Simulation of diblock copolymer surfactants. I. Micelle free energies

Joshua A. Mysona, Alon V. McCormick, and David C. Morse

Phys. Rev. E **100**, 012602 — Published 8 July 2019

DOI: 10.1103/PhysRevE.100.012602

,

Simulation of Diblock Copolymer Surfactants. 1. Micelle Free Energies

Joshua A. Mysona, Alon V. McCormick, and David C. Morse*
Department of Chemical Engineering and Materials Science,
University of Minnesota, 421 Washington Ave. SE, Minneapolis, MN 55455, USA

Semi-grand hybrid Monte Carlo (MC) simulations are used to measure equilibrium properties of micelles formed in a simple bead-spring model of asymmetric A-B diblock copolymer surfactant molecules in an A homopolymer solvent, over a range of values of surfactant solubility. Simulations are used to accurately measure the free energy of formation of micellar clusters as a function of aggregation number over a wide range of values, and to characterize the crossover from spherical to rod-like micelle shape with increasing aggregation number. Dynamical properties of the same model are considered in an accompanying paper.

I. INTRODUCTION

This work and a companion article [1] present a detailed study of both equilibrium and dynamical properties of a simple simulation model of micellar solutions of non-ionic surfactants. Both papers consider behavior of a mixture of asymmetric AB diblock copolymers dissolved in a liquid of A homopolymers, in which the copolymers form spherical micelles. Equilibrium properties such as micelle free energies are discussed in this paper, while the companion article discusses dynamical properties. The work reported in both papers was originally motivated by an interest in modelling rare dynamical processs in micellar solutions. An understanding of equilibrium properties is, however, a prerequisite to the quantitative study of dynamics. When attempting to model rare processes that create and destroy entire micelles, it is necessary to know how micelle free energies depend on aggregation number. Specifically, one needs accurate estimates of free energies for micelles with aggregation numbers that are rare in equilibrium but play a critical role in dynamics [2–5]. This need motivated the development of improved methods for computing micelle free energies, which are reported here along with our results for a particular simulation model.

Molecular dynamics (MD) simulations of surfactant micelles are complicated by the fact that equilibration of the total number of micelles in a solution can be extremely slow [6, 7]. Exchange of surfactant molecules between micelles can occur comparatively rapidly via expulsion and reinsertion of individual molecules, as can repartitioning of material between micelles and free molecules [8, 9]. These single molecule processes become infrequent in systems of very sparingly soluble surfactants, but can be observed in simulations of systems with more soluble surfactants [10–12]. Processes that create or destroy entire micelles are much less frequent, but are required in order to equilibrate the number concentration of micelles in a solution [2–4, 6, 13–15]. In any MD simulation of a closed system, with a fixed number of surfactant molecules, the system must also be large enough to contain many micelles to avoid otherwise severe finite size effects. This combination of slow relaxation and finite size effects make it difficult or impossible to accurately determine the equilibrium average micelle aggregation number and the critical micelle concentration by straightforward MD simulations [16]. Consequently, many MD and dissipative particle dynamics (DPD) simulations have instead focused on properties that equilibrate more rapidly, such as chain conformations, counterion distribution, and changes in micelle shape with changes in aggregation number [11, 12, 17–20]. Some studies have attempted to identify the CMC and equilibrium aggregation number by MD or DPD simulations of simple coarse-grained models with relatively soluble surfactants [11, 21–25]. Few of these, however, have demonstrated that equilibrium has been reached, which requires that the simulation be long enough for micelles to be repeatedly created and destroyed after the system appears to have equilibrated.

Many limitations of MD simulations can be avoided by the use of biased Monte Carlo (MC) simulation techniques in an open ensemble. Such simulations use nonphysical moves that allow the system composition to fluctuate. Panagiotopoulos and coworkers have emphasized how otherwise severe finite-size effects can be minimized by the use of Monte Carlo simulations in an open ensemble [26–28]. Bolhuis *et al.* have also used a similar approach using free energy methods to determine the micelle formation energy as a function of aggregation number [29].

In the current work, we perform a biased MC simulation in a semi-grand ensemble using a technique similar to one used previously by Cavallo, Müller and Binder to treat a somewhat specialized class of polymeric systems [30]. The technique used both here and in Ref. [30] can only be applied solutions of AB diblock copolymer dissolved in an A homopolymer in which homopolymer and copolymers contain the same number of beads. In this special case, the algorithm provides extremely fast sampling of micelle aggregation number in this special case. The simulation is performed in a semi-grand ensemble in

* Corresponding author, email: morse012@umn.edu

which the number of copolymer molecules can fluctuate, using a MC move that can transform a copolymer into a homopolymer, or vice versa. The combination of the use of an open ensemble and a bias potential that creates a nearly uniform distribution of micelle aggregation numbers allows micelles to be created and destroyed relatively rapidly, and allows us to measure how the free energy of formation of a micelle depends on aggregation number. Cavallo *et al.* simulated a lattice model and considered a sequence of systems containing chains with different values for number of beads, N, but equal values of χN, in order to compare to self-consistent field theory predictions with increasing N. Here, we use a continuum bead-spring model to simulate chains with a fixed value of $N = 32$ over a range of values of the effective χ parameter, in order to study the effects of increasing AB repulsion (or decreasing solubility) upon micelle thermodynamics. The present work also introduces several technical improvements in the methods used to compute how the free energy required to create a micelle or cluster depends on aggregation number. These improvements were introduced in order to obtain reliable results for the free energy of very small and very large micelles that are rare in equilibrium but that play an important role in the dynamical processes considered in the accompanying paper.

The remainder of this paper is organized as follows. Sec. II contains details of the simulation system that is considered in both this paper and the accompanying paper. Sec. III describes the details of our Monte Carlo simulations. Sec. IV discusses the analysis and results for several properties of a hypothetical semi-grand canonical ensemble, in which the state of the system is specified by the value of an exchange chemical potential. Sec. V presents our analysis of semi-grand canonical formation free energies for clusters of varying aggregation number. Sec. VI presents an anaysis of how the shape of a micelle changes with changing aggregation number, which shows the existence of a crossover from spherical to rod-like morphology. Conclusions are summarized in VII.

II. SIMULATION MODEL

The simulations presented here all use a simple bead-spring model of a system containing a highly asymmetric AB diblock copolymer surfactant in an A homopolymer solvent. In the model used here, the copolymer is a chain of 32 beads containing a block of 28 A beads and 4 B beads, in which the minority B block forms the micelle core. Each homopolymer "solvent" molecule is a chain of 32 A beads.

Interactions are controlled by a potential energy model similar to one used previously by our group in studies of block copolymer melts [31–35]. Neighboring beads within each chain are attracted to one another by a harmonic bond potential of the form

$$U_{\mathrm{bond}}(r) = \frac{1}{2}\kappa r^2 \tag{1}$$

where r is the distance between beads and κ is a spring constant. All beads interact through a repulsive non-bonded pair potential of the form

$$U_{\mathrm{pair}}(r) = \frac{1}{2}\epsilon_{ij}\left(1 - \frac{r}{\sigma}\right)^2 \tag{2}$$

for any pair of beads of types i and j that are separated by a distance $r < \sigma$, with $U_{\mathrm{pair}} = 0$ for $r > \sigma$, where σ is the range of the non-bonded interaction. The parameter ϵ_{ij} controls the strength of repulsion between i and j beads, with $\epsilon_{ji} = \epsilon_{ij}$. All simulations presented here use a spring constant $\kappa = 3.048 k_B T/\sigma^2$ and a non-bonded repulsion $\epsilon_{AA} = \epsilon_{BB} = 25kT$ for interactions between beads of the same type.

