Article

Finite-Size Effects and Optimal System Sizes in Simulations of Surfactant Micelle Self-Assembly

Published as part of The Journal of Physical Chemistry virtual special issue "Dave Thirumalai Festschrift".

Jonathan J. Harris, George A. Pantelopulos, and John E. Straub*

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The spontaneous formation of micelles in aqueous solutions is governed by the amphipathic nature of surfactants and is practically interesting due to the regular use of micelles as membrane mimics, for the characterization of protein structure, and for drug design and delivery. We performed a systematic characterization of the finite-size effect observed in single-component dodecylphosphocholine (DPC) micelles with the coarse-grained MARTINI model. Of multiple coarse-grained solvent models investigated using large system sizes, the non-polarizable solvent model was found to most accurately reproduce

SANS spectra of 100 mM DPC in aqueous solution. We systematically investigated the finite-size effect at constant 100 mM concentration in 23 systems of sizes 40−150 DPC, confirming the finite-size effect to manifest as an oscillation in the mean micelle aggregation number about the thermodynamic aggregation number as the system size increases. The oscillations in aggregation number mostly diminish once the system supports the formation of three micelles. Similar oscillations were observed in the estimated critical micelle concentration with a mean value of 1.10 mM, which is in agreement with experiment to 0.1 mM. The accuracy of using a multiscale simulation approach to avoid finite-size effects in the micelle size distribution and SANS spectra using MARTINI and CHARMM36 was explored using multiple long time scale 500 DPC coarse-grained simulations, which were back-mapped to CHARMM36 all-atom systems. It was found that the MARTINI model generally occupies more volume than the all-atom model, leading to the formation of micelles that are of a reasonable radius of gyration but are smaller in aggregation number. The systematic characterization of the finite-size effect and exploration of multiscale modeling presented in this work provide guidance for the accurate modeling of micelles in simulations.
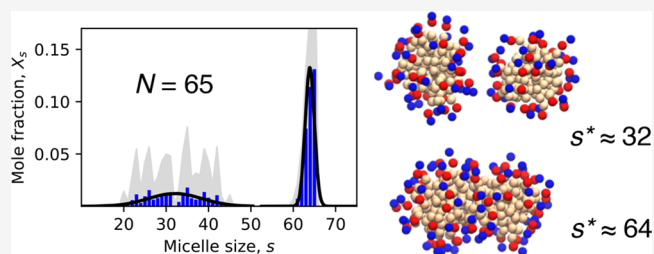
1. INTRODUCTION

The spontaneous self-assembly of surfactant molecules in aqueous solution is of great practical importance and fundamental interest due to the many applications of micelles and the forces which cause their formation. A significant challenge in the modeling of this process is capturing micelle size distributions, radii of gyration, and other characteristics that are in agreement with experiment, due in part to uncertainty in model predictions and variations in experimental observations. Dodecylphosphocholine (DPC), which has a dodecyl hydrocarbon tail and a zwitterionic phosphocholine head, is a commonly used surfactant. DPC micelles have been characterized through the methods of ultracentrifugation,[1] dynamic light scattering,[1] NMR,[2−4] small-angle neutron scattering (SANS),[5] and small-angle X-ray scattering (SAXS).[6,7] Molecular dynamics (MD) simulations of DPC micelle self-assembly have been conducted with coarse-grained (CG)[8−10] and all-atom (AA) modeling with explicit[11−16] and implicit[17−19] solvent. Fundamental aspects of the mechanism of self-assembly, the equilibrium state of micellar solutions, and the impact of micelle

encapsulated impurities on micelle size, however, remain poorly understood.

A special consideration in designing micelle simulations with periodic boundary conditions (PBCs) is the size of the system, specifically the number of surfactants present in one unit cell. In complex lipid bilayer mixtures, it has been shown that the liquid−liquid phase separation transition critically depends on system size.[20,21] In their self-assembly work with dissipative particle dynamics (DPD), Johnston et al. consider an adequate system size to be one which contains a sufficient number of surfactants for the formation of five micelles.[22] Kindt predicted the impact of finite-size on the aggregation number distribution when the total number of surfactants present is significantly
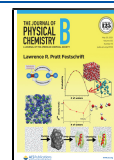
smaller than the thermodynamic limit.[23,24] According to Kindt, even systems with as few as one or two micelles can potentially have size distributions that approximate what is expected in the thermodynamic limit.[23] Because of the microsecond-order time scales associated with surfactant aggregation,[9] it is of computational interest to determine the smallest possible system size that will produce a realistic micelle size distribution.

In Tanford's treatment of micelle self-assembly, the set of aggregation numbers of a micellar system is governed by forces that promote and limit aggregation.[25] In particular, aqueous micelle formation is promoted by the tendency of the hydrocarbon tails to associate with each other, rather than with water. It is limited, in the case of zwitterionic DPC, by the comparatively stable solvation of the head groups as free monomers.[25] The preferred aggregation number of the micellar system occurs at the size that optimally balances the promoting and limiting forces, resulting in a unimodal and approximately Gaussian size distribution.[26]

Aggregation numbers for DPC micelles have been determined by various experimental methods. The particle weight obtained by ultracentrifugation of a 20 mM DPC solution by Lauterwein et al. was used to derive an aggregation number of 56 ± 5 surfactants per micelle.[1] Kallick et al. used NMR experiments to obtain an aggregation number of 44 ± 5 surfactants for DPC at 228 mM.[4] SANS experiments with 100 mM DPC conducted by Pambou et al. resulted in an aggregation number of 70.6 ± 5 surfactants.[5] SAXS methods have been used to obtain a range of aggregation numbers for DPC between 68 and 80 at 77 mM surfactant concentration.[6] The large observed differences in aggregation number appear to result from the differences in surfactant concentration, as well as the indirect nature of observation, and the sensitivity of the models used in data analysis. The range of aggregation numbers (s^*) from the literature is summarized in Table 1.

Table 1. Summary of s^* Values for DPC from Various Experiments and Computer Simulations

preferred aggregation number, s^*	conc (mM)	system size, $N_{surfactants}$	method	temp (K)
56 ± 5	20	experimental	Ultracentrifugation[a]	293
44 ± 5	228	experimental	NMR[b]	310
70.6 ± 5	100	experimental	SANS[c]	295
68–80	77	experimental	SAXS[d]	298
54	200	54	CHARMM36[e]	300
40–70	40	400	MARTINI[f]	
45	126	175	MARTINI[g]	370

[a]Lauterwein et al.[1] [b]Kallick et al.[4] [c]Pambou et al.[5] [d]Lipfert et al.[6] [e]Abel et al.[12] [f]Marrink et al.[9] [g]Sanders et al.[10]

Simulation studies of CG and AA DPC micelle formation have also been performed with not only varying surfactant concentrations, but also with different system sizes. In the AA single micelle self-assembly simulations conducted by Abel et al., 54 DPC surfactants were simulated from random initial configuration at a concentration of 200 mM.[12] 54 DPC were chosen on the basis of the Lauterwein experiments,[1] as well as proton NMR experiments,[2−4] which reported similar aggregation numbers. Marrink et al. simulated 54 DPC at 460 and 120 mM, finding that a wormlike micelle formed at the higher concentration, and a spherical micelle formed at the lower concentration, both with an aggregation number of 54

surfactants.[8] In MARTINI simulations of 400 DPC at 40 mM, Marrink et al. reported a range of aggregation numbers from 40 to 70 but noted that the simulation had not fully converged after 1 μs.[9] In a MARTINI self-assembly simulation with 175 DPC at 370 K and 126 mM, Sanders et al. found an equilibrium size distribution that was unimodal and centered around 45 surfactants.[10]

In order to efficiently simulate micelles, it is of interest to determine the lower limit of the number of surfactants necessary for a self-assembly simulation to produce a realistic size distribution. In 1996, Palmer et al. simulated micelle self-assembly using a CG model with 100 surfactants. They reported multimodal size distributions with sharp, separated peaks,[27] suggesting the presence of a finite-size artifact. Kindt predicted that, in the region of small numbers of surfactants, the mean number of micelles and mean aggregation number oscillate around the values expected at the thermodynamic size limit, at which the number of micelles varies linearly with the total number of surfactants in the system.[23] Although this finite-size effect has been predicted in other work,[23,24] no study has been conducted to fully characterize the effect on self-assembly simulations over an interval of system sizes approaching the thermodynamic limit.

