Derivation of micelle size-dependent free energies of aggregation for octyl phosphocholine from molecular dynamics simulation

Xiaokun Zhang, Jorge G. Arce, and James T. Kindt
Department of Chemistry, Emory University
Atlanta, GA 30322
Accepted for publication in *Fluid Phase Equilibria*

Abstract
Molecular dynamics (MD) simulations of the zwitterionic surfactant octyl phosphocholine (OPC) in water have been performed with two force fields, both using the TraPPE united-atom alkane parameters for tailgroup-tailgroup interactions but with tailgroup-water interactions adjusted to reproduce hydration free energies for two different water models (SPC and TIP4P-2005). Micelle size distributions from a number of trajectories were analyzed using the PEACH (Partition-Enabled Analysis of Cluster Histograms) method to yield free energies of aggregation for premicelles and micelles over the full range from 2 to over 40 molecules. The dependence of free energy on aggregation number was consistent with the functional form derived from the "quasi-droplet" model of micellization. Free energies of aggregation were used to calculate concentration-dependent micelle size distributions, from which critical micelle concentrations (cmc's) could be inferred. The cmc values derived from PEACH free energies for the SPC and TIP4P-2005 models were 40% and 27% higher respectively than the experimentally reported value of 114 mM. In systems with enough monomers to form multiple micelles, the micelles tended to stick together, complicating the analysis. Micelle size distributions were used to weight small-angle X-ray scattering (SAXS) form factors derived from micelle structures, generatinga composite form factor reflecting the polydispersity of aggregate sizes with contributions in the range from 20-40 monomers. The composite scattering profile was quite similar to that of the mass-weighted average micelle size, suggesting that SAXS profiles for small micelles can be adequately fit using a single typical micelle even in the presence of significant polydispersity.

1. Introduction

Surfactant micelles, disordered aggregates of amphiphiles in aqueous solution, are typically formed above a critical micelle concentration (cmc) as the hydrophobic attraction of the surfactant tails overcomes the entropy of mixing individual surfactant monomers (unimers) and small clusters (premicelles) throughout the solution. The structural disorder, polydispersity, and responsiveness to conditions present challenges to the unambiguous determination of the degree of aggregation and structure of surfactant micelles through experiment alone. Molecular simulation has long been a source of insight into these features.[1-3] Alkyl phosphocholines have been a focus of particular attention[4-6] in part because they are zwitterionic (and so bypass complications associated with counterions) and in part because they share their headgroup structures with the well-studied bilayer-forming phosphatidylcholine glycero-phospholipids.

A number of ingredients are required to achieve a level of confidence in the structural and thermodynamic properties displayed through simulation models, which by necessity rely on certain approximations. The use of a sufficiently realistic force field (or at least

one whose limitations are well understood and characterized) is one such ingredient. The ability to extrapolate from simulation data obtained on a finite size system to the limit of a macroscopic number of molecules and clusters is another, and the ability to make connections between simulated structures and experimental observables is a third. This report will describe a simulation study on micellization of a simple surfactant, octyl phosphocholine (OPC), in which steps are taken to address all three ingredients. The force fields used for surfactant tails are based on the TraPPE-UA model,[7] and so have been parameterized to fit alkane-alkane thermodynamic properties, but with modifications to interactions with water made to yield accurate free energies of hydration.[8] The PEACH statistical analysis method[9, 10] allows the size-dependent cluster free energy to be determined over a range of micelle sizes from a series of simulations with relatively few surfactants; the resulting free energy function allows the cmc to be calculated in a way that facilitates comparison with experiment. Finally, the effect of incorporating polydispersity into predictions of the x-ray scattering of OPC (where polydispersity is relatively high) will be investigated.

2. Methods

2.1 General elements of force fields.

Two united-atom (UA) force fields, in which similar models for octyl phosphocholine were matched with different water potentials, were used and compared. In both, bonded interactions of the alkyl tails, as well as all pairwise interactions of CH_x groups with other CH_x groups (including headgroup methyl and methylenes) used parameters from the TraPPE-UA force field.[11] Parameters for the OPC headgroup (bond lengths, bending and torsional potentials, partial charges, and Lennard-Jones parameters) were taken from the Gromos lipid forcefield of Poger *et al*., [12] derived from the Gromos G53A6 force field[13] but including an adjustment to the interaction between the choline methyl group with non-ester phosphate oxygens that was made to bring lipid bilayer properties in line with experiment.

For simulations using TIP4P-2005 water, Lennard-Jones parameters for interactions between water oxygen and all CH_x groups were taken from the HHAlkane model;[8] this force field will be referred to as "HHA". For simulations using SPC water (denoted "Alk-SPC") these parameters were calculated as described below.

2.2 Alkane-SPC water interactions.

The TraPPE-UA model was reparameterized to reproduce experimental hydration free energies for n-alkanes in SPC water, by adjusting the nonbonded alkane site-water interactions (CH_x-OW) following the concept of Ashbaugh et al.[8] The excess hydration free energy is defined as the free energy to transfer the solute molecule from an ideal gas to an ideal-dilute aqueous solution at the same concentration. This quantity was calculated from the g_bar utility in Gromacs 5.0[14] which is based on the BAR (Bennett Acceptance Ratio) method[15]. In the implementation of BAR method, the system of alkane with water was sampled for multiple coupling states with a parameter λ to adjust the strength of the Van der Waals interactions between the alkane and the solvent. For our calculations, 20 points of λ were chosen from 0 to 1 with an increment of 0.05. Free energy differences for each two adjacent points i and i+1 are calculated based on $\langle U_{i+1}-U_i \rangle_i$ and $\langle U_i-U_{i+1} \rangle_{i+1}$.

The simulation for each λ point was set up with one alkane in a box of ~420 SPC water using Gromacs 5.0[14]. The pressure is maintained at 1 bar and a compressibility of 4.5×10^{-5} by Parrinello-Rahman barostat[16] with $\tau_P = 2\ ps$. The stochastic dynamics

integrator with a 2 fs time step and coupling time $\tau_T = 2\,ps$ was used for integration of equations of motion and to maintain a constant temperature of 300 K. The Verlet[17] cutoff-scheme was applied for short-range non-bonded interactions with a cutoff of 9 Å. The long-range dispersion correction was applied for energy and pressure. Particle mesh Ewald summation[18] was used to account for long-range electrostatics with a real space cutoff of 9 Å. After a 100 ps pre-equilibration, a 1 ns simulation was used for each λ-point in the hydration free energy calculation.

To optimize the nonbonded CH_x-OW interactions for accurate interaction free energies of alkyl tail in SPC water, the LJ well depth $\varepsilon_{CHx\text{-}OW}$ and LJ diameter $\sigma_{CHx\text{-}OW}$ are adjusted to fit the hydration free energies of n-alkanes. The initial cross interactions were determined by Lorentz-Berthelot combining rules

$$\sigma_{ij} = \frac{1}{2}\left(\sigma_i + \sigma_j\right), \epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j} \quad (1)$$

The strength of the alkane-water attraction was varied by scaling $\varepsilon_{CHx\text{-}OW}$ while adjusting $\sigma_{CHx\text{-}OW}$ to maintain constant thermal radius[19] r_{iw}^{therm}. Given as below, r_{iw}^{therm} is fixed during the fitting to maintain the solute excluded volume.

$$r_{iw}^{therm} = \sigma_{iw}\left[\frac{2}{1+(kT/\varepsilon_{iw})^{1/2}}\right]^{1/6} \quad (2)$$

2.3 Simulations of OPC and PEACH analysis

Simulations of OPC in SPC and TIP4P-2005 water were performed with the Gromacs 5.0[14] software package. For PEACH calculations, we set up a series of simulations of different concentrations (20-55 OPCs with an increment of 5 OPC's, each paired with two sizes of solvent box). This set is chosen to obtain sufficient samplings across the full range of cluster sizes and allowing K_i to be obtained for cluster sizes $2 < i <$ 40-50. The choices of N and V used in PEACH analysis are given in Table 1.

The temperature was maintained at 300K by velocity rescaling thermostat[20] with $\tau_T = 2\,ps$. The pressure is maintained at 1 bar and a compressibility of 4.5×10^{-5} bar^{-1} by Berendsen barostat[21] with $\tau_P = 2\,ps$. The Gromacs default (leap-frog) integrator with a 2 fs time step was used for integration of equations of motion. The Verlet[17] cutoff-scheme was applied for short-range non-bonded interactions with a cutoff of 1.4 nm. The long-range dispersion correction was applied for energy and pressure. Particle-mesh Ewald summation[18] was used to account for Coulomb interactions with a real space cutoff of 1.4 nm.

Microsecond simulations (1.0 μs for HHA, 1.2 μs for Alk-SPC) with a 20 ns pre-equilibration were performed for each trial and used for PEACH analysis. (An additional large system simulation was performed with 250 OPC solvated by 30330 SPC waters and run for 200 ns.) The trajectory was analyzed by a modified version of Gromacs utility g_clustsize to determine the distribution of cluster sizes. Two OPC molecules are considered to be neighbors if the distance between any of their alkyl tail sites is lower than a cut-off distance r_{cut}. Values of r_{cut} at intervals of 0.05 nm between 0.45 and 0.65 nm were used, and results derived from these are compared as described below. The chains related by neighboring interactions are considered to be of the same cluster. Histograms of $\langle n_i \rangle$, representing the mean numbers of clusters of aggregation number i averaged over each trajectory, were used as input for PEACH (Partition-Enabled Analysis of Cluster Histograms) to find a globally optimized set $\{K_i\}$ of equilibrium association constants for each cluster size observed. As described in previous

publications,[9, 10] the method involves assigning an association constant to each possible cluster size *i*, generating a model cluster size distribution for each simulation based on its total *N* and *V*, and adjusting it iteratively to find the best set K_i to reproduce the cluster size distributions from simulations. The core algorithm uses the following equation for calculation of cluster size distribution from the current model $\{K_i\}$:[22]

$$< n_i >= Q(N,V)^{-1} \sum_{j=1}^{p(N)} n_{ij} Q(\{n_i\}_j, V) = Q(N,V)^{-1} \sum_{j=1}^{p(N)} \left(n_{ij} \prod_{i=1}^{N} \frac{(K_i V c^o)^{n_{ij}}}{n_{ij!}} \right) \quad (3)$$

in which $< n_i >$ is the ensemble average number of *i*-mer clusters present, p(*N*) represents the number of possible compositions of clusters that add up to *N* (called "partitions of *N*" in number theory), $Q(\{n_i\}_j, V)$ is the canonical partition function for a possible composition $\{n_i\}_j$, and $Q(N,V)$ is the sum of these partition functions over all compositions. This calculation can be simplified[10], avoiding the need to generate all partitions of *N*, using a generating function approach. Briefly, at fixed *V*, *Q*(*N*) is obtained as the coefficient of λ^N calculated from the polynomial expansion of the grand partition function $\Xi(\lambda)$ (where λ is now the thermodynamic activity, unlike in section *2.1*):

$$\Xi(\lambda) = \exp\left(\sum_{i=1}^{\infty} q_i \lambda^i\right) = \sum_{N=0}^{\infty} \lambda^N Q(N) \quad (4)$$

Evaluation of *Q*(*N*) can be achieved by differentiation of $\Xi(\lambda)$:

$$Q(N) = \frac{1}{N!} \left[\frac{\partial^N \exp\left(\sum_{i=1}^{\infty} q_i \lambda^i\right)}{\partial \lambda^N} \right]_{\lambda=0}$$