The repulsion between A and B beads is taken to exceed ϵ_{AA} by an amount controlled by a parameter

$$\alpha = (\epsilon_{AB} - \epsilon_{AA})/k_B T \quad , \tag{3}$$

which we vary to control copolymer solubility and preferred micelle size. Extensive simulations have been performed at four values of $\alpha = 10$, 12, 14, and 16. Systems with α significantly less than 10 do not exhibit stable micelles.

The "solvent" in this model was chosen to be a polymer with the same number of beads as the copolymer in order to allow the use of efficient semi-grand ensemble sampling techniques [30]. The copolymer was taken to be highly asymmetric polymer, with a core-forming B block much shorter than the corona-forming A block, to favor the formation of spherical rather than wormlike micelles in equilibrium.

All simulations are performed in the NPT ensemble at a constant pressure $P = 20.249 k_B T/\sigma^3$. This results in a bead concentration of approximately $3.0\sigma^{-3}$ for long chains [34]. Monte Carlo simulations were used to measure equilibrium properties, such as the cluster formation free energy, which are reported in this paper. Molecular dynamics simulations of the same model were used to measure dynamical properties, which are reported in [1]. All simulations presented here were performed using the open-source Simpatico simulation package, which was developed in our research group. Source code is available via the github repository, at github.com/dmorse/simpatico.

When analyzing simulations performed with this model, we use a cluster identification algorithm to identify both micelles and short-lived submicellar clusters. In this algorithm, two copolymer molecules are taken to be in direct "contact" if any two B (i.e., core block) monomers from those two chains are separated by a distance less than 0.8σ. Any two molecules that are in contact are assigned to the same cluster. This definition assigns every copolymer in the simulation unit cell to a

unique cluster of one or more molecules. Isolated free molecules are single-molecule clusters. When applied to a system that contains one micelle, the algorithm typically identifies one large cluster (the micelle), and a few free molecules. The algorithm also identifies some small submicellar clusters of two or more molecules that happen to have core block monomers in contact, but that do not have the long lifetimes characteristic of proper micelles.

III. EQUILIBRIUM SIMULATIONS

Equilibrium properties for this simulation model have been determined by performing semi-grand Monte Carlo simulations at constant temperature and pressure. In a semi-grand ensemble, the composition of the simulation is allowed to fluctuate via an "alchemical" transformation move that allows a diblock polymer molecule to be transformed into a homopolymer or vice versa [30]. Such a transformation move may only be efficiently used in dense liquid systems if the two species are approximately the same size and shape. We have chosen to study a system in which the surfactant and solvent molecules are chains with the same number of beads in order to allow the use of this technique.

The semi-grand MC simulations reported here were all performed on a system with a periodic cubic simulation unit cell that contains a total of 1500 copolymer and homopolymer molecules. At the pressure used here, this yields a simulation unit cell with edges of average length $L \simeq 25.2\sigma$. For the range of parameters studied here, a cell of this size was found to be large enough to prevent steric interaction of a micelle with its periodic images.

A. Monte Carlo Moves

All MC simulations presented here used two types of MC move: a hybrid MC move and a semi-grand transformation move.

In a hybrid MC move [36], a short molecular dynamics simulation is used to generate an attempted Monte Carlo move. In an NPT ensemble, each attempted hybrid MC move is generated by running a short constant enthalpy (NPH) MD simulation, using a reversible, symplectic integrator and an Anderson barostat. Initial velocities for each attempted hybrid move (i.e., each short MD simulation) are chosen at random from a Maxwell-Boltzmann distribution with the desired temperature. Attempted hybrid moves are accepted or rejected on the basis of an acceptance probability that depends on the change in system enthalpy. When a hybrid move is rejected, particle positions are reset to the values they had at the beginning of the attempted move. Because the measured change in enthalpy arises only from integration discretization errors, the algorithm yields 100% acceptance in the limit of a small MD integration time step Δt. The resulting algorithm has been proven to satisfy detailed balance in the NPT ensemble even for nonzero values of $\Delta t \neq 0$, and thus converges to a canonical NPT distribution [36]. In our work, each hybrid move consists of 250 molecular dynamics integration steps with an integration time $\Delta t = 0.005\tau_0$. Here, τ_0 is the the Lennard-Jones time unit, given by $\tau_0 = \sigma\sqrt{m_b/k_BT}$, where m_b is a bead masss that is taken to be the same for all beads.

The semi-grand canonical transformation move changes a single copolymer molecule into a homopolymer, or vice versa. An attempted transformation move is performed by choosing a chain at random and simply toggling the bead type of the four beads of that chain that form the B block of the copolymer, between type B (to obtain a copolymer) and type A (to obtain an A homopolymer), while leaving the 28 beads of the corona block unchanged. This move allows the total number of copolymer molecules in the simulation to fluctuate.

Umbrella sampling is used to obtain a nearly flat probability distribution for the number of copolymer molecules in the simulation, in order to increase the probabilities of otherwise rare states. Let N denote the total number of copolymers in the simulation at any instant. Each attempted transformation of a single molecule is accepted or rejected using an acceptance criterion that is designed to sample the equilibrium distribution e^{-U'/k_BT} produced by a modified potential

$$U' = U - V(N) \quad , \tag{4}$$

in which U is the physical potential energy of the model as a function of the bead positions and current assignment of molecule types, and $V(N)$ is an umbrella potential that depends only on N, the number of copolymer molecules. The addition of an umbrella potential $V(N)$ has no affect on the acceptance criterion for hybrid MC moves, since hybrid moves do not change N. The acceptance criterion for a transformation move is a conventional Metropolis MC criterion based on the change in the value of U' induced by a change in the monomer type of the last 4 monomers on a randomly chosen chain. Each simulation was restricted to a finite range of values of N, by prohibiting transformation moves that would yield values of N outside that range. Results from simulations performed over overlapping ranges of N were combined to obtain final results.

In the main loop of the MC algorithm, a decision is made at beginning of each step whether to initiate either a hybrid MC/MD move or a "sweep" of a pre-specified number N_{sweep} of consecutive attempted transformation moves. Within a sweep of transformation moves, each attempted transformation of a single molecule is accepted or rejected using the acceptance criterion described above. A decision whether to reject a hybrid MC move is instead made only at the end of the attempted move (i.e., at the end of a short MD simulation). In what follows, we will refer to both hybrid MC moves and sweeps of transformation moves as "composite" MC moves.

The equilibrium probability $P_{\rm sim}(N)$ of obtaining a state with a specified number of molecules in the box within a biased simulation is given by

$$P_{\rm sim}(N) \propto e^{-[G(N)-V(N)]/k_BT} \quad , \tag{5}$$

in which $G(N)$ is the Gibbs free energy for a system with N copolymers, and $V(N)$ is the biasing potential used in the simulation. Solving for $G(N)$ yields

$$G(N) = -k_BT \ln P_{\rm sim}(N) + V(N) \quad , \tag{6}$$

to within an arbitrary constant.

The umbrella potential $V(N)$ used in our final calculations of $G(N)$ was chosen so as to give a nearly flat histogram $P(N)$. This potential was determined iteratively, by running a sequence of simulations in which the estimate of $G(N)$ obtained from each simulation is used as the estimate of $V(N)$ for the next iteration [37].