In this work, we extensively characterize the finite-size effect on the aggregation of DPC surfactants using the CG MARTINI 2 model. Using the SANS spectra of 100 mM DPC micelles[5] of Pambou et al. as a reference,[5] we evaluate the ability of four MARTINI 2 solvent models to accurately simulate DPC micelle self-assembly. We determine the nonpolarizable solvent model to produce the most accurate SANS spectrum. We then investigate the finite-size effect on micelle size distributions with this model at 100 mM surfactant concentration by simulating systems of 40−150 surfactants, at an interval of 5 surfactants, resulting in the formation of 1, 2, and 3 micelles throughout this range. We observe oscillations in the preferred aggregation numbers and number of micelles across the different system sizes, similar to the predictions of Kindt.[23] We find that the finite-size effect is largely diminished once there are sufficient surfactants for three micelles to form at an aggregation number consistent with what is expected in the thermodynamic limit. On the basis of our observations, as a rule of thumb we suggest that systems used in simulations of surfactant self-assembly be composed of at least 3 times the known experimental aggregation number of the surfactant that is being studied. The micelle size distributions observed for each increasing system size are analyzed by using the thermodynamic model of Tanford.[26] We also test the efficacy of a multiscale approach by using large 500 DPC MARTINI simulations to simulate micelle self-assembly and then back-mapping with the backward.py method[28] to CHARMM36 AA representations. The CG and back-mapped AA simulations are compared directly to an experimental SANS spectrum.[5] We find that the back-mapped AA micelles shrink to smaller than expected sizes due to the significantly larger volume of each MARTINI DPC molecule, suggesting that MARTINI 2 size distributions cover a range of significantly smaller aggregation numbers than CHARMM36 at equilibrium, which indicates that the back-mapped systems are not in equilibrium.

Taken together, our results provide a consistent picture of DPC surfactant self-assembly at a range of system sizes, evaluate the accuracy of the MARTINI model for micelle simulation, and establish appropriate system sizes for the study of DPC micelles, which may be generalized as a method for determining the lower

limit of the number of a variety of surfactants necessary for self-assembly.

2. METHODS

2.A. Identifying the Most Accurate MARTINI Model for Micelle Self-Assembly. Four systems of the same composition were built with different versions of the MARTINI 2 force field for water and ions: nonpolarizable water (W),[29] polarizable water (PW),[30] refined polarizable water (refPOL),[31] and refPOL with polarizable ions (polIon).[32] All of the systems used the v2.0 force field for DPC,[9,29] which consisted of approximately 100 mM DPC with 150 mM NaCl, and contained 500 DPC molecules, 750 Na^+ ions, 750 Cl^- ions, and 68 775 MARTINI waters. In the W system, 10% of the water molecules were WF antifreeze. The choice of 500 DPC molecules was made in an effort to avoid finite-size effects. We assumed that the systems would behave as expected in the thermodynamic limit. Six replicates of each system were simulated from random initial starting configurations. The W systems were simulated for 5.2 μs, and PW, refPOL, and polIon were simulated for 4.0 μs. For fairness in model comparison, only data up to 4.0 μs were used in the SANS spectrum and ergodic measure calculations for the W model. In addition, the same nonbonded interaction schemes that are usually applied to MARTINI polarizable water models were also applied to the nonpolarizable water model in order to avoid any confounding differences between these schemes. The MARTINI W water model was determined to be the most appropriate model for DPC micelle self-assembly as it led to the best agreement with the experimental SANS spectrum.[5] In further simulations investigating finite-size effects on micelle formation, the W model was used.

2.B. CG Systems for Probing Finite-Size Effects. To characterize this finite-size effect, systems with $N = 40-150$ DPC were constructed, at an interval of 5 surfactants. In order to obtain reliably averaged data, 20 replicates of each system were simulated from random initial starting configurations. In order to assess the impact of the number of replicates on the resulting micelle size distributions, 20 additional replicates of the $N = 75$ system were completed. The 20 replicate and 40 replicate equilibrium ensembles were compared and found to have very similar preferred aggregation numbers and size distributions (see Figure S8 in the Supporting Information). The length of each simulation varied depending on the system size and was adjusted to achieve an equilibrium micelle size distribution. System sizes with $N = 40-100$ were simulated for 2 μs, while system sizes $N = 105-150$ were simulated for 5 μs. In each system, the numbers of water beads and NaCl ions were adjusted so that the concentration of DPC for each system was held to approximately 100 mM and the concentration of NaCl was held to approximately 150 mM.

2.C. CG to AA Back-Mapping for Micelle Characterization. To test the efficacy of a multiscale approach to micelle simulation, DPC micelles were self-assembled using the MARTINI model with nonpolarizable solvent (W) and 500 surfactants. The resulting equilibrium configuration was back-mapped to the AA CHARMM36 force field. The differences between the micelles in the two representations were evaluated in terms of radius of gyration, size distributions, preferred aggregation numbers, and SANS spectra, along with the experimental SANS spectrum.[5] The AA systems were constructed by back-mapping the W systems with $N = 500$ DPC using backward.py[28] and the DPPC.map file which was modified to suit DPC. The resulting AA systems contained 500

DPC, 275 100 waters, and 750 each of sodium and chloride ions. This proportion of species corresponds to a DPC concentration of approximately 100 mM, in concurrence with the experimental conditions.[5] All 6 of the MARTINI W systems were back-mapped for DPC atom positions only from their 5200 ns configurations and simulated for an additional 20 ns with the AA representation.

2.D. CG Model Simulation Details. The MARTINI model test systems were composed using the GROMACS 2018.3 insert-molecules function, which randomly places molecules in the simulation box. Dodecahedral periodic boundary conditions were used. All of the MARTINI systems were energy minimized, equilibrated, and performed at constant NPT using GROMACS 2018.3 on GPUs.[33,34] The systems were energy minimized using steepest descent. For MD, the leapfrog integrator was used with a 20 fs time step. The Verlet cutoff scheme for neighbor searching was applied. For nonbonded interactions, particle-mesh Ewald (PME) electrostatics was applied over a range of 0−1.1 nm with an "epsilon-r" relative dielectric constant of 2.5 for all systems except for W, which had a dielectric constant of 15. Lennard-Jones was applied with a shifting function from 0.9 nm to the cutoff at 1.1 nm. For the thermostat, velocity-rescaling was used with a coupling time of 1 ps at 295 K. Isotropic Parrinello−Rahman pressure coupling was used at 1 bar with a coupling time of 12 ps and 4.5×10^{-5} bar^{-1} compressibility. Bead coordinates and system energies were written every 1 ns.