The single-cluster partition function q_i is related to the association constant K_i, the system volume *V*, the standard-state concentration *c*°, and the standard unimer partition function $q°_1$ as:

$$q_i = K_i c° V (q°_1)^i \quad (6)$$

The mean number of clusters of a given size can then be calculated as[10]:

$$\langle n_i \rangle = \frac{q_i Q(N-i)}{Q(N)} \quad (7)$$

The Gibbs free energy of formation of an *i*-mer cluster from *i* unimers, under conditions where both unimer and *i*-mer concentrations are given by *c*, can be expressed as:

$$\frac{\Delta G_i}{kT} = -lnK_i - (i-1) * ln\frac{c}{c^0} \quad (8)$$

Table 1. Details of simulations used in PEACH fitting

Simulation details for OPC cluster aggregation in SPC water (Alk-SPC)									
	Set One			Set Two			Set Three		
N_{OPC}	N_{water}	Box Size/nm^3	c_{tot}/mM	N_{water}	Box Size/nm^3	c_{tot}/mM	N_{water}	Box Size/nm^3	c_{tot}/mM
20	3812	123.6	282	5623	179.1	191	7000	221	153
25		125.4	352		180.9	238		223	191
30		127.7	422		182.6	286		225	230
35		129.4	493		184.5	334		227	268
40		131.4	563		186.7	381		229	306
45		133.3	633		188.6	429		231	345
50		135.1	704		190.4	477		233	383
55		136.9	774		192.3	524		234	421

Simulation details for OPC cluster aggregation in TIP4P/2005 water (HHA)						
	Set One			Set Two		
N_{OPC}	N_{water}	Box Size/nm^3	c_{tot}/mM	N_{water}	Box Size/nm^3	c_{tot}/mM
15	3812	120.3	220	5025	156.8	165
20		122.2	293	5053	159.5	219
25		124.1	366	5035	160.8	275
30		126.3	439	5030	162.6	330
35		128.2	512	5047	165.3	384
40		130.4	585	6021	196.5	367
45	4549	154.4	547	6017	198.3	414
50	-	-	-	6024	200.6	459
55	-	-	-	6768	234.7	450

2.4. *WAXSiS analysis for polydisperse micelles*

To model the small-angle X-ray scattering (SAXS) profile for micelles of a specific aggregation number, structures of single OPC micelles were submitted to the WAXSiS server.[23] The server re-solvates the structure and calculates averages over a short MD trajectory to average over small fluctuations of lipid and solvent configurations. Larger shape fluctuations were observed over the trajectory, thus scattering profiles are averaged across 5~10 configurations of a single micelle observed throughout a trajectory to represent the profile for that size.

PEACH-derived association constants or free energies can be used to calculate the full distribution of concentrations for micelles of each possible aggregation number. Neglecting inter-micelle interactions, the expected form factor of the mixture should mirror the number-weighted average of the individual form factors. To avoid having to generate scattering profiles for 40 distinct micelle sizes, we instead chose representative sizes at intervals of 5. The r_{cut} =0.45 nm cluster criterion was used, based on the reasoning that more compact clusters would contribute more strongly to scattering.

Because the scattering intensity of a spherical object scales approximately with the square of its volume[24], clusters of size similar to these representative clusters are assumed to scatter as the nearest representative size micelle, scaled by the square of the ratio of their aggregation numbers;

$$I = I_{20}\sum_{i=14}^{22}(\frac{i}{20})^2 * c_i + I_{25}\sum_{i=23}^{27}(\frac{i}{25})^2 * c_i + I_{30}\sum_{i=28}^{32}(\frac{i}{30})^2 * c_i + I_{35}\sum_{i=33}^{37}(\frac{i}{35})^2 * c_i + I_{40}\sum_{i=38}^{42}(\frac{i}{40})^2 * c_i + I_{45}\sum_{i=43}^{47}(\frac{i}{45})^2 * c_i + I_{50}\sum_{i=48}^{52}(\frac{i}{50})^2 * c_i \quad (9)$$

3. Results and Discussion

3.1 Force field parameter selection

A key element of a force field for accurate simulations of micelle formation is to reproduce the interaction free energies of alkyl surfactant tails with each other and with water. As shown in previous simulation studies of micellization[25], a number of force fields yield cmc estimates below experiment by a factor of 2 or worse. One of these, the TraPPE-UA model[7] was designed to provide an accurate description of alkane-alkane interactions, reproducing the thermodynamics of alkane liquid-vapor equilibria. The likely weakness therefore lies in hydration free energies that are consistently too high – presumably due to unaccounted polarization effects between water and alkanes. To address this, Ashbaugh *et al.*[8] used TraPPE as a basis for their HH-Alkane model, where the CH_x-OW Lennard-Jones interactions were adjusted slightly to improve agreement of alkane hydration free energies upon solvation with TIP4P/2005 water.[26] Here we use the HH-Alkane parameters with TIP4P/2005 in one of our two force fields for OPC. For future use in studies of the important area of micelle-protein interactions,[27] we sought to develop a similar set of CH_x-OW parameters for alkane interactions with the SPC water model to aid in compatibility with the Gromos G54a7 UA protein force field,[28] which was designed for use with SPC. To do so, we again started with the TraPPE parameters and followed an abbreviated version of the steps taken for the HH-Alkane force field to optimize hydration free energies of small alkanes.

The initial parameters for the fitting were determined using Lorentz-Berthelot combining rules based on the LJ parameters for TraPPE alkane sites and SPC water. It is shown in Table 2 that the initial set of parameters used in ethane predicts higher hydration free energies than experiments i.e. stronger repulsions between SPC water and alkanes; thus we increase the LJ well depth $\varepsilon_{CHx\text{-}OW}$ iteratively in seek of a best set of parameters to fit the hydration free energies. The LJ diameter $\sigma_{CHx\text{-}OW}$ was adjusted at the same time to keep the thermal radius fixed. CH_3 group parameters were first adjusted to fit to the experimental hydration free energy of ethane at 300 K. Keeping those parameters fixed, CH_2 parameters were adjusted to give agreement with the experimental hydration free energy of *n*-propane at 300 K. To check the transferability of the parameters, the free enegy of hydration of *n*-butane was then evaluated using the CH_2 and CH_3 parameters. CH-OW parameters were also adjusted according to experimental hydration free energy of isobutane, while the neopentane hydration free energy (which presumably is most sensitive to the CH_3-OW parameters as the quaternary carbon is shielded from contact with water) was close to experiment without further adjustment of the C-OW parameters. The results are given in Table 2. As for the HH-Alkane force field, small adjustments in

attractive strength (by 1.6% for CH_3, 3.5% for CH_2) were sufficient to bring simulated hydration free energies in line with experiment.

Table 2. LJ Parameters and test on hydration free energy of alkanes					
Cross LJ parameters before and after optimization					
	Initial parameters		New (Alk-SPC) parameters		
	σ_{ij} /nm	ε_{ij} /K	σ_{ij} /nm	ε_{ij} /K	r_{iw}^{therm} /nm
CH3-OW	0.3458	87.5451	0.3455	88.9805	0.32596
CH2-OW	0.3558	59.9790	0.3551	62.0652	0.32837
CH-OW	0.38985	30.8780	0.38981	30.9283	0.34565
C-OW	0.46175	10.2410	-	-	0.38026

Test on hydration free energy for new parameters			
	Hydration free energy (kJ/mol)		
	Initial Parameter	New (Alk-SPC) Parameter	Experiment
Ethane	8.30(0.12)	7.75(0.17)	7.82
Propane	8.61(0.26)	8.43(0.26)	8.37
n-Butane	8.80(0.19)		8.91
Isobutane	9.24(0.09)	9.47(0.24)	9.58
Neopentane	10.69(0.33)	-	10.71

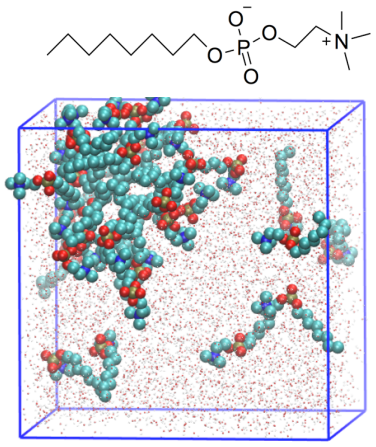


Figure 1. Structure of OPC and snapshot of 35 OPC in 3812 SPC water, including a 28-mer micelle. Bead colors represent corresponding elements; Cyan-carbon, red-oxygen, blue-nitrogen, yellow-phosphorus.

3.2. Simulation and PEACH analysis of OPC

Simulations over a range of system sizes from 20 to 55 OPC (the structure shown in Fig.1) were performed at two concentrations for both force fields. A snapshot of a typical system is shown in Fig.1. For systems with at least 30 OPC, one large micelle is present in the simulation box for most of the trajectory in equilibrium with several monomers or small clusters. The rare event of one large micelle splitting into two occurs a few times in the simulation timescale. Micelle size distributions evaluated with a cluster cut-off criterion r_{cut} = 0.45 nm are shown in Fig. 2. For each force field, free energies of aggregation for all cluster sizes *j* were determined iteratively to optimize the global fit to the full set of size distributions using the PEACH algorithm. Agreement between the model fits (curves, Fig. 2) and simulation data (symbols, Fig. 2) is reasonably good. The fitting is poorer in the range from cluster size 10 to 20 due to less sampling. Distributions and fits derived using other cut-offs are given in Supplemental Information Fig. S1.

The HHA data sets appear noisier than the ALK-SPC data primarily because of poorer statistical sampling associated with markedly slower fluctuations in micelle size, as shown in Supplemental Information Fig. S2. Although a quantitative consideration of the OPC exchange kinetics is outside of the scope of this paper, slower size fluctuations under the same conditions reflect a slower rate of monomer exchange between the micelle and the solvent. One possible contribution is that the viscosity of SPC water is known to be lower than that of TIP4P-2005;[29] this is consistent with our observation that the diffusion constant of dilute OPC using Alk-SPC interactions in SPC water is about 60% higher than using HHA in TIP4P-2005 water at 300 K (results not shown). This difference in diffusion rate is not, however, sufficient to account for the large apparent difference in monomer exchange rates. Given the fairly slight difference in aggregation free energies, it appears that the two solvent models produce significant differences in the activation barrier for OPC exchange.