B. Prohibiting Multiple Micelles

To simplify some aspects of the analysis, MC simulations were performed using an acceptance criterion that prohibits the formation of states containing more than one micelle within the simulation cell. For this purpose a "micelle" is defined to be a cluster of aggregation number greater than or equal to some cutoff value, denoted by $n_{\rm cut}$. We used a cutoff $n_{\rm cut} = 8$ in all of the simulations presented here. A cluster analysis is performed after each composite MC move, i.e., after each sweep of transformation moves or each tentatively accepted hybrid MD move. If more than one cluster of aggregation number greater than $n_{\rm cut}$ is found to be present, the entire composite move is rejected and the system is returned to the state it had before the composite move was attempted. This cluster analysis is performed only at the end of each composite move, rather than after each attempted transformation move or each time step of the hybrid move, in order to reduce the computational cost of the required cluster analysis.

The constraint described above allows sampling of states that have an arbitrary number of small clusters with $n < n_{\rm cut}$. It thus does not suppress appearance of states with multiple unimers, dimers and other small aggregates. The prohibition of states with more than one cluster of size $n > n_{\rm cut}$ was introduced in order avoid potential problems with adequately sampling states in which the system contains a large number of copolymer molecules. Because the bias potential $V(N)$ that we use to enhance the probability of rare states depends only on the total number of copolymers in the simulation cell, it does not distinguish between states with one large micelle and states with two smaller micelles. For sufficiently large values of N, states in which the system contains a single large micelle can have a higher free energy than states with two smaller micelles of comparable size. If the simulations were perfectly ergodic, this would cause large micelles to undergo rapid spontaneous fission, and cause poor sampling of the small population of very large micelles. Moreover, because fission and fusion are rare events that are not effectively accelerated by our biasing potential, it could be difficult to ergodically sample the relative equilibrium probabilities of states with one large micelle or two smaller micelles. The bias potential used here was designed to dramatically accelerate the rate of creation and destruction of a single micelle by stepwise association and dissociation. This bias allows us to adequately sample relative probabilities of states with zero or one micelle, but not relative probabilities of states with different nonzero numbers of micelles. The use of biased simulations with a prohibition on the formation of multiple micelles avoids this potential problem, and thereby allows us to reliably measure free energies for large micelles that would otherwise be susceptible to spontaneous fission.

The only effect of this rejection rule upon the thermodynamics of the system is to constrain the sampling to a subset of microstates in which the system contains no more than one micelle. The free energy $G(N)$ that is obtained from a constrained simulation is thus the constrained free energy of a model that is confined to this set of microstates, as given by a partition function that is defined as a sum over these allowed states.

IV. SEMIGRAND CANONICAL ENSEMBLE

Results of biased simulations can be reweighted to obtain results for physical properties in a hypothetical semigrand canonical ensemble, by correcting for the effects of the artificial bias $V(N)$. A semi-grand canonical ensemble describes a system with a fixed total number of molecules that can swap molecules with a reservoir at a fixed exchange chemical potential, causing fluctuations in N. The exchange chemical potential, denoted by $\Delta\mu$, is the difference between chemical potential of the copolymer and homopolymer species. Such an ensemble can be used to simulate a small open subsystem within a macroscopic micellar solution.

Semi-grand canonical average values for any dynamical variable can be obtained by reweighting the results of an adaptively biased MC simulation. Let f denote the instantaneous value of any variable that can be computed from knowledge of system configuration (i.e., from the bead positions and the type of each molecule in the system). The semi-grand canonical average of any such variable can be obtained as a re-weighted average

$$\langle f \rangle = \frac{\langle f e^{[\Delta\mu N - V(N)]/k_BT} \rangle_{\rm sim}}{\langle e^{[\Delta\mu N - V(N)]/k_BT} \rangle_{\rm sim}} \quad , \tag{7}$$

where $\langle \cdots \rangle_{\rm sim}$ denotes an average over a sequence of states generated by a biased MC simulation. Somewhat simpler expressions can be given for quantities that only require knowledge of the Gibbs free energy $G(N)$ as a function of N, as discussed below.

A. Probability Distribution for M

The equilibrium distribution of values of N in a semi-grand canonical ensemble, denoted by $P_{eq}(N, \Delta\mu)$, is given by a ratio

$$P_{eq}(N, \Delta\mu) = e^{-\Phi(N,\Delta\mu)/k_B T}/\Xi \quad , \tag{8}$$

in which Ξ is a semi-grand partition function,

$$\Xi = \sum_{N=0}^{\infty} e^{-\Phi(N,\Delta\mu)/k_B T} \tag{9}$$

and in which

$$\Phi(N, \Delta\mu) \equiv G(N) - N\Delta\mu \tag{10}$$

is the constrained semi-grand canonical free energy for a system with a known number of copolymers. Here, $G(N)$ is the Gibbs free energy that we obtain by analyzing the biased MC simulations, as described above.

Figure 1 shows MC results for the free energy $\Phi(N, \Delta\mu)$ vs. N at $\alpha = 10$, 12, 14 and 16 at the value of $\Delta\mu$ for which the free molecule concentration is exactly equal to the critical micelle concentration CMC. The definition of the CMC used here is specified in subsection IV C. At values of $\Delta\mu$ near this critical value, $\Phi(N, \Delta\mu)$ is a function with two well-separated minima.

States in which N has a value near the smaller value of N at which $\Phi(N)$ is minimum contain only free surfactant molecules and occasional small clusters, but no proper micelle. The value of N at this minimum is thus the most probable value of the number of free surfactant molecules in a system with no micelle. This is approximately 23 for $\alpha = 10$ and about 1 for $\alpha = 16$. Near this first minimum, $P_{eq}(N)$ is well approximated by a Poisson distribution

$$P_{eq}(N) \propto \langle N_1 \rangle^N / N! \quad , \tag{11}$$

where $\langle N_1 \rangle$ is the average number of free surfactant molecules. For this distribution, the value of N at which $P_{eq}(N)$ is maximum and $\Phi(N)$ is minimum is equal to zero if $\langle N_1 \rangle < 1$ and nearly equal to $\langle N_1 \rangle$ for $\langle N_1 \rangle > 1$.

The second minimum in $\Phi(N, \Delta\mu)$ corresponds to states that almost always contain a single proper micelle coexisting with zero or more free molecules. The value of N at this minimum is thus approximately equal to the most probable micelle aggregation number plus the average number of free surfactant molecules that would coexist with such a micelle within our simulation cell. A simple estimate of the most probable aggregation number at a given value of $\Delta\mu$ may thus be obtained by taking the difference between the two minima in such a plot of $\Phi(N, \Delta\mu)$.

Between these two minima is a maxima of $\Phi(N)$, at a critical value of N that we denote by $N_{\rm cr}$. This maximum corresponds to a transition state for the formation of a micelle by stepwise association or for the destruction of a micelle by stepwise dissociation, which is important in the next article in this series.

