable Force Field for Molecular Dynamics Simulations of Saturated and Unsaturated Zwitterionic Lipids. *J. Chem. Theory Comput.* **2017**, *13* (9), 4535–4552. https://doi.org/10.1021/acs.jctc.7b00262.

(3) Frisch, Ae.; Frisch, M. J. *Gaussian 03 Pocket Reference*; Gaussian, Incorporated, 2003.

(4) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. Efficient Diffuse Function-Augmented Basis Sets for Anion Calculations. III. The 3-21+G Basis Set for First-Row Elements, Li–F. *J. Comput. Chem.* **1983**, *4* (3), 294–301. https://doi.org/10.1002/jcc.540040303.

(5) Woon, D. E.; Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. III. The Atoms Aluminum through Argon. *J. Chem. Phys.* **1993**, *98* (2), 1358–1371. https://doi.org/10.1063/1.464303.

(6) Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90* (2), 1007–1023. https://doi.org/10.1063/1.456153.

(7) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem.* **1983**, *4* (2), 187–217. https://doi.org/10.1002/jcc.540040211.