Several 3-parameter models of the form:

$$\frac{\Delta G(i)}{kT} = Ai^a - Bi^b + Ci^c \qquad (10)$$

have been proposed to describe the dependence of micelle free energy on aggregation number. [30-32] In the droplet model of Shchekin and co-workers[29] and the model of Maibaum and Chandler[30], both derived from classical nucleation theory, the first term arises from the surface free energy and so scales as the surface area of a compact object with *a*=2/3. The second term is related to the free energy of transfer of the surfactant from solvent to a bulk-like hydrophobic environment and is therefore linear, *b*=1 and $B=\Delta\mu/kT$. For the the last term, which accounts for the structural limitations to the size of the micelle, considerations of dipole-dipole interactions[30] or entropic costs associated with arranging headgroups at the micelle surface[31] have motivated choices of *c* = 4/3 and *c* = 5/3 respectively. A third “quasi-droplet” model also proposed by Shchekin et al.[32] is derived from a model for the micelle that includes partially hydrated segments of chains protruding from a compact hydrophobic core; the evolving balance between core and hydrated segments leads to combinations of terms with *a*=1, *b*=3/2, and *c*=2. In testing these models for their ability to fit the PEACH-derived free energies from simulation, we modify them slightly to ensure that $\Delta G(1)=0$ as follows:

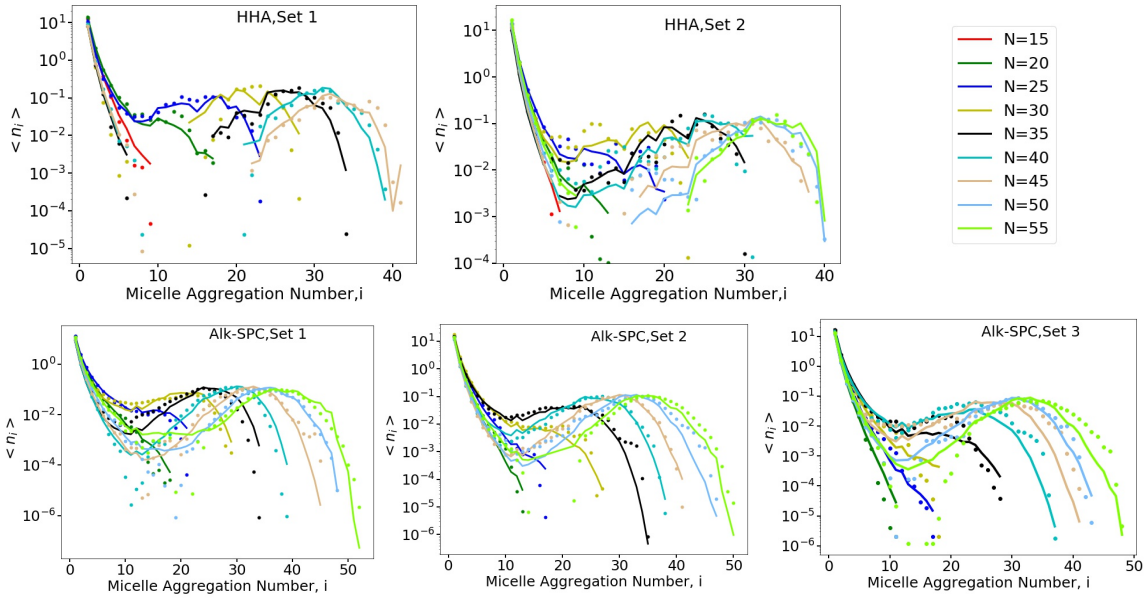


Figure 2. Cluster size distribution for simulations (symbols) and the PEACH fit (curves) for sets of MD simulations of OPC, with force fields and sets as labelled (set corresponding to Table 1) and number *N* of OPC given according to the color legend.

$$\frac{\Delta G(i)}{kT} = A(i^a - 1) - B \times \left(i^b - 1\right) + C \times (i^c - 1) \qquad (11)$$

Rationalizations for this modification, in the context of nucleation free energies, have been offered in the literature.[9, 33] Attempts to fit the free energy curves using eq. 11 with both the Maibaum/Chandler model and the quasi-droplet model are shown for Alk-SPC and HHA free energies (derived using a r_{cut} = 0.45 nm) in Fig. 3. (Agreement with the droplet model, fit not shown, was worse than with the Maibaum/Chandler model.) The quasi-droplet model fit the free energies of association derived from Alk-SPC simulations better than the Maibaum/Chandler model; as shown in Supplemental Information Fig. S3, this held true for all tested values of r_{cut}. Both models performed equally well for the less complete and noisier sets of HHA-derived data. We note that a similarly good fit could also be obtained using the combination *a*=2/3, *b*=1, *c*=4, but we do not know of any physical justification for the choice of *c*=4. It will be of interest to compare results from future simulations of other micelle-forming systems to tell whether the quasi-droplet model is generally the best descriptor for free energy of aggregation of globular micelles.

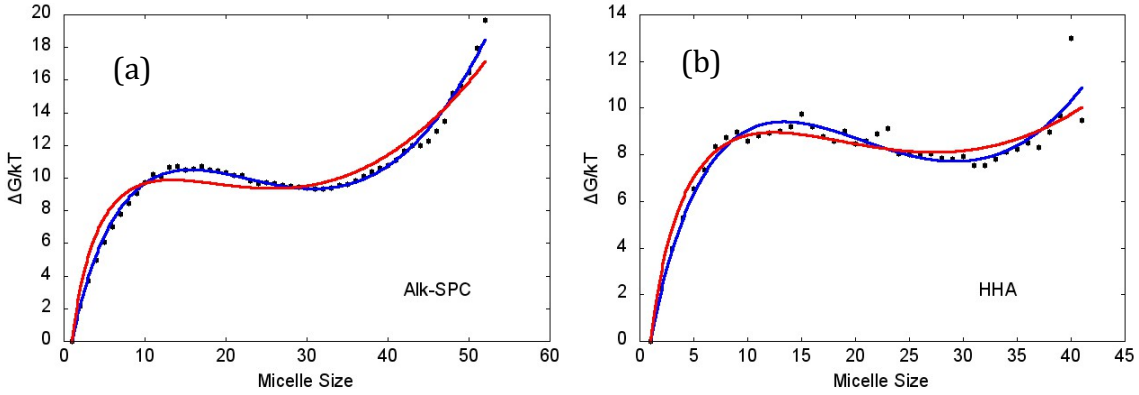


Figure 3. PEACH-derived cluster free energy curves (symbols) for OPC, simulated using (a)Alk-SPC and (b) HHA force fields, calculated for a total OPC concentration of 116 mM respectively. Free energy curves correspond to r_{cut} = 0.45 nm. Curves correspond to best fits by eq. 11 using quasi-droplet model (blue) and Maibaum-Chandler model (red).

Table 3. Best-fit parameters for the quasi-droplet model for r_{cut}=0.45 nm. (Parameters for other cutoffs given in Supplemental Information Table S1.)		
	Alk-SPC	HHA
A	3.864	3.821
B	1.122	1.133
C	0.08781	0.09102

Using the free energies of aggregation $\{\Delta G_i\}$ (or the equilibrium constants $\{K_i\}$ to which they are related by eqn. 8), the distribution of cluster sizes can be easily determined for any bulk concentration, which can then be used to derive information about cmc and average micelle size. The onset of micellization can be qualitatively gauged by considering the fraction of monomers that belong to clusters greater than a given size. Here we choose i=11, near the maximum in the cluster free energy curves, as the minimum cluster size to be defined as a micelle. Figure 4 shows this micellar fraction calculated from equilibrium constants derived using different cluster neighbor site distance cut-offs r_{cut}. Predictions are almost identical over the range of r_{cut} from 0.45 to 0.55 nm. Since only contacts between the alkyl tail sites are counted towards neighbour definitions, the insensitivity of the predictions to r_{cut} over this range suggests that the clusters that are captured by this definition have a well-defined micellar organization with with each surfactant's tail and forming redundant site-site contacts with other tails, so that distance fluctuations affecting one pair of sites does not affect its membership in the cluster. For Alk-SPC, the onset of micellization shifts slightly between r_{cut} = 0.55 nm and 0.60 nm and then strongly between r_{cut} = 0.60 nm and r_{cut} = 0.65 nm. Observation of the structures of clusters defined using this largest cut-off shows that under this criterion, two OPC molecules may be counted as neighbors even when their tails are separated by solvent (see Fig. S4 in Supplemental Information), which suggested to us that this cut-off length is too permissive for description of micelles as they are conventionally understood.

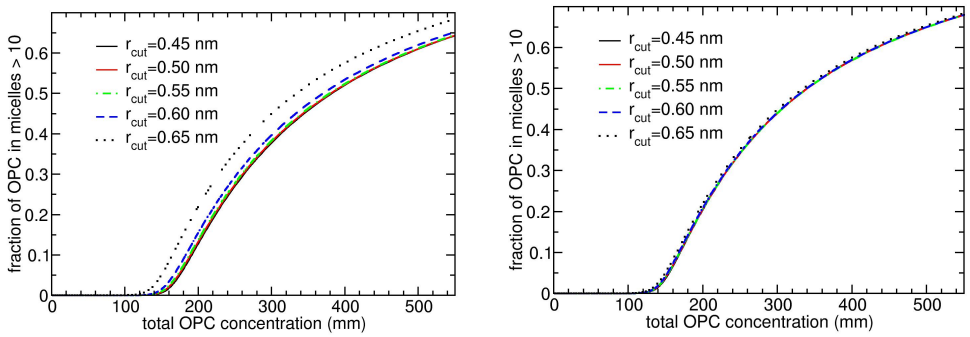


Figure 4. Fraction of OPC in clusters of size > 10 vs. total OPC millimolal concentration, as calculated from free energies derived from MD simulations performed with a) Alk-SPC force field and b) HHA force field. Different curves represent predictions based on free energies derived using different tail site cut-off distances r_{cut} in the neighbor criterion for defining clusters.

As has been noted[34], in comparing with experimental data for the cmc, it is important to consider what experiment was performed to determine cmc. A cmc value of 114 mM is reported by Anatrace, which markets OPC under the name FOS-choline-8. The experimental technique used in that determination was to measure the degree of capillary rise vs. surfactant concentration and find the intersection of straight lines fit to the low-concentration (below cmc) and high-concentration (above cmc) regimes.[35] Although we do not obtain interfacial tensions directly from simulations, we may presume from thermodynamic principles that surface tension is related linearly to surfactant activity, which in turn is approximately given by the concentration of unimers. The predicted dependences of unimer concentration on total OPC concentration, calculated from PEACH-derived $\{K_i\}$ using a cluster cut-off definition of r_{cut}=0.45 nm, are shown in Fig. 5a-b; data generated from $\{K_i\}$ derived using other cut-off values are presented in Supplemental Information Fig. S5. Bilinear fits give cmc estimates of 159 mM (Alk-SPC) and 143 mM (HHA), 40% and 25% higher than experiment. (The point of intersection depends on the concentration ranges used in the fits. Here we used the concentration ranges from 0.5 to 0.75 times the cmc and from 1.5 to 2.0 times the cmc, iterating as needed to reach a self-consistent intersection point.) The dependence of apparent cmc on r_{cut} is shown in Table 2.

The crossover concentrations correspond rather well to the onset of formation of micelles shown in Fig. 4. Average micelle size from the predicted cluster size distribution increases rapidly with total concentration just above the cmc but slowly after about 200 mM, as shown in Fig. 6. Increasing r_{cut} yields modest increases in average micelle size. The Alk-SPC model, in spite of having weaker attractions as reflected by the higher cmc, favors somewhat greater average micelle sizes than the HHA model. Concentration-dependent size distributions for the Alk-SPC model will be considered in more detail in section 3.4.

The concentration of unimers at equilibrium with micelles, or total concentration of unimers and small “premicelle” clusters, is often used to estimate the cmc. Pitfalls associated with this definition have been identified in the literature: the unimer concentration tends to drop off with increasing concentration above the cmc for ionic surfactants,[25] and excluded volume [36] influences the effective concentration. In

simulations, finite-number effects also play an important role;[22] it is therefore interesting to compare estimates results obtained from simulations of small systems to the cmc extracted from the global PEACH fits to cluster statistics. In Fig. 6, the mean concentration of OPC present as unimers and small clusters can be seen to oscillate as the total number of OPC is varied, increasing at low *N* until the first micelle is formed and then dropping sharply. In the regime after this drop, where the unimers and small clusters are at equilibrium with a micelle, their concentration ranges from 150 mM to 225 mM in the Alk-SPC system and from 125 mM to 180 mM in the HHA system. In such a small system, the unimer concentration is strongly influenced by the number of monomers left over after a micelle of optimal size is formed, and so indeed can only be used for a very rough measure of cmc.