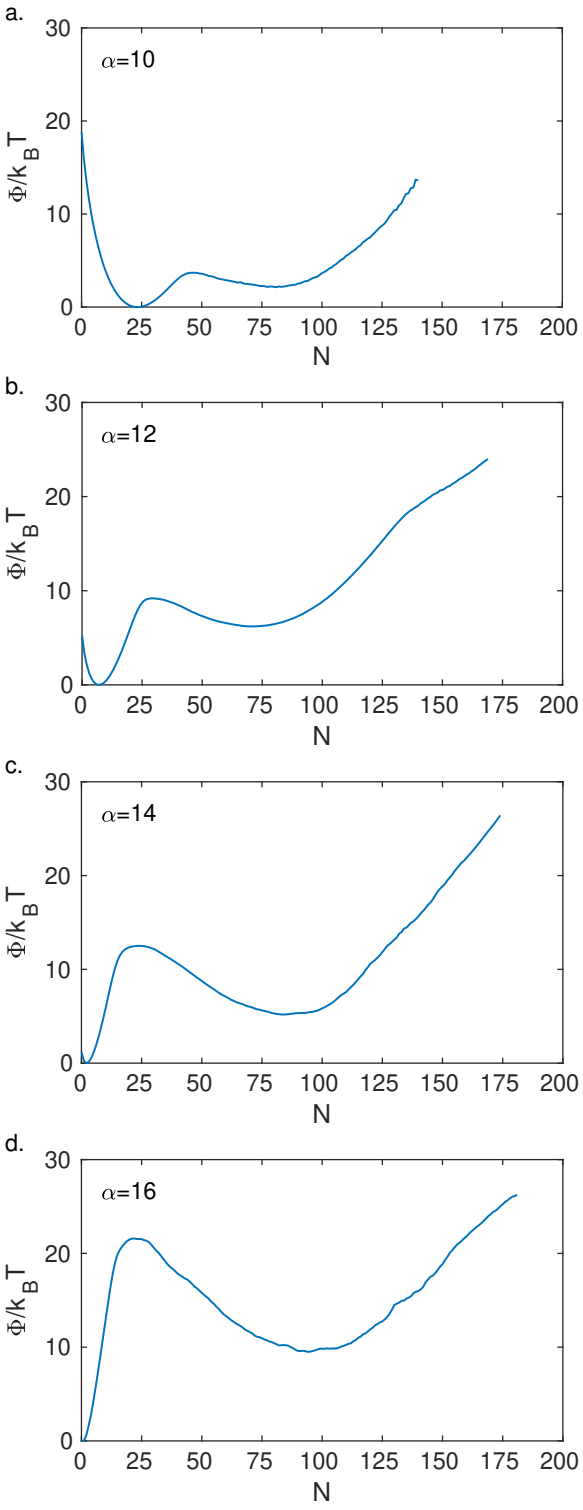

FIG. 1. Semi-grand canonical free energy $\Phi(N)$ of the simulation cell as a function of N (the total number of copolymer molecules) at $\Delta\mu = \Delta\mu_c$, for $\alpha = 10$, 12, 14, and 16, top to bottom. Curves have been shifted vertically such that $\Phi = 0$ at the first minimum.

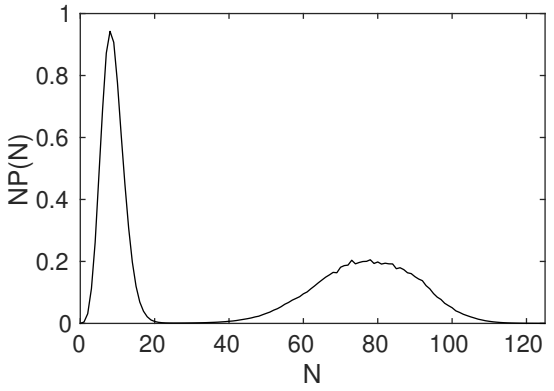

FIG. 2. Product $NP(N)$ vs. the number of copolymer molecules N in the simulation cell at $\Delta\mu = \Delta\mu_c$ for $\alpha = 12$, where $P(N)$ is the probability distribution for N.

B. Free Copolymer Concentrations

Knowledge of the equilibrium probability $P_{eq}(N, \Delta\mu)$ can be used to compute how the total concentration and the concentration of free molecules depends on the exchange chemical potential $\Delta\mu$.

The total copolymer concentration c in a system with exchange chemical potential $\Delta\mu$, including contributions from both free molecules and clusters of all sizes, is given by the ratio $c = \langle N\rangle/V$ in which $\langle N\rangle$ is given by the average,

$$\langle N\rangle = \sum_{N=0}^{\infty} N P_{eq}(N) = \frac{\sum_{N=0}^{\infty} N e^{-\beta\Phi(N,\Delta\mu)}}{\sum_{N=0}^{\infty} e^{-\beta\Phi(N,\Delta\mu)}} \quad . \qquad (12)$$

This sum is generally dominated by values of N near one or both of the minima of $\Phi(N, \Delta\mu)$ or (equivalently) the maxima of $P_{eq}(N, \Delta\mu)$. When $\Delta\mu$ is near the critical value - at which the concentration is equal to the critical micelle concentration - the summand $MP_{eq}(M)$ has similar contributions from both maxima in P_{eq}, as shown graphically in Figure 2.

To estimate the concentration of free molecules, it useful to consider a hypothetical system in which micellization is suppressed. Micelle formation can be artificially suppressed by adding a fictitious infinite free energy penalty to microstates with $N > N_{\rm cr}$, where $N_{\rm cr}$ is the value at which $P_{eq}(N, \Delta\mu)$ has a local minimum. This is equivalent to simply setting $P_{eq}(N, \Delta\mu) = 0$ for all $N > N_{\rm cr}$. Let c_1 denote the concentration of free molecules at a specified value of $\Delta\mu$ in a system with such a truncated probability distribution. This concentration is given by a ratio

$$c_1 = \langle N\rangle_f/V \quad , \qquad (13)$$

where V is the average volume of the simulation unit cell,

α	ϕ_c	$\Delta\mu_c/k_BT$	n_e	σ_m	$\Delta W_d/k_BT$	$\Delta\mu^0/k_BT$	a
10	1.63%	5.49	55	12.1	3.00	9.8	-12.8
12	0.54%	5.96	70	12.8	4.82	11.3	-22.2
14	0.17%	6.17	83	12.1	7.87	12.6	-32.2
16	0.072%	6.54	97	12.2	14.1	13.8	-40.7

TABLE I. Properties of micellar solutions at several values of α. Shown here are values for the copolymer mole fraction ϕ_c at the critical micelle concentration, the corresponding critical exchange chemical potential $\Delta\mu_c$, the most probable micelle aggregation number n_e at $\Delta\mu = \Delta\mu_c$, the standard deviation σ_m of the micelle aggregation number at $\Delta\mu = \Delta\mu_c$, the dissociation barrier ΔW_d defined in Eq. (30), the standard state exchange chemical potential $\Delta\mu^0$, and the second virial coefficent introduced in Eq. (17). The standard deviation σ_m reported here is computed by considering only clusters with $n > n_t$, in order to exclude unimers and other submicellar clusters, and computing the standard deviation of the distribution $P_n \propto e^{-W_n/k_BT}$ for this subpopulation of proper micelles.

and where

$$\langle N\rangle_f = \frac{\sum_{N=0}^{N_{\rm cr}} N e^{-\beta\Phi(N,\Delta\mu)}}{\sum_{N=0}^{N_{\rm cr}} e^{-\beta\Phi(N,\Delta\mu)}} \qquad (14)$$

is the average of N averaged only over states with $N \leq N_{\rm cr}$. Here, we denote the average over such states by the symbol $\langle\cdots\rangle_f$. This constrained average is dominated by values of N near the first minimum in $\Phi(N, \Delta\mu)$, which are states with no micelle.

C. Critical Micelle Concentrations

We define the critical micelle concentration c_c to be the free molecule concentration c_1 in a state in which the average number of free copolymer molecules is equal to the average number in micelles, or, equivalently, in which the total concentration is twice the free molecule concentration. We thus determine the corresponding value of $\Delta\mu$ in this state, denoted by $\Delta\mu_c$, as the value of $\Delta\mu$ at which

$$c = 2c_1 \quad , \qquad (15)$$

and at which we define $c_c = c_1 = c/2$. All results for $\Phi(M)$ shown in Fig. 1 were computed at $\Delta\mu = \Delta\mu_c$, corresponding to systems with $c_1 = c_c$. Values of $\Delta\mu_c$ and mole fraction of surfactant at $\Delta\mu_c$ are given in Table I.