The systems designed to extensively investigate finite-size effects used the same parameters described above for the MARTINI W test system. The particle definitions for water, ions, and DPC correspond to the MARTINI v2.0 force field.[9,29]

2.E. AA Model Simulation Details. For the AA system, the CHARMM36 force field[35] was used for all particle definitions with CHARMM36 TIP3P water. After back-mapping the DPC positions, the systems were solvated as described above using insert-molecules and simulated using GROMACS 2020 with GPUs.[33,34] The systems were minimized using the steepest descent minimization algorithm. For MD, the leapfrog integrator with a 2 fs time step was used. The "Verlet" cutoff scheme for neighbor searching was applied with updates every 20 steps. For nonbonded interactions, particle-mesh Ewald (PME) electrostatics from 0 to 1.2 nm and "Shift" Lennard-Jones van der Waals from 1.0 to 1.2 nm were used. For the thermostat, velocity-rescaling was used with a coupling time of 1 ps at 295 K. Isotropic Parrinello−Rahman pressure coupling was used at 1 bar with a coupling time of 2 ps and 4.5×10^{-5} bar^{-1} compressibility. Atom coordinates and system energies were written every 100 ps.

2.F. Thermodynamic Analysis Using Tanford Model. The DPC molecules were clustered into micelles so that the equilibrium size distribution could be constructed and characterized. For all simulations in this study, single-link hierarchical clustering from Scipy was used with a cutoff of 9 Å for the MARTINI systems and 12 Å for the all-atom systems to determine the DPC cluster sizes.[36] A micelle was defined as a cluster with more than 10 surfactants. The cutoffs used for clustering were applied to the terminal tail carbon atom (for AA) or bead (for CG) distances and were determined on the basis of visual inspection, radial distribution functions (for CG), radius of gyration (for AA), and plots of the number of micelles over time (see Figure S6 in the Supporting Information). The micelle definition was determined on the basis of the populations of the size distributions, which generally do not contain clusters between 3 and 10 surfactants at equilibrium, no matter the

system size. The clustering and micelle characterization was conducted using Python with Scipy and MDAnalysis libraries.[37,38] All MD visualizations and snapshots were made using Visual Molecular Dynamics (VMD).[39]

The time to reach equilibrium for each CG self-assembly trajectory was determined on the basis of the time series of the number of clusters present in each system over time. In order to quantify the point at which each trajectory had converged into a stable micelle size distribution, the ergodic measure of Thirumalai and co-workers[40−44]

$$\Omega(t) = \frac{1}{N}\sum_{j=1}^{N}[f_j(t) - F(t)]^2 \tag{1}$$

was calculated, where $f_j(t)$ is the number of micelles in the jth replicate simulation at the tth frame, and $F(t)$ is the mean number of micelles at this frame between all N replicates.

In order to evaluate the thermodynamics of the finite-size effect MARTINI simulations, Tanford's treatment of equilibrium micelle size distributions was applied. Tanford's fundamental equation[26] defines the mole fraction X_s of aggregate size s

$$\ln(X_s) = -\frac{s\Delta G_s^0}{RT} + s\ln(X_1) + \ln(s) \tag{2}$$

where ΔG_s^0 is the free energy change contribution from each of s monomers in the formation of a micelle of size s, and X_1 is the mole fraction of monomers in the system. The value of s at which the slope of $\ln(X_s)$ is zero corresponds to the preferred aggregation number, s^*, of the surfactant system. A micelle of size s^* has the optimal balance of entropy and enthalpy, which mainly correspond to surfactant tail conformational distributions and head group solvation, respectively.[26]

Tanford also constructed a Gaussian model that can be used to approximate the preferred aggregation number, s^*, from a micelle size distribution which has the form

$$X_s = X_s^* \, e^{-a(s-s^*)^2} \tag{3}$$

where X_s is the mole fraction of micelles with s members, X_s^* is the mole fraction of micelles with s^* members, and a is a constant parameter.[26]

2.G. Comparison with Experimental SANS Spectrum. The results of the all-atom simulations were used to construct simulated SANS profiles through the use of the Debye scattering equation.[45] The Debye scattering equation solves for the scattering intensity defined in terms of the interatomic distances, r_{ij}, within each aggregate, such that

$$\frac{I(Q)}{I(Q_0)} = \frac{\sum_{i,j} b_i b_j^* \frac{\sin(Qr_{ij})}{Qr_{ij}}}{\sum_{i,j} b_i b_j^* \frac{\sin(Q_0 r_{ij})}{Q_0 r_{ij}}} \tag{4}$$

where b_i and b_j are the scattering lengths[46] of atoms i and j, and Q is the scattering vector.[47] To make a direct comparison between the results of the all-atom simulations and the experimental results of Pambou et al.,[5] the intensity vs Q points for the fully deuterated sample of DPC at 100 mM and 295 K were extracted using Engauge Digitizer.[48] The scattering lengths used for each DPC atom in femtometers are 0.66 for ^{12}C, 0.65 for ^{2}H, 0.94 for ^{14}N, 0.58 for ^{16}O, and 0.50 for P.[46]

In addition, the radius of gyration (R_g) of the micelles was calculated from the distance distribution function,[49,50] $p(r)$, which is the probability function of interatomic distances, r_{ij}, within each micelle. R_g^2 is half of the second moment of the pair distance distribution function[50]

$$R_g^2 = \frac{\int_0^\infty r^2 p(r)\mathrm{d}r}{2\int_0^\infty p(r)\mathrm{d}r} \tag{5}$$

or $\sqrt{3/5}$ times the mean radius of the pair distance distribution function[12] of an approximately spherical micelle. The radius of gyration was used as another way to compare the simulated and experimental DPC micelles.

3. RESULTS AND DISCUSSION

3.A. Identifying the Optimal CG Model for DPC Micelle Self-Assembly. As a way of determining which MARTINI solvent model is most appropriate for micelle self-assembly, the equilibrium distributions of micelles resulting from the large (N = 500 DPC) simulations of each model were compared directly to the SANS profile produced by Pambou et al.[5] The systems were considered to have reached equilibrium after 2.5 μs based on the convergence of the ergodic measure of the average number of micelles present (see Figure S2a in the Supporting Information). Further analysis investigated the mean relaxation time for micelle formation for each solvent system, using the function[11]

$$G(t) = \frac{N_m(t) - N_m(T)}{N_m(0) - N_m(T)} \tag{6}$$

where $N_m(T)$ is the equilibrium number of micelles, which ranged from 10 to 16 across the solvent systems. The stretched exponential[11]

$$\phi(t) = \exp\left(-\left(\frac{1}{\tau}\right)^\alpha\right) \tag{7}$$

was fit to the $G(t)$ function for each solvent system. Values of the exponent α ranged from 0.65 to 0.76, consistent with the observed degree of dynamical heterogeneity. The average relaxation times were obtained using[11]

$$\langle t \rangle = \frac{\tau}{\alpha}\Gamma\left(\frac{1}{\alpha}\right) \tag{8}$$

and found to be 287, 491, 727, and 707 ns for simulations using W, PW, refPOL, and polIon water models, respectively (see Figures S2 and S3 in the Supporting Information). Because the MARTINI representation does not include every atom present in DPC, scattering lengths for each DPC bead were approximated on the basis of the 157.63 fm deuterated head and 246.53 tail scattering lengths[5] reported by Pambou et al. so that the scattering lengths are 78.65 fm for each of the two head beads and 82.18 fm for each of the three tail beads. It was found from a χ^2 analysis that the nonpolarizable W model yields the most realistic micelle distribution, with a χ^2 of 1.37, while PW had 1.90, refPOL had 2.35, and polIon had 2.53 (Figure 1), where