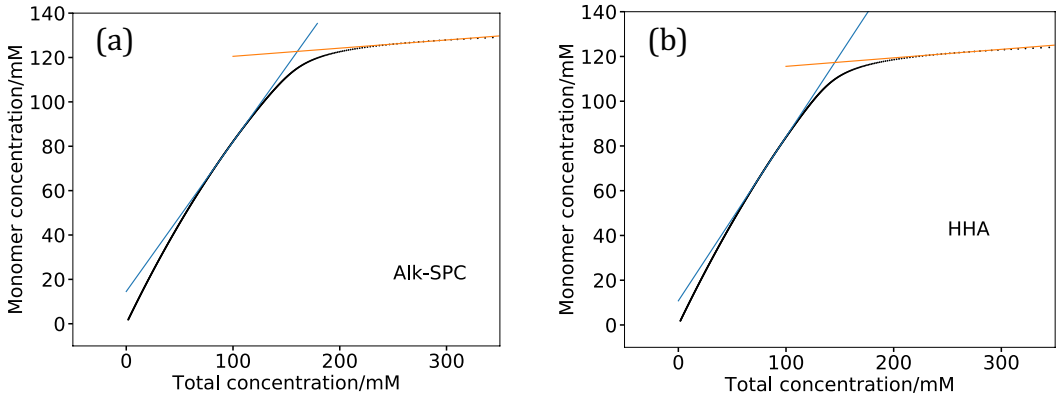


Figure 5. Free monomer vs. total OPC concentration calculated from PEACH-derived cluster free energy profile (a) with Alk-SPC (predicted cmc is 160 mM) (b) with HHA (predicted CMC is 145 mM)

Table 4: cmc values in mM predicted from PEACH free energies by force field and r_{cut}		
Cutoff/nm	Alk-SPC	HHA
0.45	160	145
0.50	158	145
0.55	156	144
0.60	152	143
0.65	141	140

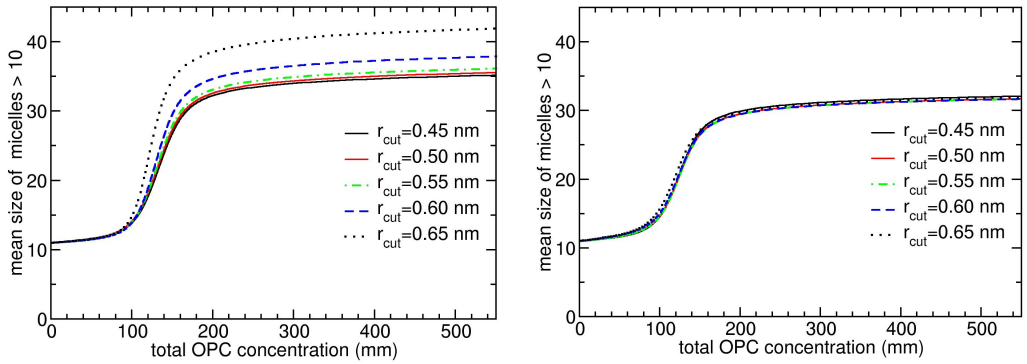


Figure 6. Mean micelle size vs. concentration, as calculated from free energies derived from MD simulations performed with a) Alk-SPC force field and b) HHA force field. Different curves represent predictions based on free energies derived using different tail site cut-off distances r_{cut} in the neighbor criterion for defining clusters.

Keeping the roughness of this measure in mind, we will use it now to make a rough comparison of the ability of the HHA model and the original TraPPE force field to reproduce the experimental cmc. (Having only performed two simulations using the original TraPPE CH_X/OW Lennard-Jones parameters, we do not have sufficient cluster statistics for a full PEACH analysis.) Two 200 ns trajectories of OPC using the original TraPPE parameters (with 40 and 30 OPC + 312 solvent) yielded concentrations of monomers in unimers and small clusters of 64 mM and 39 mM respectively. These values fall below the corresponding levels from HHA simulations (Fig. 7) by factors of 3.2 and 2.3 respectively. This suggests that full PEACH analysis of the TraPPE model would give a cmc roughly 3.2-2.3 times lower than the HHA model, *i.e.* in the range 44-62 mM. Such a range, about a factor of two lower than the experimentally reported cmc value of 114 mM, would be consistent with a previous report that surfactants modeled with TraPPE tails showed cmc values a factor of two lower than experiment.[25] We can conclude that tuning the force field to match hydration free energy, as was the goal in the HH-Alkane force field development[8], improves agreement with the experimental onset of micelle formation.

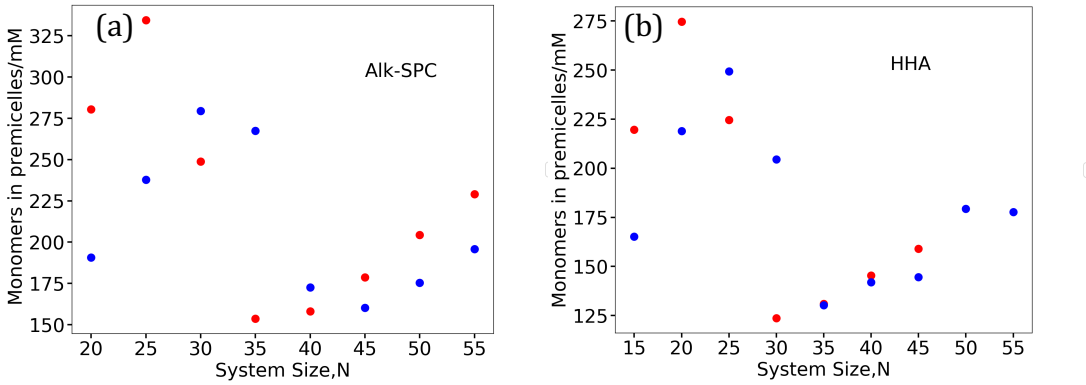


Figure 7. Concentrations of monomers in premicelles (i<=10) averaged over MD trajectories with varying total number of OPC using (a) Alk-SPC and (b) HHA force fields. (red dots for Set One, blue dots for Set Two as tabulated in Table 1.)

3.2 *Effects of limited system size and results of large system simulation*

The up-turn in cluster formation free energies evident at large cluster size *i* evident in Fig. 3 suggests that, if the phenomenological models for spherical micelle formation (like the quasi-droplet model) are applicable in this system, micelles much larger than ~50 monomers should be uncommon in solution. A large system (250 OPC, 30330 SPC; 0.46 molal) was simulated for 200 ns to test whether this is the case. Figure 8 indicates that the answer is not a simple yes or no. The cluster size histogram (Fig. 8a) does show a drop-off above $i = 50$, qualitatively consistent with the increase in free energy derived from small-*N* simulations; however, a long tail in the distribution extends to significantly higher cluster sizes. The nature of these larger clusters is suggested by a snapshot (Fig.8b) of the last frame of the trajectory, showing modest sized clusters that appear to be sticking together. The time dependence of the maximum cluster size (Fig. 8c) shows that the lifetimes of the large clusters are very short, consistent with their formation by transient bridges between more persistent smaller structures. The strong sensitivity to r_{cut} of the levels of large clusters is also consistent with an agglomeration of distinct hydrophobic cores whose alkyl tails only occasionally approach each other, rather than a continuous hydrophobic core connected through many pathways whose redundancy lowers sensitivity to fluctuations. The structures formed are not simple aggregates of the micelles present in the starting configuration of the simulation, but are thoroughly mixed as shown in an animation with the initial micellar components shown in different colors (Supplemental Information, Movie S1). We conclude that the large clusters formed are in fact aggregates of small clusters, connected by a partially solvated double layer of headgroups and transiently bridged by tail sites. Such a hierarchical structure has been invoked to explain the high flexibility of certain large micelles[37] or as a transient intermediate in the “sticky-collision” mechanism of dye exchange between micelles[38] The free energies of association of these higher-order clusters cannot be extrapolated from either the single-micelle simulations or the single-micelle theoretical models. Even considering the individual micelles that make up these higher-order clusters, their size-dependent free energies are likely to be perturbed away from what predictions based on free, fully-solvated micelles would give. Unfortunately, the computational expense required for long simulations at the large system sizes required prohibit a full quantitative investigation of these effects. Given that inter-micellar interactions should be weaker than intra-micellar interactions, it is likely that clustering would not become important until micelles start to become crowded together. Predictions related to the onset of micellization should therefore not be influenced by this phenomenon.

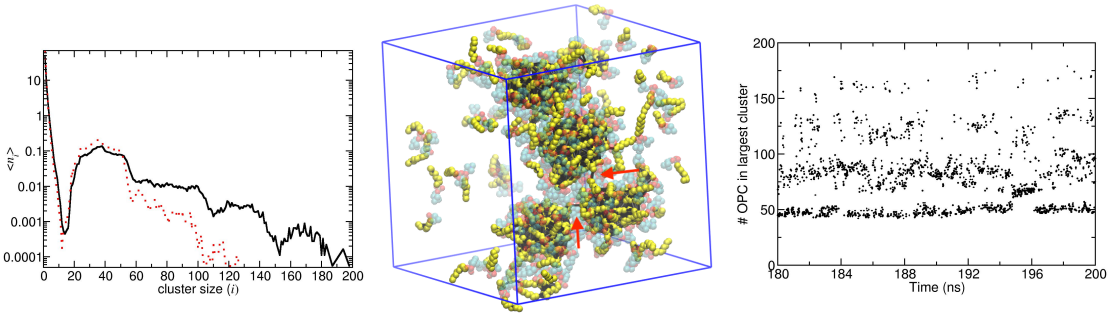


Figure 8. Results from 200 ns simulation of 250 OPC / 30330 SPC water. a) Cluster histograms generated from 20-200 ns using r_{cut} = 0.45 nm (red dashed curve) and r_{cut} = 0.60 nm (heavy black curve). b) Snapshot of last frame showing tail sites in yellow, headgroup sites in transparent colors, and "neighbor ties" connecting tail sites nearer than 0.60 nm in black. Red arrows point to neighbor connections bridging micelle clusters. c) Detail from 180-200 ns of the time dependence of the largest cluster size as evaluated using r_{cut} = 0.60 nm.

Whether inter-micelle effects of this nature are important experimentally for OPC or are an artifact of the headgroup force field parameters is not clear. On the one hand, this could explain neutron scattering results[39] that suggest dehydration and increased core-shell mixing in dodecyl phosphocholine (DPC) micelles (sharing the same headgroup as OPC but with 12-carbon tails) upon increasing concentration from 10 mM to 100 mM. On the other hand, formation of persistent supermicellar clusters would not be indicated by the reduction in apparent hydrodynamic radius obtained through dynamic light scattering in the same report.

3.3 Modeling SAXS data

Small and wide angle X-ray scattering (SAXS/WAXS) is a common technique to characterize the size and shape of micelles[40, 41]. The contribution of individual micelle structure to the overall scattering profiles, the "form factor," can be related to Fourier transforms of the electron density distributions, averaged over configurations and orientations of micelles. Information about cluster size and general shape can be inferred from fitting experimental scattering profiles to a simple model, for instance a two-shelled ellipsoid with the interior representing the hydrophobic interior and the outer shell representing the headgroups (which, due to the presence here of phosphorus, have significant contrast in electron density against the interior or the solvent). Known complications in interpreting the scattering profile include correlations between cluster positions (which grow in importance with concentration, as discussed in the previous section), polydispersity, and variations in mean micelle size with concentration[40]. Efforts have been made to generate scattering data from molecular simulation based on both implicit solvent[42] and explicit solvent[43] models.

Faramarzi and coworkers have used the WAXSiS server[23, 44] to generate scattering profiles generated from simulation coordinates can be used to select a micelle size that is most consistent with experiment.[43] In their report, they have suggested that imperfect agreement between the profile generated from simulation data (using the WAXSiS server) for dodecyl phosphocholine could be the result of polydispersity. We were unable to find SAXS data for OPC in the literature for direct comparison with the

simulation results, but instead here we aim to address the more general question: can the combined scattering from a polydisperse equilibrium mixture of micelles be adequately represented by the scattering profile of a single typical micelle size?