D. Dilute Solution Equation of State

In a sufficiently dilute solution, in which interactions between dissolved copolymers are negligible, we expect

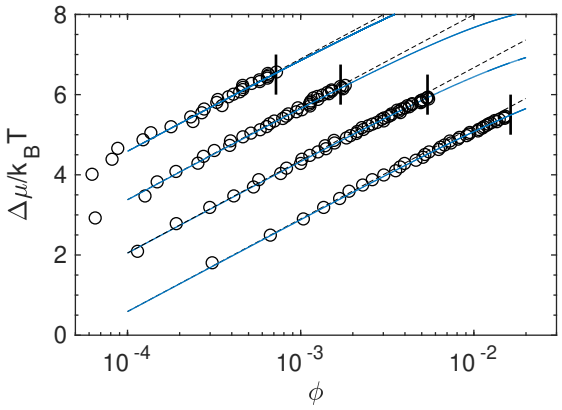

FIG. 3. Exchange chemical potential $\Delta\mu$ vs. mole fraction ϕ of surfactant in dilute solutions for $\alpha = 10$, 12, 14, and 16. Short vertical lines indicate the location of the CMC for each value of α. Solid blue lines through the data are fits to the virial equation of state given in Eq. (17). Dashed lines are dilute solution asymptotes obtained by setting $a = 0$.

the exchange chemical potential to be related to the mole fraction ϕ of copolymers by a dilute solution equation of state

$$\Delta\mu = \Delta\mu^\circ + k_B T \ln\left(\frac{\phi}{1-\phi}\right) \quad . \tag{16}$$

where $\Delta\mu^\circ$ is a standard state exchange chemical potential. At concentrations near the CMC, we find that interactions between individual copolymers cause small but measurable deviations from this dilute solution law. To more accurately describe that data, we have thus fit our data to a virial equation of state

$$\Delta\mu = \Delta\mu^0 + k_B T \left[\ln\left(\frac{\phi}{1-\phi}\right) + a\phi\right] \quad , \tag{17}$$

in which a is a dimensionless second virial coefficient.

Figure 3 shows a fit of our simulation results for the relationship of ϕ and $\Delta\mu$ at concentrations below the critical micelle concentration to Eq. (17), in which $\Delta\mu^\circ$ and a have been treated as fitting parameters. The resulting parameters are reported in Table I.

The Flory-Huggins theory for the type of mixture considered here predicts that, in the dilute limit, $\Delta\mu$ should have an equation of state of the form given in Eq. (16), with

$$\Delta\mu^\circ/k_B T = \chi N_B \quad , \tag{18}$$

where $N_B = 4$ denotes the number of monomers in the core block of a copolymer molecule. The value of $\Delta\mu^\circ$ given in Table I can thus equally well be interpreted as an estimated value of χN_B.

V. CLUSTER FREE ENERGIES

The free energy $\Phi(N, \Delta\mu)$ shown in Fig. 1 is a free energy for entire simulation cell with a known total number of copolymers N. In this section we present our analysis and results for the free energies required to form clusters of specified aggregation number.

We consider a solution that can contain aggregates or clusters of all possible sizes. In what follows, we use the symbol n to denote the aggregation number (i.e., the number of copolymer surfactant molecules) of a particular species of aggregate, where $n = 1$ denotes a free molecule. Let c_n denote the number concentration of aggregates that contain n surfactant molecules, or "n-mers", while c_1 is the concentration of free surfactant, or "unimers".

A. Analysis of Simulation Data

We infer cluster free energies from an analysis of the average number of clusters of each aggregation number within the simulation unit cell. Let N_n denote the number of clusters of aggregation number n in a particular microstate state of the simulation cell, as determined by our cluster identification algorithm. The average value $\langle N_n \rangle$ in a hypothetical semi-grand canonical ensemble with any specified value of $\Delta\mu$ can be obtained from a trajectory of states generated by a biased simulation by applying Eq. (7) to the dynamical variable N_n.

The joint probability $P(N_1, N_2, \ldots)$ of obtaining a state with a specified list of numbers of clusters of each size (N_1 unimers, N_2 dimers, etc.) is given by a Boltzmann factor

$$P(N_1, N_2, \ldots) \propto e^{-\Phi(N_1, N_2, \ldots)/k_B T} \tag{19}$$

in which $\Phi(N_1, N_2, \ldots)$ is the semi-grand canonical free energy of a system constrained to have a specified number of clusters of each aggregation number.

Our analysis treats the system as a dilute solution of non-interacting clusters. The limitations of this assumption are discussed in subsection V C. In this limit, the constrained free energy $\Phi(N_1, N_2, \ldots)$ of our simulation cell can be expressed as a sum

$$\Phi(N_1, N_2, \ldots) = \Phi_h + \sum_{n \geq 1} \left[N_n \Phi_n^{\rm ex} + \ln(N_n!)\right] \quad . \tag{20}$$

Here, Φ_h is the semi-grand canonical free energy of a homopolymer reference system, with no copolymer molecules, while $\Phi_n^{\rm ex}$ is the excess free energy required to form a n-mer anywhere in our simulation cell. Our goal is to infer values of $\Phi_n^{\rm ex}$ for all n from our simulation data.

In our simulations, we have imposed the constraint that there can be no more than one cluster of size n greater than or equal to a critical size $n_{\rm cut} = 8$. The effects of this constraint must be taken into account in our

analysis of the frequency of large clusters, $n > n_{\rm cut}$. It is useful to begin, however, by considering the behavior of a simpler model, in which no such constraint is imposed. This can be shown to correctly describe the statistics of small clusters with $n < n_{\rm cut}$ that are unaffected by the constraint.

Unconstrained Systems: In the absence of any artificial constraint on the allowed values of N_n, this model of non-interacting clusters given in Eq. (20) yields a joint probability distribution

$$P(N_1, N_2, \ldots) \propto \prod_{n=1}^{\infty} \frac{e^{-\beta \Phi_n^{\rm ex} N_n}}{N_n!} \quad . \tag{21}$$

This yields a distribution in which values of N_n for different values of n are statistically independent, and in which the probability distribution N_n is given for each n by a Poisson distribution with an average value

$$\langle N_n \rangle = e^{-\beta \Phi_n^{\rm ex}} \quad . \tag{22}$$

In the absence of any artificial constraint on the number of micelles, we can thus infer $\Phi_n^{\rm ex}$ from a measurement of $\langle N_n \rangle$ by setting $\Phi_n^{\rm ex} = -k_B T \ln \langle N_n \rangle$.

Constrained System: We now consider how to take into account the constraint prohibiting the appearance of more than one large cluster. Because the constraint does not involve values of N_n for clusters of size $n < n_{\rm cut}$, values of N_n for these small clusters remain statistically independent of one another, and of values of N_n for larger clusters. As a result, we may correctly infer values of $\Phi_n^{\rm ex}$ for these smaller clusters by setting $\Phi_n^{\rm ex} = -k_B T \ln \langle N_n \rangle$, exactly as in the absence of the constraint. Values of N_n for larger clusters are, however, strongly correlated by the requirement that no more than one of them can have a value $N_n = 1$. The allowed states of the list of values of N_n for $n \geq n_{\rm cut}$ are thus: (a) A state with no micelle, for which $N_n = 0$ for all $n \geq n_{\rm cut}$, and (b) a set of states in which the system contains a single micelle of some specified size greater than $n_{\rm cut}$. Let P_0 denote the probability that a randomly chosen state in a semi-grand canonical ensemble will not contain any cluster of size $n \geq n_{\rm cut}$. For $n \geq n_{\rm cut}$ let $\langle N_n \rangle$ denote the average number of clusters of aggregation number n in the same ensemble, which is also equal to the probability that one such cluster exists. These probabilities must satisfy

$$1 = P_0 + \sum_{n=n_{\rm cut}}^{\infty} \langle N_n \rangle \quad . \tag{23}$$

In a constrained system, the probability $\langle N_n \rangle$ finding a cluster of specified size $n > n_{\rm cut}$ is related to the probability P_0 by

$$e^{-\Phi_n^{\rm ex}/k_B T} = \frac{\langle N_n \rangle}{P_0} \quad . \tag{24}$$

Eq. (24) has been used to compute $\Phi_n^{\rm ex}$ each $n \geq n_{\rm cut}$ from measured values of P_0 and $\langle N_n \rangle$.