$$\chi^2 = \frac{[I(Q)/I(Q_0)^{\mathrm{sim}} - I(Q)/I(Q_0)^{\mathrm{exp}}]^2}{I(Q)/I(Q_0)^{\mathrm{exp}}} \tag{9}$$
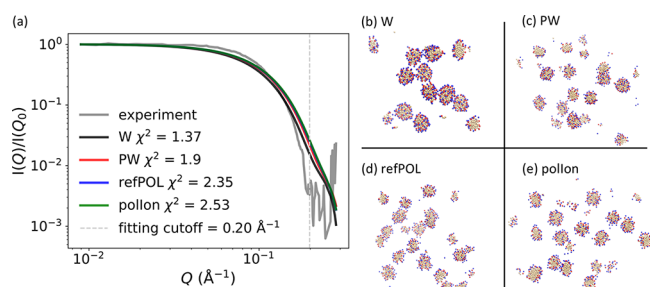
Figure 1. (a) Normalized SANS spectra computed from DPC micelle data for each MARTINI simulation with solvent models W (black), PW (red), refPOL (blue), and polIon (green), respectively. The experimental spectrum from Pambou et al. (gray)[5] shows the best agreement with DPC in the W water model. The dashed line is the cutoff Q value for the χ^2 analysis between each model and the experimental profile. Snapshots of the DPC micelle systems at equilibrium with the (b) W water model, (c) PW water model, (d) refPOL water model, and (e) polIon water model.

for each Q value from 0.0090 Å^{-1} up to 0.20 Å^{-1}. The cutoff of Q = 0.20 Å^{-1} was applied due to the amount of noise past this point in the experimental spectrum. The χ^2 values are not overly sensitive to the chosen cutoff (see Figure S7 in the Supporting Information).

The performances of the MARTINI solvent models were also evaluated on the basis of which size distribution is most reasonable when compared to previous DPC micelle experiments. In order to quantify the micelle size distributions produced by each water model, Tanford's Gaussian approximation (eq 3) was used to calculate preferred aggregation numbers (s^*). The Gaussian function was applied to the size distributions which contained data from 6 replicate simulations of each water model (see Figure S1 in the Supporting Information). It was found that the W model yielded s^* = 44.2 ± 7.0, PW = 36.2 ± 9.9, refPOL = 31.8 ± 7.7, and polIon = 33.0 ± 7.5 surfactants. It is fortuitous that the W model is also the most computationally affordable of the MARTINI solvents, with production speeds 4 times those of the other models.

3.B. Finite-Size Effects on DPC Self-Assembly Using CG Models. The initial characterization of finite-size artifacts was completed by computing the average number of micelles, the preferred aggregation number (s^*) for each system size, and the aggregation number expected in the thermodynamic limit, s^{thermo} (Figure 2). The micelle equilibrium was considered to be achieved after 1.5 μs for systems N = 40−100 and 3.0 μs for systems N = 105−150. Equilibrium was determined by the time series analysis of the convergence of the ergodic measure based on the time averaged number of clusters per system (see Figure S3 in the Supporting Information).

For the purposes of data analysis, a micelle is defined as a cluster with more than 10 DPC. Qualitatively, the plot of the number of micelles ($n_{micelles}$) as a function of the total number of surfactants (N) in the systems reveals the finite-size behavior previously proposed by Kindt.[23] We see characteristic plateaus associated with integer numbers of micelles for smaller N systems and convergence to an approximately linear relationship at larger N (Figure 2). Kindt demonstrated that, in the thermodynamic limit of large N, the relationship between the N and $n_{micelles}$ is linear.[23] As a way of approximating this correlation, a linear fit was obtained,
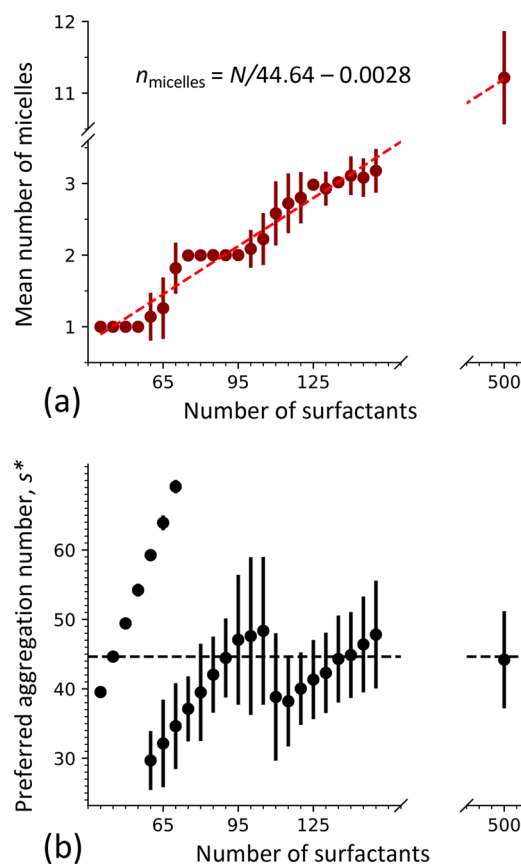
Figure 2. (a) Mean number of micelles in each system as a function of the total number of surfactant molecules in the system (N). Error bars represent the standard deviation across replicates. Linear fit to data defined as $n_{micelles} = N/44.64 - 0.0028$, which corresponds to s^{thermo} = 44.64 and c = −0.0028 (eq 10). (b) Preferred aggregation number, s^*, for each system as a function of total number of surfactants (N). The s^* values were obtained by Tanford's Gaussian approximation for micelle size distributions (eq 3). The horizontal line corresponds to s^{thermo} = 44.64. System sizes N = 60, 65, and 70 have two s^* values due to their bimodal size distributions. Error bars represent the standard deviation of the fitted normal distributions to the micelle size distributions.

$$n_{micelles} = \frac{N}{s^{thermo}} + c \tag{10}$$

where the slope, $1/s^{thermo}$, corresponds to the number of micelles per surfactant in the thermodynamic limit, and the inverse of the slope represents the number of surfactants per micelle at this limit. The number of surfactants per micelle in the thermodynamic limit is a special aggregation number, s^{thermo}, which was found to be 44.64, with a small correction constant, c = −0.0028, which was observed due to unsampled data at additional system sizes. The constant c should approach zero as samples from more system sizes are added to the linear fit. As expected, systems N = 45, 90, and 135, which contain multiples of s^{thermo}, fall on the line defined by eq 10 (Figure 2). We therefore suggest that these system sizes are less impacted by finite-size effects because they contain appropriate numbers of surfactants to form 1, 2, and 3 micelles of size s^{thermo}. The relationship between each s^* micelle size and the total number of surfactants is also in agreement with previous predictions by Kindt.[23] The plot of the s^* values presents a damped oscillation and converges to the $s^{thermo} \approx 45$ number of surfactants (Figure 2).
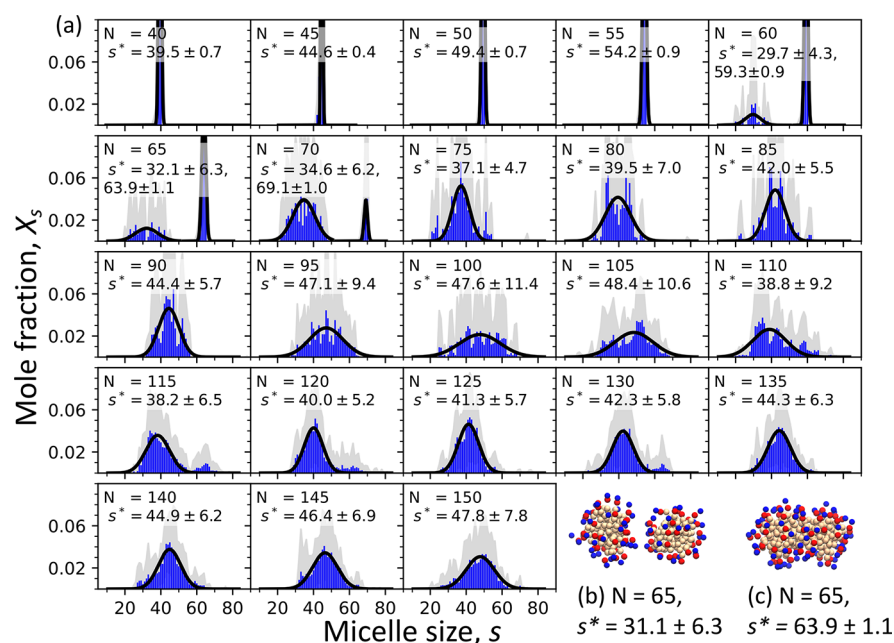
Figure 3. (a) Micelle size distributions for each system (blue bars) measured as the mole fraction X_s of aggregates of size s with standard deviation across replicates (gray fill). Fits to a Gaussian distribution centered at s^* defined by Tanford (black line). s^* values reported with standard deviation of the normal distribution. Snapshots depicting equilibrium micelle ensembles from systems (b) $N = 65$, small s^*, and (c) $N = 65$, large s^*.