Fig. 9(a) shows the model which the WAXSiS server used to calculate the scattering profile. The hydration layer and excluded solvent were constructed from a short trajectory of simulation with position-restraining potentials to keep the structures close to initial structures.[23] As is shown in Fig. 9(b), variations were observed for SAXS curves of the same micelle size. This variation is due to the slightly different configuration of a micelle, thus an average of multiple curves for each micelle size was calculated. In these averaged SAXS profiles (Fig. 9(c)), a shift to lower q was observed with the increase of micelle size. The peaks are in the range of 2.3-2.7 nm^{-1}. Table 5 gives the radius of gyration calculated on the WAXSiS server based by Guinier analysis, which shows the radius is about 1.97 nm for a cluster size of 45. These features of OPC are consistent with previous experimental and simulation results on dodecyl phosphocholine.[41, 43]

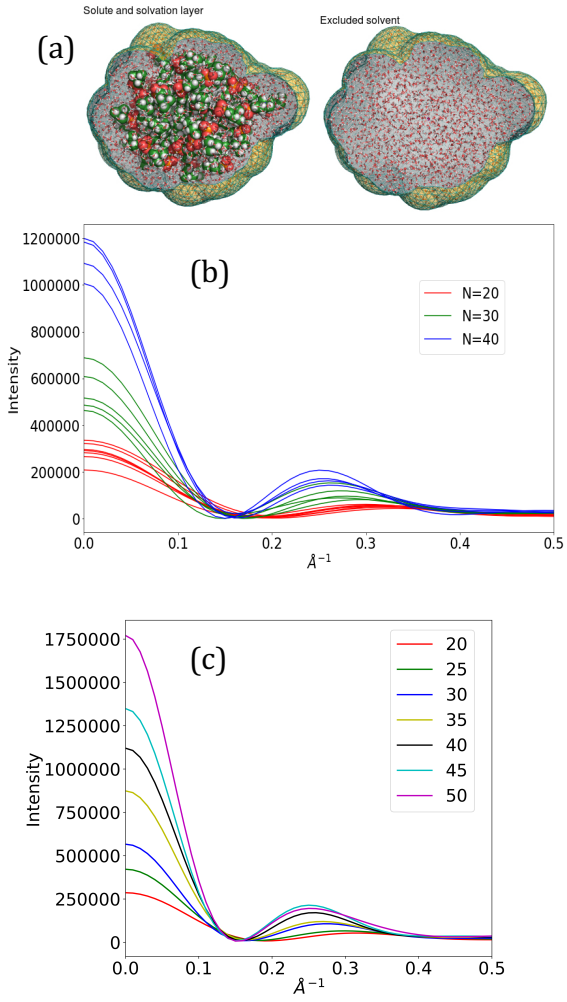


Figure 9. (a) The model of OPC micelle generated by WAXSiS server (b) Comparison of SAXS scattering profiles for different configurations of micelles with micelle size N=20,30,40 (c) Averaged SAXS scattering profiles for each cluster size from 20 to 50 at intervals of 5.

Based on the cluster free energy profile obtained from PEACH method, we can calculate the cluster size distribution for polydisperse systems for chosen concentrations. Fig. 10(a) shows the cluster size distribution at total OPC concentrations 198 and 787 mM. Composite SAXS curves were calculated using Eqn 9. Concentrations for each aggregation number *i* (curves in Fig. 10) are combined into weighting coefficients (bars in Fig. 10) for *i* = 20, 25, 30, etc. according to Eqn. 9. These coefficients are used in calculating a weighted sum of scattering profiles (Fig. 9c) to yield an approximate composite representation of the scattering from the size distribution. Premicelles for cluster sizes below 13 were not counted as they do not have SAXS features in the spectral range of interest. To compare this polydisperse scattering curve with ones for monodisperse micelle, the curves of N=30,35,40 were normalized based on I(q=0) to compensate the concentration variation. As seen in Fig. 11(a), the scattering profile of the mixture is quite similar to the profile of monodisperse 35-mers, with no obvious signature of the polydispersity presenting itself as a feature. From this we conclude that the deviation observed between experiment and the WAXSiS-derived results by Faramarzi *et al.*, where is unlikely to have its origin in polydispersity.

The specific deviation that was observed was overprediction by the model of the intensity of low-*q* scattering relative to a peak at q~1.7 nm^{-1}. We note that the ratio of intensities of the peak at q=0 to the peak at 1.7 nm^{-1} for our OPC structures is quite sensitive to even small changes in the choice of background solvent electron density used in the WAXSiS calculation. The default selection for this parameter is 334 nm^{-3} [44] whereas for the buffer used in the experiments is reported[40] to have slightly higher density at 3.4×10^2 nm^{-3}. Increasing the background solvent electron density from 334 to 340 nm^{-3} lowers the zero-*q* peak by 17% without affecting the height of the q=1.7 nm^{-1} peak. (not shown) This is not to say that the difference in density between solvent and buffer accounts for the discrepancy between simulated and experimental scattering profiles, rather that this element of the scattering profile is highly sensitive to details of contrast in electron density.

Table 5. Average radius of gyration for each cluster size

Micelle Size	Average radius of gyration(Å)
20	15.8443
25	17.4305
30	18.7757
35	19.1788
40	19.7028
45	20.2235
50	20.8528

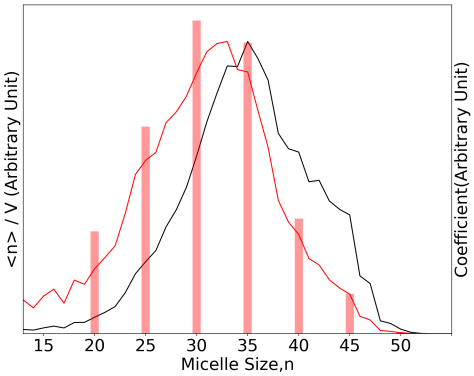


Figure 10. Micelle size distribution for c_{tot}= 198 mM. [$c_{monomer}$= 127 mM] and c_{tot}= 787 mM [$c_{monomer}$= 141 mM]. The bars represent the weighting coefficients for each group of cluster sizes for c_{tot}= 198 mM.

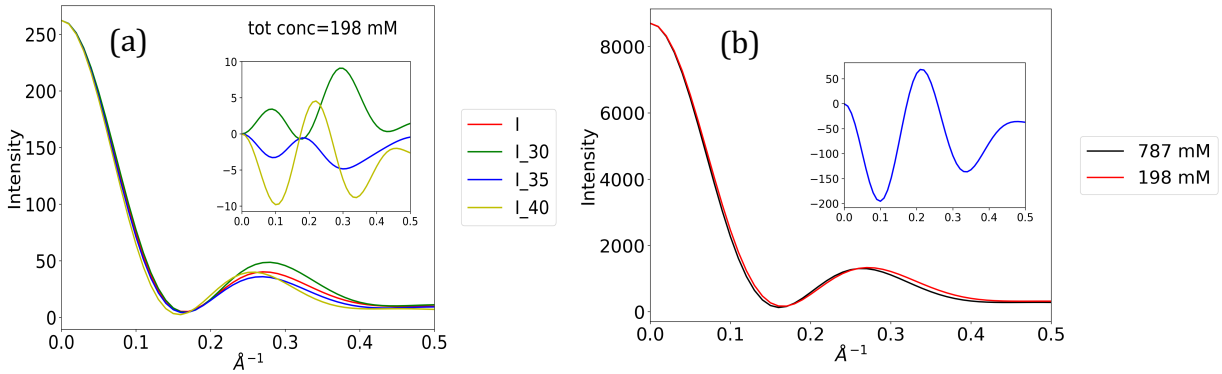


Figure 11. (a) Predicted SAXS scattering profile for polydisperse mixture at 198 mM, compared with contributions from individual size components; residual error is shown in inset. (b) Predicted SAXS scattering profile based on composite weightings at 198 mM and 787 mM. All profiles are scaled to have equal $I(q=0)$.

Even if the scattering profile at a single composition is not very sensitive to the breadth of the micelle size distribution, this underlying ΔG_i is reflected in how concentration affects the average micelle size. As we consider how concentration affects the scattering profile, we first note that the contribution from the structure factor (not considered here) will be strongly affected by concentration. Considering the form-factor alone, increasing concentration drives a small shift in distribution of aggregation numbers as shown in Fig. 10(a), with the peak size increasing from 33 to 35. The result is a slight shift in the small peak towards lower q in the calculated form factor (Fig. 11b).

4. Conclusions

In this work we used molecular simulation to study the micellization equilibria of octyl phosphocholine using the HH-Alkane force field and the newly developed Alk-SPC parameters to represent tail-tail and tail-water interactions. PEACH analysis was used

to obtain the free energy of association over a broad range of cluster sizes, which followed a dependence on aggregation number that was well-fitted by the quasi-droplet model for micelle assembly. The free energy of association was used to predict the concentration-dependent onset of micellization and cluster size distributions. The critical micelle concentration was in fair agreement with experiment; the HH-Alkane force field showed distinct improvement relative to TraPPE parameters on which it was based, reflecting an improved representation of the hydration free energy of the alkyl tails. Although the optimal cut-off distance to be used in defining a cluster is not clear, a range from 4.5 nm-6.0 nm produced only modest changes in predicted micellization properties. Use of the small-*N* approach offers simplicity but introduces the risk of neglecting features that emerge from larger-scale simulations, such as the significant tendency for micelles to cluster into higher-order aggregates seen here in a simulation of 250 OPC. Scattering profiles generated via the WAXSiS server from individual micelle structures were combined, using weights derived from the PEACH analysis, to represent a composite form factor at different OPC concentrations; it was found that only minor differences separated the composite scattering profile from the profile generated using a single typical micelle size. Further study of the dynamics of monomer exchange may be useful combined with these equilibrium calculations to provide details for general kinetic models of micellization.[45]

Acknowledgments

This work used the Extreme Science and Engineering Discovery Environment (XSEDE) Comet cluster at the San Diego Supercomputer Center, which is supported by National Science Foundation Grant No. ACI- 1548562, Allocation No. TG-MCB110144. This work also used the resources of the Cherry L. Emerson Center for Scientific Computation.

References

[1] B. Jönsson, O. Edholm, O. Teleman, Molecular dynamics simulations of a sodium octanoate micelle in aqueous solution, J. Chem. Phys., 85 (1986) 2259-2271.
[2] J. Shelley, K. Watanabe, M.L. Klein, Simulation of Sodium Octanoate Micelles in Aqueous Solution, Electrochimica Acta, 36 (1991) 1729-1734.
[3] E.N. Brodskaya, Computer simulations of micellar systems, Colloid Journal, 74 (2012) 154-171.
[4] T. Wymore, X.F. Gao, T.C. Wong, Molecular dynamics simulation of the structure and dynamics of a dodecylphosphocholine micelle in aqueous solution, J. Mol. Struct., 485–486 (1999) 195-210.
[5] D.P. Tieleman, D. van der Spoel, H.J.C. Berendsen, Molecular Dynamics Simulations of Dodecylphosphocholine Micelles at Three Different Aggregate Sizes: Micellar Structure and Chain Relaxation, The Journal of Physical Chemistry B, 104 (2000) 6380-6388.
[6] S. Abel, F.-Y. Dupradeau, M. Marchi, Molecular Dynamics Simulations of a Characteristic DPC Micelle in Water, Journal of Chemical Theory and Computation, 8 (2012) 4610-4623.
[7] M.G. Martin, J.I. Siepmann, Transferable Potentials for Phase Equilibria. 1. United-Atom Description of n-Alkanes, The Journal of Physical Chemistry B, 102 (1998) 2569-2577.
[8] H.S. Ashbaugh, L. Liu, L.N. Surampudi, Optimization of linear and branched alkane interactions with water to simulate hydrophobic hydration, The Journal of Chemical Physics, 135 (2011) 054510.