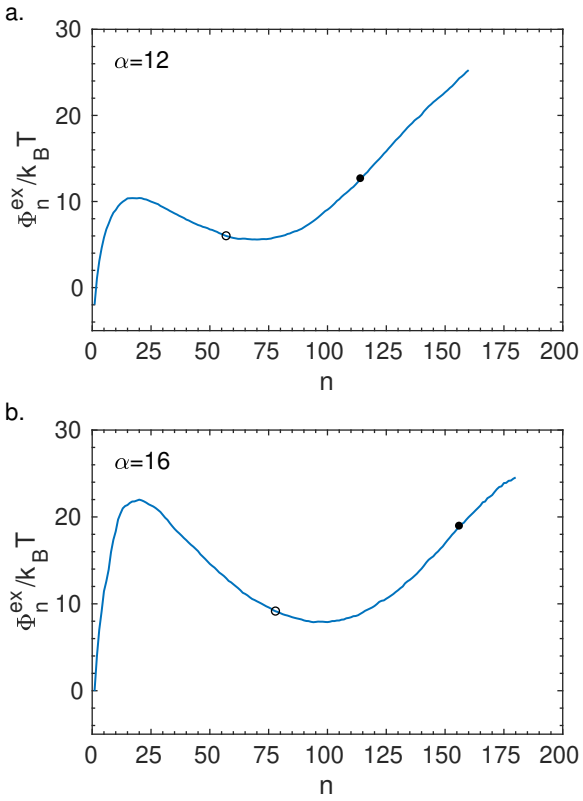

FIG. 4. Results for the excess free energy $\Phi_n^{\rm ex}$ to form a cluster of aggregation number n plotted vs. n for systems with $\Delta\mu = \Delta\mu_c$ (i.e., $c_1 = c_c$), at $\alpha = 12$ (top plot) and $\alpha = 16$ (bottom plot). The filled black dot on each curve indicates the value of n above which, in the absence of any constraint on the number of large micelles, the appearance of two micelles of aggregation number $n/2$ within our simulation cell would become more probable than the appearance of a single micelle of aggregation number n. The corresponding unfilled dot on the curve marks the aggregation number $n/2$ of the corresponding fission products.

Results for $\Phi_n^{\rm ex}$ are shown plotted vs. n for systems with $\Delta\mu = \Delta\mu_c$ for several values of α in Fig. (4). Results for $\Phi_n^{\rm ex}$ were obtained using Eq. (22) for $n < n_{\rm cut}$ and Eq. (24) for $n \geq n_{\rm cut}$. Reassuringly, the resulting estimate of $\Phi_n^{\rm ex}$ is continuous at $n_{\rm cut} = 8$. The location of the local micellar minimum in $\Phi_n^{\rm ex}$ corresponds to the most probable micelle aggregation number, which we denote by $n_{\rm e}$. Values of $n_{\rm e}$ at $\Delta\mu = \Delta\mu_c$ increases with increasing α by almost a factor of 2 over the range studied here, from $n_{\rm e} = 55$ for $\alpha = 10$ to $n_{\rm e} = 97$ for $\alpha = 16$. The value of n at which $\Phi_n^{\rm ex}$ has a local maximum, denoted here by $n_{\rm t}$, is the critical aggregation number for creation of a micelle by stepwise association and for destruction of a micelle by stepwise dissociation.

Plots of $\Phi_n^{\rm ex}$ vs. n shown in Fig. 4 broadly resemble the corresponding plots of $\Phi(N)$ vs. N in Fig. 1, except for a few key differences. The behavior is qualitatively different for small values of n or N, because $\Phi_n^{\rm ex}$ always has a minimum at $n = 1$, while $\Phi(N)$ shows a mini-

mum at a value of N corresponding to the most probable number of free copolymers in the simulation cell. Any attempt to describe the prevalance of different sizes of small, subcritical clusters must thus be based on a computation of $\Phi_n^{\rm ex}$, rather than $\Phi(N)$. The two functions are more qualitatively similar for $n > n_{\rm t}$. In this case, the most obvious difference is a shift in the absicca arising from the fact that the value of N in this region is the sum of the aggregation number of a large cluster and the number of unimers that coexist with it within the unit cell, whereas $\Phi_n^{\rm ex}$ is plotted as a function of the true aggregation number.

In appendix B, we consider the question of whether the behaviour of $\Phi_n^{\rm ex}$ for $n > n_{\rm t}$ can be inferred from knowledge of $\Phi(M)$ alone. We consider there how well one can approximate $\Phi_n^{\rm ex}$ for $n > n_{\rm t}$ simply by re-expressing $\Phi(M)$ as a function of the average number of molecules in the micelle, when we compute the average number of molecules in a micelle by subtracting the average number of free molecules from M. We find that this yields an accurate estimate for $\Phi_n^{\rm ex}$ to within a constant for $n \geq n_{\rm t}$ in systems with large values of α (and thus very few free molecules), but that this procedure fails at lower values of α for which the number of free molecules in the simulation is larger.

The motivation for our introduction of a constraint prohibiting the formation of more than one micelle can be understood by considering when spontaneous fission would be otherwise become likely. The solid dot on each plot in Fig. 4 indicates the value of n above which, in the absence of this constraint, the appearance of two clusters of size $n/2$ within the simulation cell would be predicted to be more probable than the appearance of one cluster of aggregation number n, for a simulation cell of the size used here. These probabilities become equal where $\Phi_n^{\rm ex} = \Phi_{n/2}^{\rm ex} + \ln(2!)$. Open circles shows the corresponding value of $n/2$. The value of n above which existence of a single large cluster of n molecules would be less probable than the existence of any two smaller clusters of the same total aggregation number is actually somewhat less than this estimate, because of the possibility of forming two clusters of unequal size. The constraint is intended to improve the reliability of sampling these large, thermodynamically unstable clusters, and to avoid the need to accurately sample the relative probability of one- and two-micelle states.

The quantity $\Phi_n^{\rm ex}(\Delta\mu)$ is the semi-grand canonical free energy required to form a single cluster of aggregation number n anywhere in our simulation cell. The contribution of translational entropy thus causes Φ_n to depend on the size chosen for the simulation cell: Values for $\Phi_n^{\rm ex}$ obtained from simulations performed using simulation cells of different volumes V_1 and V_2 would differ by $-k_BT\ln(V_2/V_1)$ because of the larger translational entropy of the micelle in a larger cell.

B. Standard Free Energies

When defining a free energy of formation for a mobile cluster, there is no way to avoid introducing an arbitrary convention regarding the volume to which the cluster is confined. The above definition of $\Phi_n^{\rm ex}$ uses the volume of the simulation cell for this purpose, which is convenient when analyzing simulation data. The more conventional approach is to express the chemical potential of each species relative to that obtained at some standard concentration. In this approach, the chemical potential for clusters of aggregation number n (n-mers) in a diute solution is expressed as a sum

$$\mu_n = \mu_n^\circ + k_BT\ln(c_n/c^\circ) \tag{25}$$

where c° is an arbitrary standard concentration and μ_n° is a corresponding standard-state chemical potential. For simplicity, we take c° to be the same for all species.