**Table 2. Summary of s^* Values Calculated by Gaussian Fitting to X_s and Quadratic Fitting to $\ln(X_s)$**

	N						
	90	95	100	105	110	115	120
s^* (Gaussian)	44.4 ± 5.7	47.1 ± 9.4	47.6 ± 11.4	48.4 ± 10.6	38.8 ± 9.2	38.2 ± 6.5	40.0 ± 5.2
s^* (quadratic)	45.1 ± 4.9	47.5 ± 1.9	48.0 ± 1.6	46.5 ± 0.9	46.2 ± 0.9	45.5 ± 1.5	46.7 ± 1.7
	N						
	125	130	135	140	145	150	500
s^* (Gaussian)	41.3 ± 5.7	42.3 ± 5.8	44.3 ± 6.3	44.9 ± 6.2	46.4 ± 6.9	47.8 ± 7.8	42 ± 7.0
s^* (quadratic)	43.1 ± 1.8	45.6 ± 1.1	44.9 ± 1.7	45.0 ± 1.0	46.9 ± 1.9	45.7 ± 0.8	45.1 ± 1.5

In order to derive the preferred aggregation number, s^*, for each system size from the equilibrium micelle distribution, we consider Tanford's Gaussian approximation for modeling micelle size distributions (eq 3).[26] After fitting the Gaussian function to selected regions of the computed mole fraction distribution of each system size, the finite-size effect is again apparent (Figure 3). Systems of 60, 65, and 70 DPC exhibit unique, bimodal size distributions due to their relatively small finite-size and, consequently, have two s^* values. The s^* values for each system are summarized in Table 2. A notable result in the large N region is that, past $N = 135$, the s^* value increases. When $N = 135$, $s^* = 44.3$. If we assume that the previously calculated thermodynamic limit s^* value of 44.6 is accurate, we find that increasing the total number of surfactants from $N = 135$ to $N = 150$ brings us further away from our expected result. When $N = 150$, the Gaussian fit leads to an s^* value of 47.8. Therefore, even at the system size of $N = 150$, there remains, although diminished, an artifact of the finite-size effect. After applying the same Gaussian fitting to the $N = 500$ system described in the previous section, which had the same concentration and solvent model as the finite-size effect simulations, we obtain $s^* = 44.2 \pm 7.0$ (see Figure S1 in the Supporting Information). In this large N system, we find that the preferred aggregation number is similar to that observed in the

smaller $N = 135$ system, which further validates our method of calculating s^* values using relatively small system sizes.

An alternative way of computing s^* from the finite-size simulations makes use of Tanford's fundamental micelle equation (eq 2).[26] The value of s which maximizes the equation corresponds to the s^* value of the size distribution. As X_s can be modeled as a Gaussian function, $\ln(X_s)$ can be modeled with a quadratic function that can be used to identify s^* for each system size. The quadratic function was successful for fitting to the medium and large N systems but did not work well for small N systems due to both nongaussian and bimodal size distributions (Figure 4). The obtained s^* values for each system are summarized in Table 2 and plotted as a function of system size (see Figure S9 in the Supporting Information). It is notable that the method of quadratic fitting (Figure 4, Table 2) is less sensitive to finite-size effects compared to the Gaussian fitting (Figure 3). Therefore, for future studies which aim to derive micelle aggregation numbers from small surfactant systems, we posit that it is preferable to use the quadratic fitting method.

To further evaluate the thermodynamics of micellization as observed in the finite-size simulations, we can invoke Tanford's micelle equation (eq 2) to derive ΔG_s° from the distribution of the mole fraction, X_s.[26] If we consider each micelle of size s as being in equilibrium with monomers in each system, we expect that
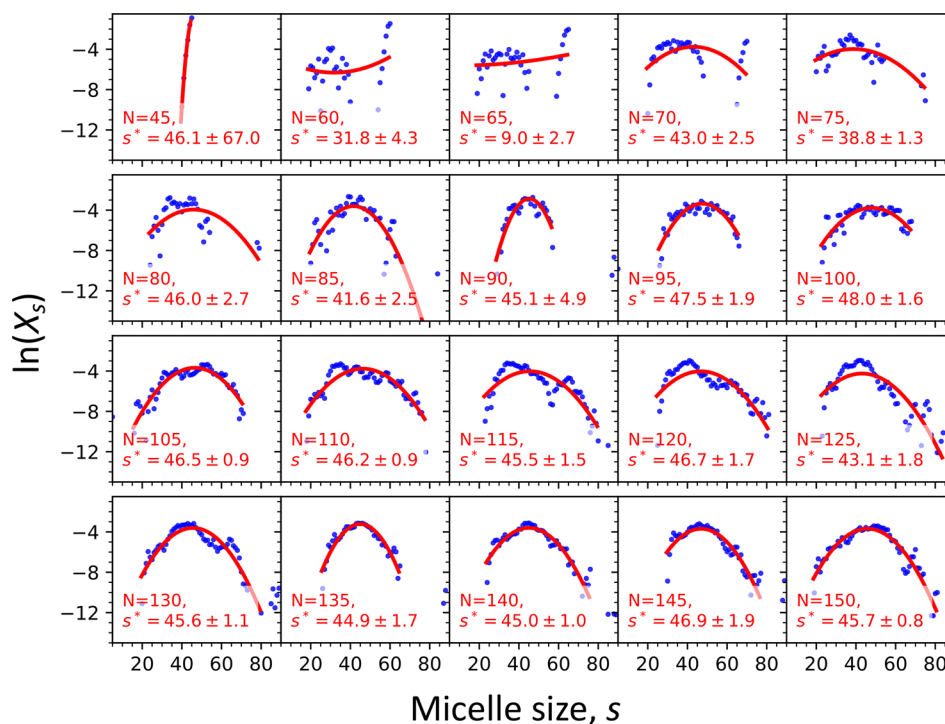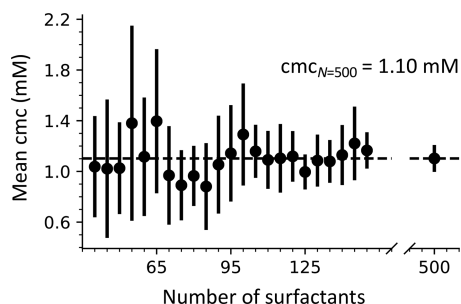
Figure 4. Natural logarithm of the mole fraction of each cluster size for each finite-size effect system micelle size distribution. Tanford's fundamental micelle equation (eq 2)[26] is applied to derive the $s*$ value for each system size. The $s*$ value is the maximum of the quadratic fit to the points.

$$\Delta G_s^{\circ} = -\frac{RT}{s}\ln(K_s) \qquad (11)$$

where $K_s = X_s/sX_1^s$ is the equilibrium constant for the formation of a micelle with s members from s surfactants. The value of ΔG_s° corresponds to the free energy change[26] contributed by each of s monomers in the formation of a micelle of size s. The resulting ΔG_s° values are on the order of -1 kJ/mol (see Figure S4 in the Supporting Information). In the bimodal region of the free energy plots, $N = 60$, 65, and 70, we see an interesting trend in the ΔG_s° values, which have double wells around two separate preferred aggregation numbers. The finite-size effect on the free energy of micellization is especially apparent for these systems.