[9] L.A. Patel, J.T. Kindt, Cluster Free Energies from Simple Simulations of Small Numbers of Aggregants: Nucleation of Liquid MTBE from Vapor and Aqueous Phases, Journal of Chemical Theory and Computation, 13 (2017) 1023-1033.
[10] X. Zhang, L.A. Patel, O. Beckwith, R. Schneider, C.J. Weeden, J.T. Kindt, Extracting Aggregation Free Energies of Mixed Clusters from Simulations of Small Systems: Application to Ionic Surfactant Micelles, Journal of Chemical Theory and Computation, 13 (2017) 5195-5206.
[11] M.G. Martin, J.I. Siepmann, Predicting multicomponent phase equilibria and free energies of transfer for alkanes by molecular simulation, J. Am. Chem. Soc., 119 (1997) 8921-8924.
[12] D. Poger, W.F. Van Gunsteren, A.E. Mark, A New Force Field for Simulating Phosphatidylcholine Bilayers, J. Comput. Chem., 31 (2010) 1117-1125.
[13] C. Oostenbrink, A. Villa, A.E. Mark, W.F. van Gunsteren, A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: The GROMOS Force-Field Parameter Sets 53A5 and 53A6, J. Comput. Chem., 25 (2004) 1656-1676.
[14] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX, 1-2 (2015) 19-25.
[15] C.H. Bennett, Efficient estimation of free energy differences from Monte Carlo data, Journal of Computational Physics, 22 (1976) 245-268.
[16] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, Journal of Applied Physics, 52 (1981) 7182-7190.
[17] S. Páll, B. Hess, A flexible algorithm for calculating pair interactions on SIMD architectures, Comput. Phys. Comm., 184 (2013) 2641-2650.
[18] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems, The Journal of Chemical Physics, 98 (1993) 10089-10092.
[19] D. Ben-Amotz, G. Stell, Reformulation of Weeks−Chandler−Andersen Perturbation Theory Directly in Terms of a Hard-Sphere Reference System, The Journal of Physical Chemistry B, 108 (2004) 6877-6882.
[20] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, The Journal of Chemical Physics, 126 (2007) 014101.
[21] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, The Journal of Chemical Physics, 81 (1984) 3684-3690.
[22] J.T. Kindt, Accounting for Finite-Number Effects on Cluster Size Distributions in Simulations of Equilibrium Aggregation, Journal of Chemical Theory and Computation, 9 (2013) 147-152.
[23] C.J. Knight, J.S. Hub, WAXSiS: a web server for the calculation of SAXS/WAXS curves based on explicit-solvent molecular dynamics, Nucleic Acids Research, 43 (2015) W225-W230.
[24] I.S. Dmitri, H.J.K. Michel, Small-angle scattering studies of biological macromolecules in solution, Reports on Progress in Physics, 66 (2003) 1735.
[25] S.A. Sanders, M. Sammalkorpi, A.Z. Panagiotopoulos, Atomistic Simulations of Micellization of Sodium Hexyl, Heptyl, Octyl, and Nonyl Sulfates, The Journal of Physical Chemistry B, 116 (2012) 2430-2437.
[26] J.L.F. Abascal, C. Vega, A general purpose model for the condensed phases of water: TIP4P/2005, J. Chem. Phys., 123 (2005) 234505.
[27] X. Cheng, J.-K. Kim, Y. Kim, J.U. Bowie, W. Im, Molecular dynamics simulation strategies for protein–micelle complexes, Biochim. Biophys. Acta, 1858 (2016) 1566-1572.

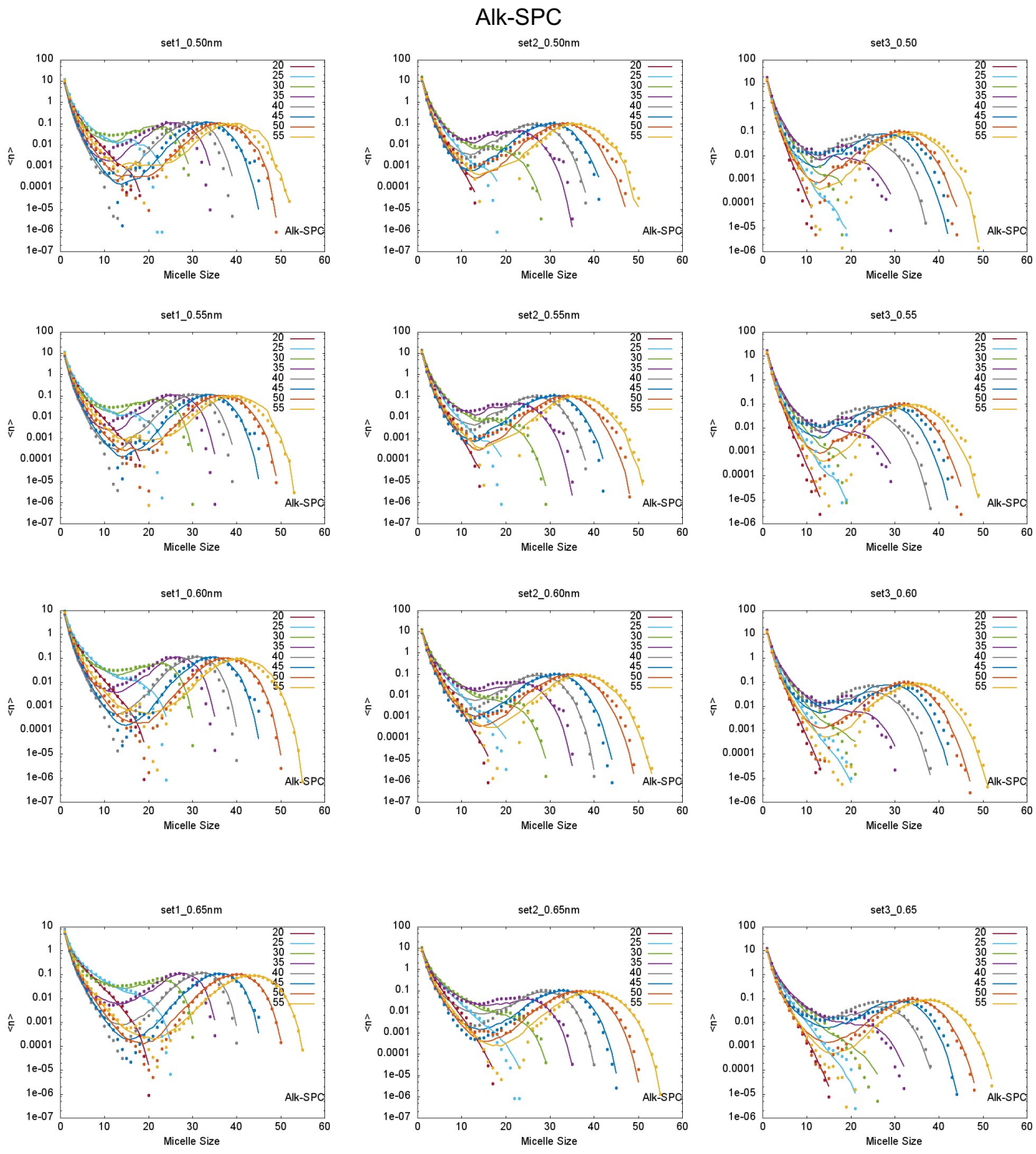
[28] N. Schmid, A.P. Eichenberger, A. Choutko, S. Riniker, M. Winger, A.E. Mark, W.F. van Gunsteren, Definition and testing of the GROMOS force-field versions 54A7 and 54B7, Eur. Biophys. J., 40 (2011) 843.
[29] Y. Mao, Y. Zhang, Thermal conductivity, shear viscosity and specific heat of rigid water models, Chem. Phys. Lett., 542 (2012) 37-41.
[30] A.I. Rusanov, F.M. Kuni, A.P. Grinin, A.K. Shchekin, Thermodynamic Characteristics of Micellization in the Droplet Model of Surfactant Spherical Molecular Aggregate, Colloid Journal, 64 (2002) 605-615.
[31] L. Maibaum, A.R. Dinner, D. Chandler, Micelle Formation and the Hydrophobic Effect, J. Phys. Chem. B, 108 (2004) 6778-6781.
[32] A.P. Grinin, A.I. Rusanov, F.M. Kuni, A.K. Shchekin, Thermodynamic Characteristics of a Spherical Molecular Surfactant Aggregate in a Quasi-Droplet Model, Colloid Journal, 65 (2003) 145-154.
[33] S.L. Girshick, C.P. Chiu, Kinetic nucleation theory: A new expression for the rate of homogeneous nucleation from an ideal supersaturated vapor, The Journal of Chemical Physics, 93 (1990) 1273-1277.
[34] M.A. Johnston, W.C. Swope, K.E. Jordan, P.B. Warren, M.G. Noro, D.J. Bray, R.L. Anderson, Toward a Standard Protocol for Micelle Simulation, J. Phys. Chem. B, 120 (2016) 6337-6351.
[35] G.D. Simth, V.G. Pena, An Inexpensive Glass Capillary 'Tensiometer' for Determining the Critical Micelle Concentration of Surfactants, The Chemical Educator, 14 (2009) 239-242.
[36] A.P. Santos, A.Z. Panagiotopoulos, Determination of the critical micelle concentration in simulations of surfactant systems, The Journal of Chemical Physics, 144 (2016) 044709.
[37] C. Tanford, Theory of micelle formation in aqueous solutions, J. Phys. Chem., 78 (1974) 2469-2479.
[38] Y. Rharbi, M. Li, M.A. Winnik, K.G. Hahn, Temperature Dependence of Fusion and Fragmentation Kinetics of Triton X-100 Micelles, J. Am. Chem. Soc., 122 (2000) 6242-6251.
[39] E. Pambou, J. Crewe, M. Yaseen, F.N. Padia, S. Rogers, D. Wang, H. Xu, J.R. Lu, Structural Features of Micelles of Zwitterionic Dodecyl-phosphocholine (C12PC) Surfactants Studied by Small-Angle Neutron Scattering, Langmuir, 31 (2015) 9781-9789.
[40] J. Lipfert, L. Columbus, V.B. Chu, S.A. Lesley, S. Doniach, Size and Shape of Detergent Micelles Determined by Small-Angle X-ray Scattering, J. Phys. Chem. B, 111 (2007) 12427-12438.
[41] R.C. Oliver, J. Lipfert, D.A. Fox, R.H. Lo, S. Doniach, L. Columbus, Dependence of Micelle Size and Shape on Detergent Alkyl Chain Length and Head Group, PLOS ONE, 8 (2013) e62488.
[42] A. Jusufi, A. Kohlmeyer, M. Sztucki, T. Narayanan, M. Ballauff, Self-Assembly of Charged Surfactants: Full Comparison of Molecular Simulations and Scattering Experiments, Langmuir, 28 (2012) 17632-17641.
[43] S. Faramarzi, B. Bonnett, C.A. Scaggs, A. Hoffmaster, D. Grodi, E. Harvey, B. Mertz, Molecular Dynamics Simulations as a Tool for Accurate Determination of Surfactant Micelle Properties, Langmuir, 33 (2017) 9934-9943.
[44] P.-c. Chen, Jochen S. Hub, Validating Solution Ensembles from Molecular Dynamics Simulation by Wide-Angle X-ray Scattering Data, Biophysical Journal, 107 (2014) 435-447.
[45] H. Diamant, D. Andelman, Free energy approach to micellization and aggregation: Equilibrium, metastability, and kinetics, Current Opinion in Colloid & Interface Science, 22 (2016) 94-98.