The relationship between the concentrations of n-mers and unimers may be established by considering a hypothetical reaction in which an n-mer is formed by association of n unimers (see, e.g., Ref. [38] for a previous discussion). The equilibrium criterion for this reaction requires that $\mu_n = n\mu_1$. Combining this criterion with Eq. (25) and solving for c_n yields an n-mer concentration

$$c_n^*(c_1) = c^\circ e^{-W_n(c_1)/k_BT} \tag{26}$$

for all $n > 1$, in which we have defined

$$W_n(c_1) \equiv \mu_n^\circ - n\mu_1 \quad . \tag{27}$$

Here and hereafter, $c_n^*(c_1)$ denotes the equilibrium concentration of n-mers in a solution with a known unimer concentration c_1. The quantity $W_n(c_1)$, which depends on both n and c_1, is the free energy required to create an additional n-mer in a hypothetical solution with a specified unimer concentration c_1 and a standard n-mer concentration $c_n = c^\circ$. The dependence of W_n on c_1 can be made explicit by writing $W_n(c_1)$ as a sum

$$W_n(c_1) = W_n^\circ + k_BT\ln\left(\frac{c_1}{c^\circ}\right) \tag{28}$$

in which W_n° is a standard free energy of formation for an n-mer, given by $W_n^\circ \equiv W_n(c_1 = c^\circ)$. Note that, if Eq. (27) is used to define W_1, then $W_1^\circ = 0$, by definition.

The quantities W_n and $\Phi_n^{\rm ex}$ can be related by comparing equivalent expressions for c_n. Because Eq. (26) is correct only in a dilute system with no artificial constraint on the number of micelles, we compare it to Eq. (22) for $\langle N_n\rangle$, which was valid in under the same conditions. By equating corresponding expressions for $c_n^* = \langle N_n\rangle/V$, we obtain

$$W_n = \Phi_n^{\rm ex} - k_BT\ln(c^\circ V) \quad . \tag{29}$$

Values of $W_n(c_1)$ and $\Phi_n^{\rm ex}(c_1)$ obtained at the same value of c_1 thus differ by a shift that is independent of n. This

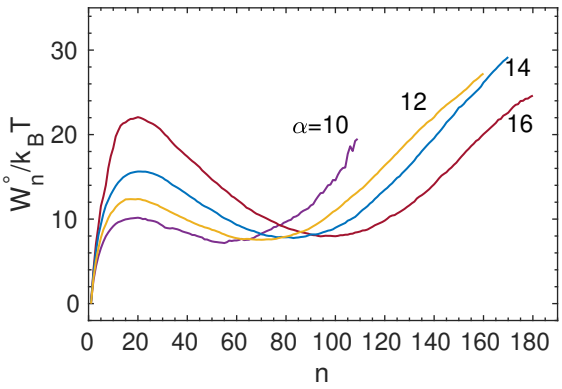

FIG. 5. Standard micelle formation free energy W_n° plotted vs. cluster aggregation number n for $\alpha = 10$, 12, 14, and 16, defined using a standard concentration $c^\circ = c_c$. The standard free energy W_n° is defined by Eqs. (25-28).

implies that W_n and $\Phi_n^{\rm ex}$ have local maxima and local minima at the same values of of n, which we denote by $n_{\rm t}$ and $n_{\rm e}$, respectively.

It is convenient to choose a standard concentration $c^\circ = c_c$ equal to the critical micelle concentration. Values of the standard free energy of formation W_n° defined using this convention are plotted vs. n in Figure 5 for systems all four values of $\alpha = 10$, 12, 14, and 16. With this convention, W_n° corresponds to the value of W_n at $c_1 = c_c$ or $c = 2c_c$.

The free energy barrier for dissociation of a micelle into unimers by stepwise processes, which we denote by ΔW_d, is given by the difference

$$\Delta W_d \equiv W_{n_{\rm t}} - W_{n_{\rm e}} \tag{30}$$

between the values of W_n at its local maximum at $n_{\rm t}$ and at its local minimum at $n_{\rm e}$. The difference ΔW_d depends on unimer concentration c_1, and decreases with decreasing c_1, but is independent of the arbitrary choice of a standard concentration c°. For the system studied here, values of ΔW_d at $c_1 = c_c$ vary from a 3 - 14 k_BT as α varies from 10 - 16. In the accompanying paper [1], we give quantitatitve predictions for the average time before a randomly chosen micelle in an equilibrated solution would be destroyed by stepwise dissociation, which we refer to as the equilibrium dissociation lifetime. Predictions for this lifetime depend exponentially on $\Delta W_d/k_BT$.

C. Inter-Micelle Interactions

Throughout the above analysis, we have neglected all effects of interactions between micelles. Because the simulations from which we infer micelle free energies never contain more than one micelle, the free energies reported here are correct for systems with very low micelle number concentrations, as would be found very near the cmc. Inclusion of the effects of inter-micelle interactions would thus have almost no effect upon our predictions of the CMC, but could effect predictions for the relationship between chemical potential and the total surfactant concentration at concentrations sufficiently far above the CMC.

Effects of interactions between micelles upon thermodynamic properties can be estimated using a second virial approximation. Because block copolymer micelles exhibit strong repulsive interactions, due to repulsion of their coronas, we may model micelles for this purpose as repulsive spheres with an effective hard-core radius d comparable to the micelle radius of gyration R_g. The micelle number concentration, which we denote by c_m, can be approximated by $c_m = (c - c_c)/n_{\rm e}$, where c is total surfactant number concentration and $n_{\rm e}$ is the most probable micelle aggregation number. The second virial coefficient B for hard spheres of effective diameter d is given by $B = 2\pi d^3/3$. The corresponding increase δW in the free energy required to form a micelle of aggregation number $n \simeq n_{\rm e}$ is $\delta W = k_BTc_mB$.

As an example, consider a system with $\alpha = 10$ at the total copolymer concentration $c = 2c_c$, which is the concentration at which we have plotted cluster free energies. Because $\alpha = 10$ is the lowest value of α we have considered, with the highest CMC, this is the system for which interaction effects are greatest at this value of c/c_c. For this system, $\phi_c = 0.0163$, $c_c = 0.00153\sigma^{-3}$, and $c_m = 2.8 \times 10^{-5}\sigma^{-3}$, corresponding to a characteristic distance $c_m^{-1/3} \simeq 33\sigma$ between micelles. Approximating $R_g \simeq 5\sigma$ and $d = 2R_g$ yields a estimated second virial coefficient $B = 2\pi d^3/2 \simeq 2 \times 10^3\sigma^3$. This estimate yields a predicted free energy shift of $\delta W \simeq 6 \times 10^{-2}k_BT$ per micelle. This free energy change would be relevant in very precise quantitative analysis, but is negligible for most purposes. The resulting change in free energy per micelle would be lower for larger values of α at the same value of c/c_c.

VI. MICELLE SHAPE

The shape of the core of micellar clusters has been characterized by examining the eigenvalues of the gyration tensor of the micelle core. The instantaeous gyration tensor, denoted by $\mathbf{S}$, is defined for a cluster of aggregation number n as a sum

$$\mathbf{S} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{r}_i\mathbf{r}_i \tag{31}$$

in which $\mathbf{r}_i$ is the position of the ith B bead in the cluster, defined relative to the micelle core center of mass, and $\mathbf{r}_i\mathbf{r}_i$ is a dyadic tensor product. Here, $N = N_Bn$ is the total number of B beads in a cluster of n molecules, where $N_B = 4$ is the number of B beads per molecule in the cluster.