A final characterization of the finite-size effect on the $N = 40-150$ systems concerns the critical micelle concentration (cmc) (Figure 5). The cmc was approximated as the concentration of surfactants which are monomers or are associated with clusters

of 10 or fewer members.[23] The behavior of the cmc as a function of the total number of surfactants, N, is in agreement with the predictions of Kindt.[23] We observed an oscillation of the cmc values about the large $N = 500$ value of 1.10 mM (Figure 5). This cmc is in good agreement with Lauterwein et al.'s experimentally derived value of 1.1 mM for DPC at 295 K from their ultracentrifugation and light scattering study.[1]

3.C. Back-Mapping MARTINI Representations to CHARMM36 All-Atom Representations Leads to Decrease in Micelle Volume. The finite-size effects on self-assembly could possibly be avoided by implementing a multiscale approach: first simulating a large, computationally affordable MARTINI system, and then back-mapping the equilibrium configuration to an all-atom representation for better comparison with experimental data. In the interest of testing the applicability of this method, the DPC positions at 5200 ns from the large $N = 500$ DPC MARTINI W systems in Section 3.A were back-mapped to CHARMM36 AA representations using the backward.py method.[28] The new AA systems were simulated for 20 ns until the time series of the mean radius of gyration (R_g) was observed to converge (see Figure S6 in the Supporting Information). The systems were analyzed to produce a micelle size distribution, SANS profile, and pair distance distribution function (Figure 5). The first nanosecond was omitted from these calculations because there was a substantial collapse in R_g after back-mapping from MARTINI to CHARMM36 (see Figure S6 in the Supporting Information).

The radius of gyration of the AA micelles was found to be 15.60 Å, in contrast with the MARTINI radius of gyration of 16.37 Å. The decrease in R_g after the back-mapping was performed is evidence, however, that the MARTINI surfactants occupy a larger volume within a smaller aggregation number, compared to the AA model. Our results with the MARTINI model are similar to the findings of Abel et al., who reported an R_g of 16.8 Å for an AA DPC micelle of size $s = 54$ with

Figure 5. Mean critical micelle concentration (cmc) in mM for each system size as a function of the total number of surfactants, N. The dashed line corresponds to cmc = 1.10 mM, which is the large $N = 500$ system mean cmc value. The error bars represent the standard deviation across the replicates.

CHARMM36.[12] In addition, the all-atom simulations of Faramarzi et al. yielded a similar R_g.[51] The preferred aggregation number of the MARTINI micelles, $s* = 44.2 \pm 7.0$ (see Figure S1 in the Supporting Information), is considerably smaller than Abel's micelle. This $s*$ value, however, is in agreement with the MARTINI self-assembly simulations of Sanders et al., which resulted in a micelle size distribution centered around 45 surfactants.[10] This insight explains why the radius of gyration decreases after back-mapping and also provides evidence that the MARTINI micelles reliably predict the preferred volume but underestimate the preferred aggregation number because the DPC molecule itself is too large in this model. This finding is remarkable considering the major computational improvements afforded by the MARTINI model (Figure 6), which sacrifices accuracy in aggregation number but does not compromise the expected R_g of the micelles.
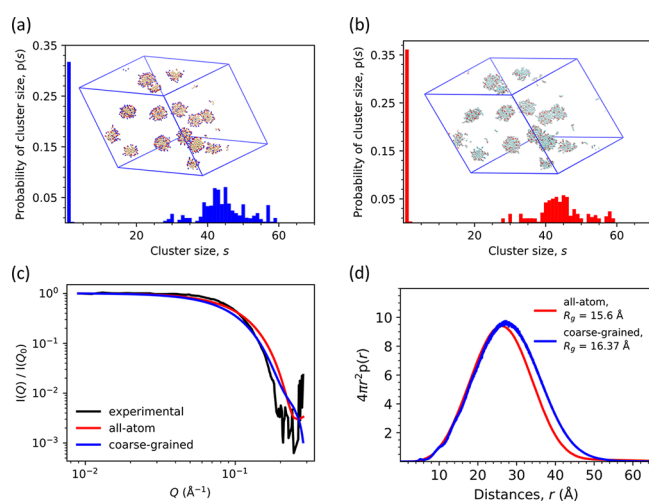
Figure 6. (a) Cluster size distribution of MARTINI W system, averaged over the last 20 ns of the 6 replicates. Superimposed is the first frame before back-mapping. (b) Cluster size distribution of the back-mapped AA system, averaged over the first 20 ns, omitting the first nanosecond. Superimposed is the first frame after back-mapping. (c) SANS profiles constructed using the Debye scattering equaion (eq 4), for the AA (red), MARTINI (blue), and experimental[5] (black) systems. (d) Distance distribution function, $4\pi r^2 p(r)$, and corresponding radius of gyration calculated using the second moment of the distribution (eq 5) for the equilibrium ensemble of the MARTINI (blue) and first 20 ns, omitting the first nanosecond, of the AA (red) systems.

At the concentration of 100 mM (used throughout our simulations), Pambou et al. found a radius of gyration of 21.45 Å, based on the reported core radius of 19.6 Å and shell radius of 7.8 Å,[5] and assuming spherical micelles for the purpose of approximation. In addition, the SAXS-derived R_g for DPC micelles was found by Lipfert et al. to be 34.5 ± 0.08 Å at 77 mM surfactant concentration.[6] As noted by Faramarzi et al., in the case of SAXS-derived R_g, the hydration shell of the experimental

micelle results in a larger value than that predicted by MD.[51] For SANS, however, we have worked under the assumption that the "water shelling" contribution is removed by selective deuteration of the surfactant.[5] In the ultracentrifugation study, Lauterwein et al. found a diameter of 47 Å at 20 mM, which converts to $R_g = 18.2$ Å under the spherical assumption.[1] Similarly, the NMR studies of Kallick et al. reported a hydrodynamic radius of 18.65 ± 0.3 Å at 228 mM.[4] We see much better agreement between the simulation-derived R_g values and the ultracentrifugation and NMR experimentally derived values. These variations in R_g are summarized in Table 3.

Considering the experimental and simulated SANS profiles (Figure 5), it is qualitatively evident that the experimental (black) and CG micelles (blue) have larger volumes, and the AA micelles (red) have smaller volumes. Furthermore, Pambou et al. reported an aggregation number of 70.6 ± 5 surfactants.[5] This aggregation number is not only larger than the value derived from the MARTINI simulations but also larger than Lauterwein et al.'s ultracentrifugation and DLS results[1] of 56 ± 5 and Kallick et al.'s NMR results[4] of 44 ± 5 surfactants. These variations in aggregation number are summarized in Table 3. The bulkiness of the MARTINI representations does not allow for an accurate number of surfactants to aggregate into a preferred micelle size. Due to this limitation of the MARTINI model, the large N, long time scale MARTINI simulations, and subsequent back-mapping to AA, while effective in avoiding finite-size effects, is not adequate for reproducing experimental micelle results.