Supplemental Information

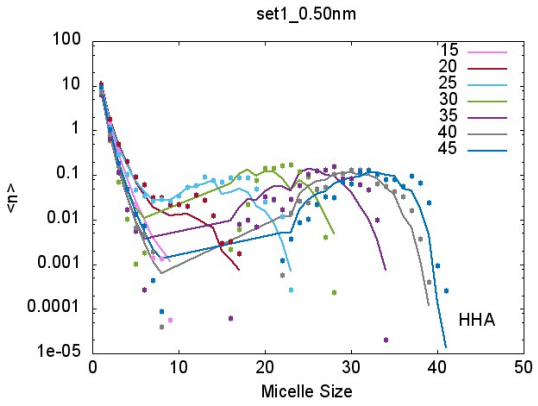
Derivation of micelle size-dependent free energies of aggregation for octyl phosphocholine from molecular dynamics simulation

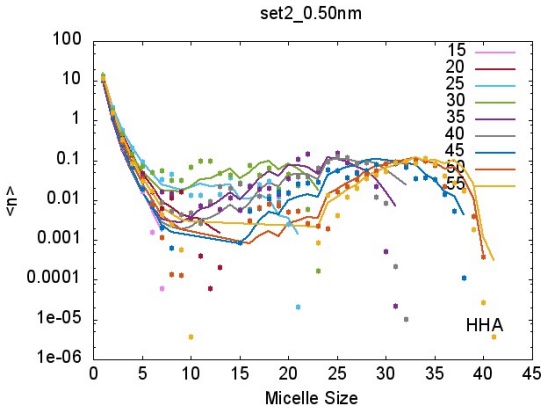
X. Zhang, J. G. Arce, and J. T.Kindt*
*jkindt@emory.edu

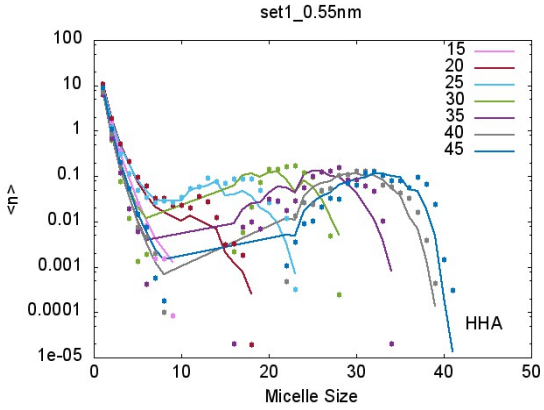
Fig S1. Cluster size distribution for simulations (symbols) and the PEACH fit (curves) for cutoffs 0.50, 0.55, 0.60, 0.65 nm; set number refers to compositions in Table I.

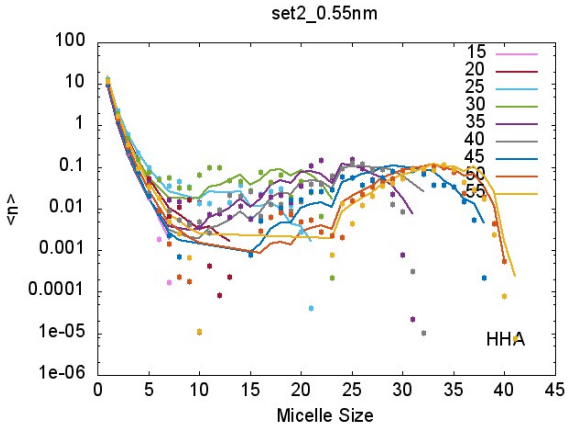


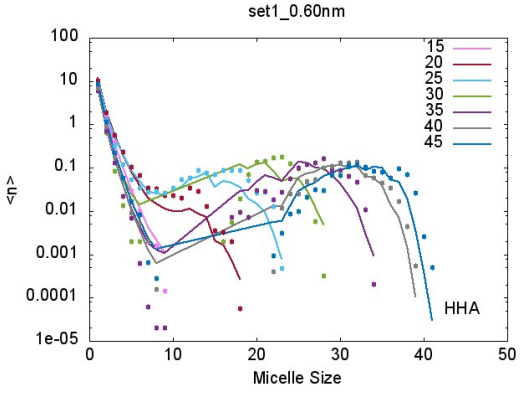
(Fig. S1 continued) HHA

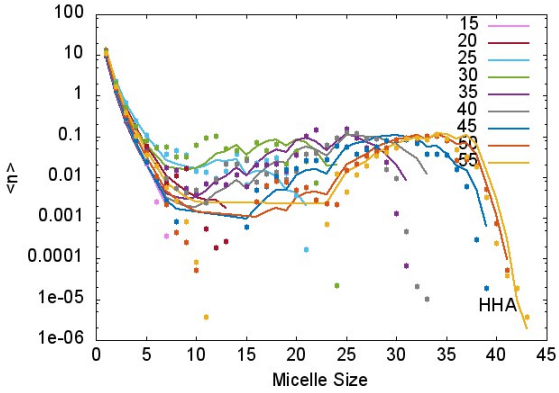

set1_0.50nm
<n>
Micelle Size
HHA
15
20
25
30
35
40
45


set2_0.50nm
<n>
Micelle Size
HHA
15
20
25
30
35
40
45
50
55


set1_0.55nm
<n>
Micelle Size
HHA
15
20
25
30
35
40
45


set2_0.55nm
<n>
Micelle Size
HHA
15
20
25
30
35
40
45
50
55


set1_0.60nm
<n>
Micelle Size
HHA
15
20
25
30
35
40
45


set2_0.60nm
<n>
Micelle Size
HHA
15
20
25
30
35
40
45
50
55

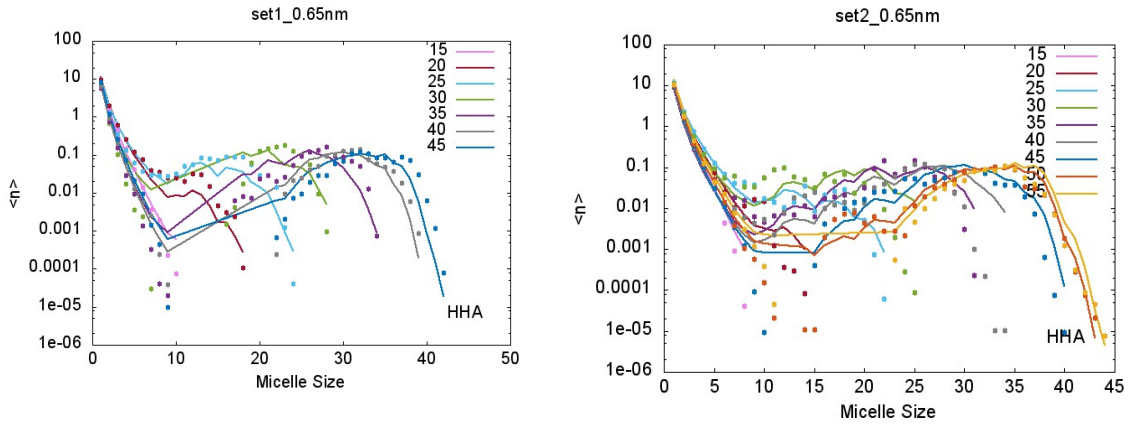


Figure S2: Comparison of dynamics of fluctuation of largest micelle over first 200 ns of simulation for Alk-SPC force field (red) and HHA force field (black). For both force fields, system is composed of 40 OPC and 3812 waters, and cluster sizes are plotted at 10 ps intervals.

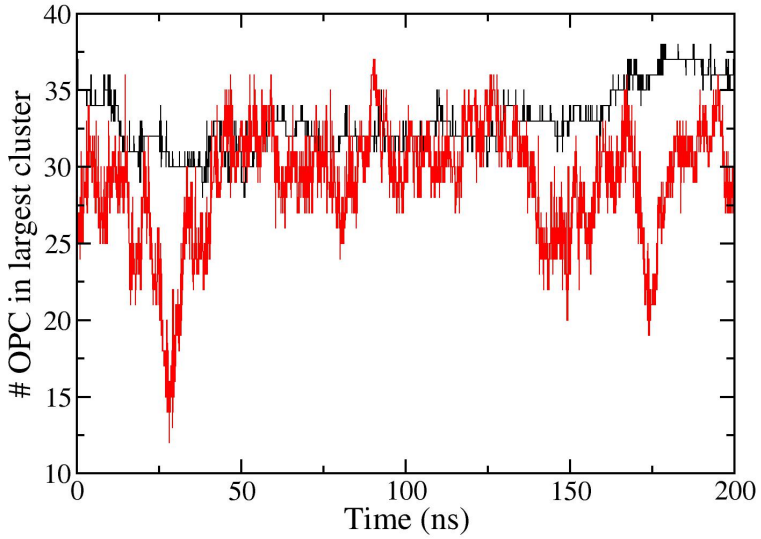
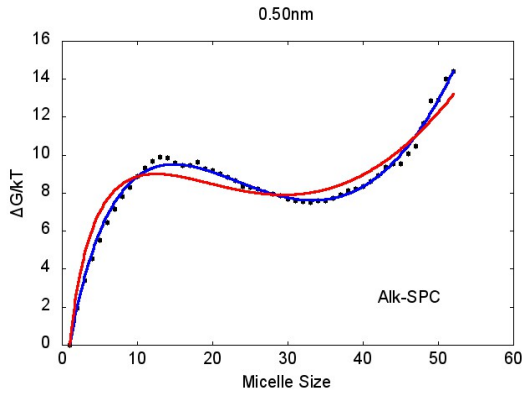
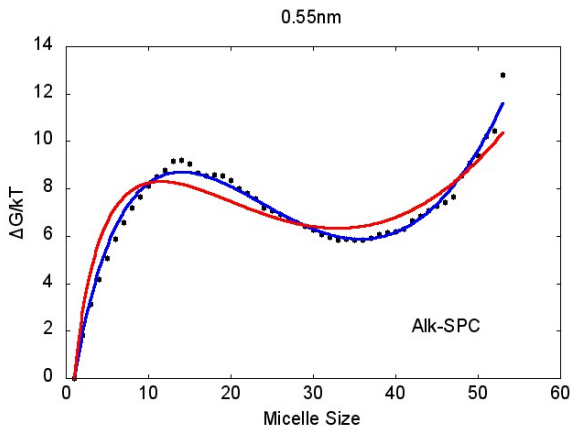
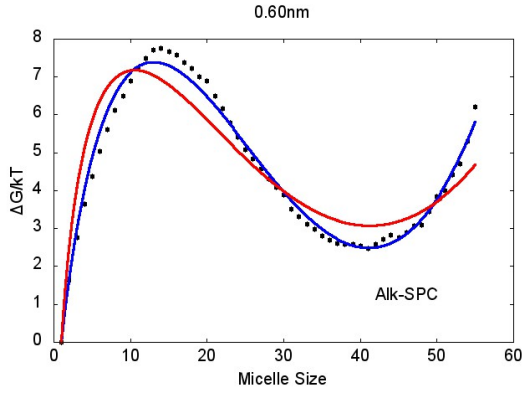


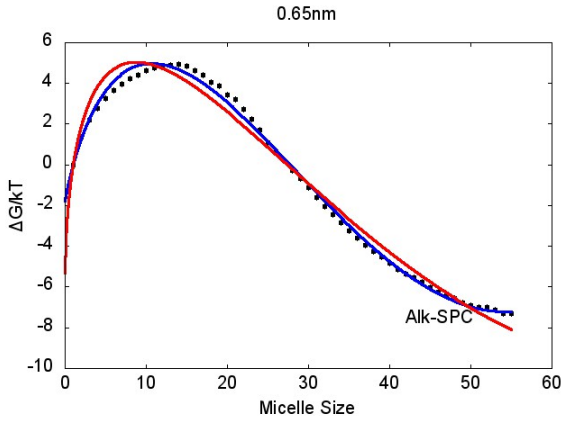
Figure S3. Fit cluster free energy curve (symbols) to two models (a) OPC simulated with Alk-SPC ($c_{monomer}$=116 mM)(b) OPC simulated with HHA ($c_{monomer}$=116 mM), with red curve for Maibaum-Chandler model and blue curve for "quasi-droplet" model.

Alk-SPC

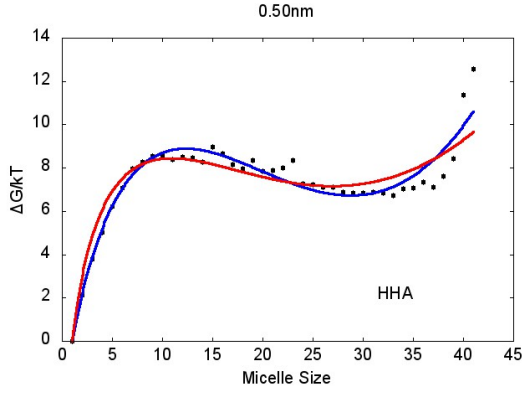


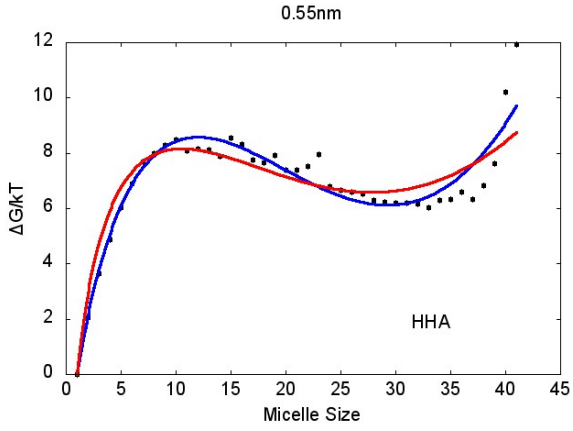


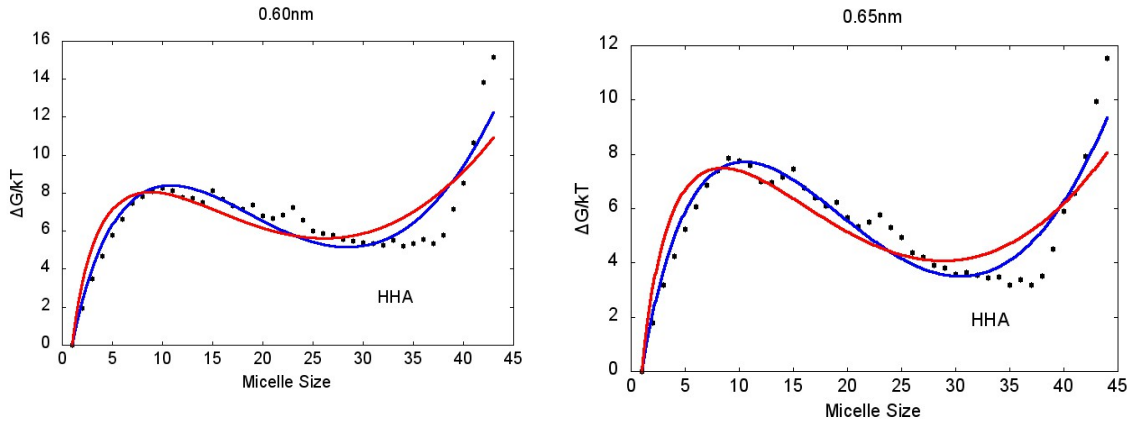




(Fig. S3 continued) HHA







$$\frac{\Delta G(i)}{kT} = A(i^a - 1) - B \times \left(i^b - 1\right) + C \times (i^c - 1)$$

a=1, *b*=3/2, and *c*=2

Table S1: Best-fit parameters for quasi-droplet model as applied to PEACH-derived cluster free energies with different force fields and cutoffs r_{cut}.

Cutoff/nm	Force Field	Parameters			
		A	B	C	RMS
0.50	Alk-SPC	3.7197	1.07384	0.0837218	0.236877
	HHA	4.33884	1.36368	0.115297	0.603107
0.55	Alk-SPC	3.51592	1.01825	0.0785707	0.290286
	HHA	4.26495	1.34579	0.11373	0.615484
0.60	Alk-SPC	3.09374	0.892887	0.0668185	0.249387
	HHA	4.68232	1.53398	0.133444	0.993096
0.65	Alk-SPC	2.42344	0.71269	0.0502369	0.260727
	HHA	4.37147	1.42875	0.122445	0.738317

Figure S4: Snapshot of 35 OPC/3812 SPC system showing tails sites in yellow, headgroup sites in elemental colors (cyan=CH_x, red=O, brown=P, blue=N), and "neighbor bonds" connecting pairs of tail sites closer less than r_{cut}=0.65 nm in black. Purple sites mark the tail of an OPC that is exterior to a large cluster but connected via a "neighbor bond". Solvents near the two neighboring sites are shown as thin sticks.

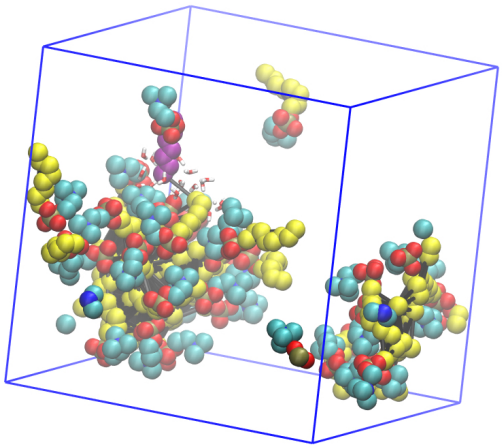
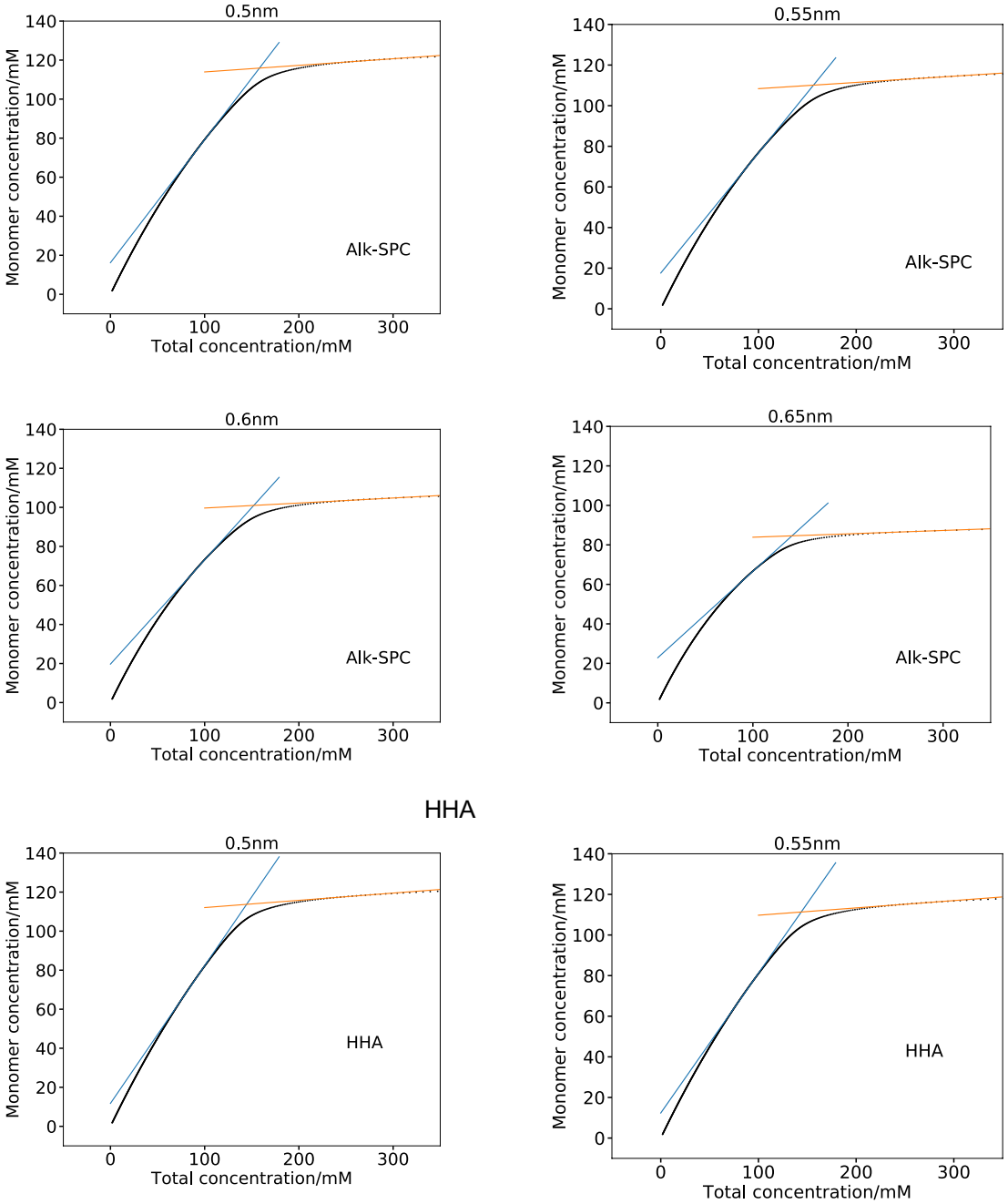
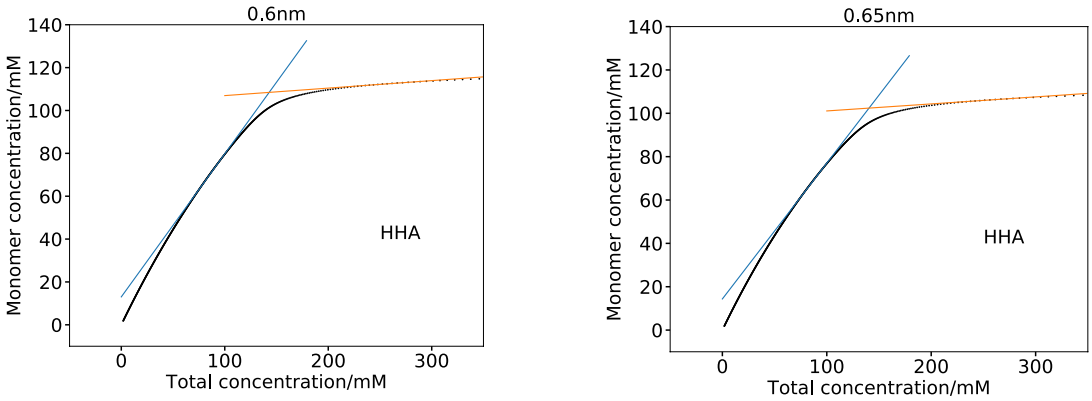


Figure S5. Free monomer concentration vs. total OPC concentration predicted from cluster free energy profile for different force fields, labelled with cluster criteria r_{cut} values in nm. Lines are fits whose intersections are used to estimate cmc.





Supplemental movie S1: Animation of 200 ns trajectory of 250 OPC and 30330 SPC waters (not shown). Alkyl tails are color-coded by the identity of the initial cluster to which they belong. Headgroups are shown in transparent gray.