4. CONCLUSION

We performed molecular dynamics simulations of DPC micelle self-assembly to equilibrium at 100 mM surfactant concentration using four different MARTINI 2 solvent models and evaluated the equilibrium size distributions by direct comparison with an experimental SANS spectrum. We determined that the widely used MARTINI 2 nonpolarizable water produces SANS spectra most similar to those from experiment. We studied finite-size effects on DPC micelle simulations in 40−150 DPC systems. We observed damped oscillations in both the number of micelles and in the preferred aggregation number ($s*$) of each system as a function of the number of surfactants in the system about the value of $s*$ in the thermodynamic limit. We observe that these damped oscillations mostly converge to s^{thermo} once the system size exceeds 3 times the value of s^{thermo}. This observation suggests that reasonably accurate micelle simulations may be performed by using three or more times a guessed s^{thermo} number of surfactants in a simulation, assuming that this value can be estimated. In addition to damped oscillations in micelle number and size as a function of system size, we also observed bimodal distributions of micelle sizes in systems of 60, 65, and 70 DPC. We note that a dodecahedral unit cell was used in our simulations. In practice, the exact nature of the finite-size effects may be influenced by the specific choice of PBC definitions, as well as the treatment of long-range interactions.

Table 3. Summary of $s*$ Values Calculated by Gaussian Fitting to X_s and Quadratic Fitting to $\ln(X_s)$

method	MARTINI[a]	CHARMM36[b]	SANS[c]	SAXS[d]	ultracentrifugation[e]	NMR[f]
preferred aggregation number, $s*$	44.2 ± 7	54	70.6 ± 5	68−80	56 ± 5	44 ± 5
radius of gyration, R_g (Å)	16.37	16.8	21.45	34.5 ± 0.08	18.2	18.65 ± 0.3

[a]100 mM DPC, from $N = 500$ MARTINI system. [b]200 mM from Abel et al.[12] [c]100 mM DPC, from Pambou et al.,[5] R_g based on spherical micelle assumption. [d]77 mM DPC, from Lipfert et al.[6] [e]20 mM DPC, from Lauterwein et al.[1] [f]228 mM DPC, from Kallick et al., reported hydrodynamic radius.[4]

As a way of testing the accuracy of a multiscale approach to micelle self-assembly, large 500 DPC MARTINI micelle configurations at equilibrium were back-mapped to AA CHARMM36 representations. The aggregation numbers, radii of gyration, and SANS profiles were compared between models and with experiment. It was found that the MARTINI model is reliable for achieving the proper radius of gyration but, due to the large volume of each MARTINI DPC molecule, underestimates the preferred aggregation number. Nevertheless, the simulated aggregation number and radius of gyration values are generally consistent with the range of experimental observations considered in this study.

These results provide a clear prescription for the accurate modeling of DPC micelle self-assembly in terms of the choice of model, minimum system size, and methods of analysis for comparison with experiment. The trade-offs between models and the feasibility of a multiscale approach to the problem of determining equilibrium micelle size distributions are elucidated. It is clear that the MARTINI model is accurate for determining the volume of DPC micelles and was reliable for fully assessing the finite-size effect and deriving the thermodynamic aggregation number for a generic surfactant. For determining a more accurate aggregation number, however, an AA model must be used, at the expense of longer computational time. We posit that the methods described here for determining the minimum system size for MARTINI DPC can be repeated and generalized for a wide variety of surfactants which may be of research interest.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcb.1c01186.

> Figures of micelle size distributions, Tanford Gaussian approximation, ergodic measure, and relaxation times for the four $N = 500$ MARTINI systems; figure of the ergodic measure for all finite-size effect systems; figures of free energy change for selected finite-size effect systems and Tanford's micelle equation for the $N = 500$ MARTINI W system; and figure of time series of the mean radius of gyration for back-mapped CG to AA systems (PDF)

AUTHOR INFORMATION

Corresponding Author

John E. Straub − Department of Chemistry, Boston University, Boston, Massachusetts 02215, United States; orcid.org/0000-0002-2355-3316; Email: straub@bu.edu

Authors

Jonathan J. Harris − Department of Chemistry, Boston University, Boston, Massachusetts 02215, United States
George A. Pantelopulos − Department of Chemistry, Boston University, Boston, Massachusetts 02215, United States; orcid.org/0000-0002-4373-1677

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jpcb.1c01186

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Lauterwein, J.; Bösch, C.; Brown, L. R.; Wüthrich, K. Physicochemical studies of the protein-lipid interactions in melittin-containing micelles. *Biochim. Biophys. Acta, Biomembr.* **1979**, *556*, 244−264.

(2) Gao, X.; Wong, T. C. Studies of the binding and structure of adrenocorticotropin peptides in membrane mimics by NMR spectroscopy and pulsed-field gradient diffusion. *Biophys. J.* **1998**, *74*, 1871−1888.

(3) Göbl, C.; Dulle, M.; Hohlweg, W.; Grossauer, J.; Falsone, S. F.; Glatter, O.; Zangger, K. Influence of phosphocholine alkyl Chain length on peptide-micelle interactions and micellar size and shape. *J. Phys. Chem. B* **2010**, *114*, 4717−4724.

(4) Kallick, D. A.; Tessmer, M. R.; Watts, C. R.; Li, C.-Y. The Use of Dodecylphosphocholine Micelles in Solution NMR. *J. Magn. Reson., Ser. B* **1995**, *109*, 60−65.

(5) Pambou, E.; Crewe, J.; Yaseen, M.; Padia, F. N.; Rogers, S.; Wang, D.; Xu, H.; Lu, J. R. Structural Features of Micelles of Zwitterionic Dodecyl-phosphocholine (C_{12}PC) Surfactants Studied by Small-Angle Neutron Scattering. *Langmuir* **2015**, *31*, 9781−9789.

(6) Lipfert, J.; Columbus, L.; Chu, V. B.; Lesley, S. A.; Doniach, S. Size and shape of detergent micelles determined by small-angle X-ray scattering. *J. Phys. Chem. B* **2007**, *111*, 12427−12438.

(7) Oliver, R. C.; Lipfert, J.; Fox, D. A.; Lo, R. H.; Doniach, S.; Columbus, L. Dependence of Micelle Size and Shape on Detergent Alkyl Chain Length and Head Group. *PLoS One* **2013**, *8*, e62488.

(8) Marrink, S. J.; Tieleman, D. P.; Mark, A. E. Molecular Dynamics Simulation of the Kinetics of Spontaneous Micelle Formation. *J. Phys. Chem. B* **2000**, *104*, 12165−12173.

(9) Marrink, S. J.; de Vries, A. H.; Mark, A. E. Coarse Grained Model for Semiquantitative Lipid Simulations. *J. Phys. Chem. B* **2004**, *108*, 750−760.

(10) Sanders, S. A.; Panagiotopoulos, A. Z. Micellization behavior of coarse grained surfactant models. *J. Chem. Phys.* **2010**, *132*, 114902.

(11) Urano, R.; Pantelopulos, G. A.; Song, S.; Straub, J. E. Characterization of dynamics and mechanism in the self-assembly of AOT reverse micelles. *J. Chem. Phys.* **2018**, *149*, 144901.

(12) Abel, S.; Dupradeau, F. Y.; Marchi, M. Molecular dynamics simulations of a characteristic DPC micelle in water. *J. Chem. Theory Comput.* **2012**, *8*, 4610−4623.

(13) Urano, R.; Pantelopulos, G. A.; Straub, J. E. Aerosol-OT Surfactant Forms Stable Reverse Micelles in Apolar Solvent in the Absence of Water. *J. Phys. Chem. B* **2019**, *123*, 2546−2557.

(14) Yoshii, N.; et al. Free energy change of micelle formation for sodium dodecyl sulfate from a dispersed state in solution to complete micelles along its aggregation pathways evaluated by chemical species model combined with molecular dynamics calculations. *Wuli Huaxue Xuebao* **2018**, *34*, 1163−1170.

(15) Takeda, K.; Fujimoto, K.; Yoshii, N.; Okazaki, S. Molecular dynamics study of solubilization of cyclohexane, benzene, and phenol into mixed micelles composed of sodium dodecyl sulfate and octaethylene glycol monododecyl ether. *J. Comput. Chem.* **2019**, *40*, 2722−2792.

(16) Zhang, X.; Arce Nunez, J. G.; Kindt, J. T. Derivation of micelle size-dependent free energies of aggregation for octyl phosphocholine from molecular dynamics simulation. *Fluid Phase Equilib.* **2019**, *485*, 83−93.

(17) Lazaridis, T.; Mallik, B.; Chen, Y. Implicit solvent simulations of DPC micelle formation. *J. Phys. Chem. B* **2005**, *109*, 15098−15106.

(18) Mori, T.; Sugita, Y. Implicit Micelle Model for Membrane Proteins Using Superellipsoid Approximation. *J. Chem. Theory Comput.* **2020**, *16*, 711−724.

(19) Chen, H.; Panagiotopoulos, A. Z. Molecular Modeling of Surfactant Micellization Using Solvent-Accessible Surface Area. *Langmuir* **2019**, *35*, 2443−2450.

(20) Pantelopulos, G. A.; Nagai, T.; Bandara, A.; Panahi, A.; Straub, J. E. Critical size dependence of domain formation observed in coarse-grained simulations of bilayers composed of ternary lipid mixtures. *J. Chem. Phys.* **2017**, *147*, 095101.

(21) Huang, J.; Feigenson, G. W. Monte Carlo simulation of lipid mixtures: finding phase separation. *Biophys. J.* **1993**, *65*, 1788−1794.

(22) Johnston, M. A.; Swope, W. C.; Jordan, K. E.; Warren, P. B.; Noro, M. G.; Bray, D. J.; Anderson, R. L. Toward a Standard Protocol for Micelle Simulation. *J. Phys. Chem. B* **2016**, *120*, 6337−6351.

(23) Kindt, J. T. Accounting for finite-number effects on cluster size distributions in simulations of equilibrium aggregation. *J. Chem. Theory Comput.* **2013**, *9*, 147−152.

(24) Zhang, X.; Patel, L. A.; Beckwith, O.; Schneider, R.; Weeden, C. J.; Kindt, J. T. Extracting Aggregation Free Energies of Mixed Clusters from Simulations of Small Systems: Application to Ionic Surfactant Micelles. *J. Chem. Theory Comput.* **2017**, *13*, 5195−5206.

(25) Tanford, C. *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*; John Wiley & Sons, Inc.: New York, 1973; p 43.

(26) Tanford, C. Theory of micelle formation in aqueous solutions. *J. Phys. Chem.* **1974**, *78*, 2469−2479.

(27) Palmer, B. J.; Liu, J. Simulations of micelle self-assembly in surfactant solutions. *Langmuir* **1996**, *12*, 746−753.

(28) Wassenaar, T. A.; Pluhackova, K.; Böckmann, R. A.; Marrink, S. J.; Tieleman, D. P. Going Backward: A Flexible Geometric Approach to Reverse Transformation from Coarse Grained to Atomistic Models. *J. Chem. Theory Comput.* **2014**, *10*, 676−690.

(29) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; De Vries, A. H. The MARTINI force field: Coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **2007**, *111*, 7812−7824.

(30) Yesylevskyy, S. O.; Schäfer, L. V.; Sengupta, D.; Marrink, S. J. Polarizable water model for the coarse-grained MARTINI force field. *PLoS Comput. Biol.* **2010**, *6*, e1000810.

(31) Michalowsky, J.; Schäfer, L. V.; Holm, C.; Smiatek, J. A refined polarizable water model for the coarse-grained MARTINI force field with long-range electrostatic interactions. *J. Chem. Phys.* **2017**, *146*, 054501.

(32) Michalowsky, J.; Zeman, J.; Holm, C.; Smiatek, J. A polarizable MARTINI model for monovalent ions in aqueous solution. *J. Chem. Phys.* **2018**, *149*, 163319.

(33) 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 super-computers. *SoftwareX* **2015**, *1−2*, 19−25.

(34) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; Van Der Spoel, D.; et al. GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29*, 845−854.

(35) Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D.; Pastor, R. W. Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *J. Phys. Chem. B* **2010**, *114*, 7830−7843.

(36) Müllner, D. Modern Hierarchical, Agglomerative Clustering Algorithms. *arXiv.org* **2011**, arXiv:1109.2378.

(37) Gowers, R.; Linke, M.; Barnoud, J.; Reddy, T.; Melo, M.; Seyler, S.; Domański, J.; Dotson, D.; Buchoux, S.; Kenney, I.; et al. MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. *Proceedings of the 15th Python in Science Conference* **2016**, 98−105.

(38) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **2011**, *32*, 2319−2327.

(39) Humphrey, W.; Dalke, A.; Schulten, K. VMD - Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.

(40) Thirumalai, D.; Mountain, R. D.; Kirkpatrick, T. R. Ergodic behavior in supercooled liquids and in glasses. *Phys. Rev. A: At., Mol., Opt. Phys.* **1989**, *39*, 3563−3574.

(41) Thirumalai, D.; Mountain, R. D. Ergodic convergence properties of supercooled liquids and glasses. *Phys. Rev. A: At., Mol., Opt. Phys.* **1990**, *42*, 4574−4587.

(42) Straub, J. E.; Thirumalai, D. Exploring the energy landscape in proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 809−813.

(43) Straub, J. E.; Choi, J. K. Extracting the energy barrier distribution of a disordered system from the instantaneous normal mode density of states: Applications to peptides and proteins. *J. Phys. Chem.* **1994**, *98*, 10978−10987.

(44) Straub, J. E.; Rashkin, A. B.; Thirumalai, D. Dynamics in Rugged Energy Landscapes with Applications to the S-Peptide and Ribonuclease A. *J. Am. Chem. Soc.* **1994**, *116*, 2049−2063.

(45) Pedersen, J. S. Analysis of small-angle scattering data from colloids and polymer solutions: Modeling and least-squares fitting. *Adv. Colloid Interface Sci.* **1997**, *70*, 171−210.

(46) Bacon, G. E.; Lonsdale, K. Neutron diffraction. *Rep. Prog. Phys.* **1953**, *16*, 1−61.

(47) Dinnebier, R. E.; Billinge, S. J. L. *International Tables for Crystallography*; American Cancer Society, 2018; Chapter 1.1, pp 2−23.

(48) Mitchell, M.; Muftakhidinov, B.; Winchen, T. *Engauge Digitizer Software*.

(49) Hammouda, B. *Probing Nanoscale Structure—The SANS Toolbox*; NIST; pp 544−545.

(50) Paradies, H. H. Shape and size of a nonionic surfactant micelle. Triton X-100 in aqueous solution. *J. Phys. Chem.* **1980**, *84*, 599−607.

(51) Faramarzi, S.; Mertz, B.; Bonnett, B.; Scaggs, C. A.; Hoffmaster, A.; Grodi, D.; Harvey, E. Molecular Dynamics Simulations as a Tool for Accurate Determination of Surfactant Micelle Properties. *Langmuir* **2017**, *33*, 9934−9943.