The eigenvalues of this gyration tensor reveal information about the shape of the micelle core. If the micelle core were perfectly spherical, then the three eigenvalues

of S would have identical values. If the micelle were instead cylindrical or rodlike there will be one larger eigenvalue and two equal smaller eigenvalues. To characterize the shape, we compute for a sequence of states the three eigenvalues of $\mathbf{S}$ which we denote by λ_1, λ_2, and λ_3. By convention, we index the three eigenvalues in ascending order, such that $\lambda_3 \geq \lambda_2 \geq \lambda_1$. For each value of the aggregation number n, we evaluate average values of the largest, middle, and smallest eigenvalues over clusters of that aggregation number. We then define three associated effective root-mean-squared radiii R_1, R_2, and R_3, defined by setting

$$R_\alpha = \langle \lambda_\alpha \rangle^{1/2} \tag{32}$$

for $\alpha = 1, 2, 3$. By construction, these lengths are also ordered such that $R_1 \leq R_2 \leq R_3$. The radius of gyration of the core of a cluster of known aggregation number, denoted by $R_{g,\mathrm{core}}$ is related to these lengths by the relation

$$R_g^2 = \langle \mathrm{Tr}[\mathbf{S}] \rangle = R_1^2 + R_2^2 + R_3^2 \quad , \tag{33}$$

where $\mathrm{Tr}[\cdots]$ denotes the trace of a tensor.

Figure 6 shows the results of a computation of the the three ordered root-mean-squared core radii as functions of aggregation number n for $\alpha = 12$ and $\alpha = 16$. The dotted line in this graph shows the common value that would be expected for all three radii if the core were a perfectly spherical region containing $N_B n$ B monomers with a volue of $v = 1/c$ per monomer, computed using the bulk density $c = 3\sigma^{-3}$ appropriate to this model in the limit of long homopolymers. At aggregation numbers near the equilibrium aggregation number the effective radii are similar and stay clustered near the value expected for a spherical core (the dotted line). The fact that the three values are never exactly equal reflects the fact that the core shape fluctuates somewhat about a typical spherical shape, and that the eigenvalues are ordered by convention. Beginning at some crossover aggregation several tens of percent above m_{eq}, the micelle undergoes a crossover to a more rodlike shape, as indicated by the onset of a steady increase in the value largest value R_3 and the two smaller values, R_1 and R_2, in which R_3 increases starts to increase approximately linearly with n, while R_1 and R_2 remain similar to one another and change comparatively little with cchanges in n. This crossover to a rod-like structure appears to begin in the range $n = 100$ - 125 for $\alpha = 12$, for which $n_\mathrm{e} = 70$, and begins over the range $n = 125 - 150$ for $\alpha = 16$, for which $n_\mathrm{e} = 97$. At values of n slightly above the onset of this crossover, there is a slight decrease in the values of R_1 and R_2, which we believe reflects the formation of dumbbell like structure with a slightly narrower neck. Images from simulation demonstrating this transition are shown in Fig. 7 which shows typical micelles for α equal to 16 and for n equal to 99 and 170.

For the model studied here, micelles are roughly spherical for values of n near the most probable value n_e, but start to become rodlike for modestly larger values of n.

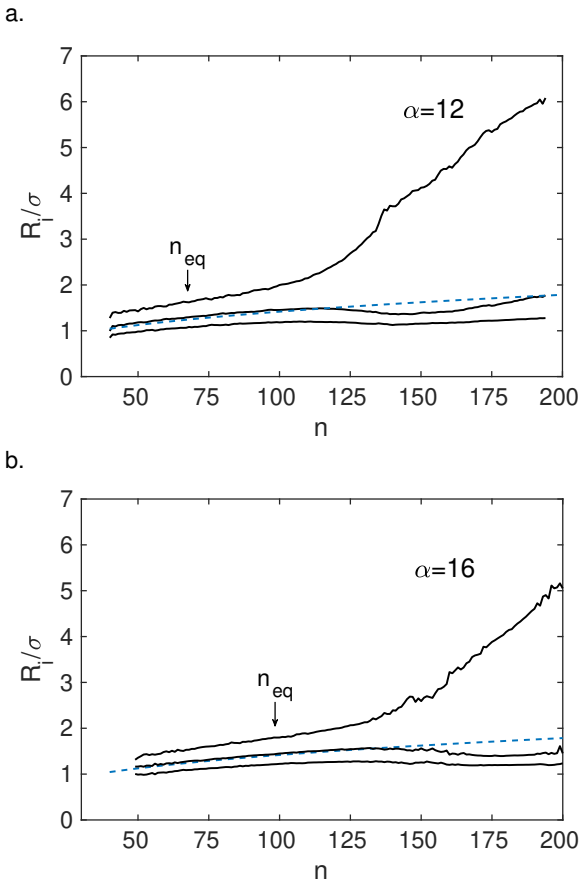

FIG. 6. Ordered root-mean-squared radii of the micelle core measured along the eigendirections of the gyration tensor, plotted as a function of micelle aggregation number n for $\alpha = 12$ (upper plot) and $\alpha = 16$ (lower plot). Solid lines represent the measured values of R_1, R_2 and R_3 (from bottom to top). The dotted line represents the corresponding value expected for a perfectly spherical core with a fixed volume per core monomer.

While acquiring data at very large values of the aggregation number n it was found that for micelles with aggregation numbers roughly twice that of the equilibrium aggregation number the micelles form two separate micelles bridged by one or more surfactant molecules. Formation of this structure is the result of our prohibition on forming more than one micelle in each simulation volume which prevents the micelles from properly separating. The reliable formation of this structure marks an end to where we are able to gather reliable equilibrium data using this technique and so we do not present data where this structure was observed.

VII. CONCLUSIONS

This work presents a methodology for accurately measuring micelle formation free energies in a particular type of model system. The method used here is a semi-grand

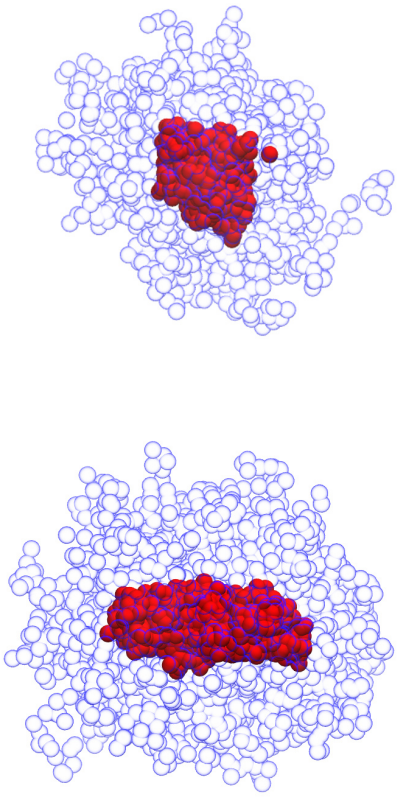

FIG. 7. Snapshots of micelles with $\alpha = 16$ with aggregation number $n = 99$ (upper image) and $n = 170$ (bottom image). The core forming B beads are shown as solid red while the corona forming A beads are only outlined in blue. Surrounding homopolymers are not shown. The aggregation number $n = 99$ of the micelle in the upper image is very close to the most probable value n_e for this value of α. Micelles undergo a crossover from spherical to elongated cores with increasing n.

MC simulation technique similar to one used previously by Cavallo *et al.* [30]. The method relies on the use of a semi-grand ensemble and a bias potential that depends on the total number of copolymer molecules to create an ensemble in which different values of the micelle aggregation numbers are almost equally probable. The main differences between the methodology used here and that of Cavallo *et al.* are (1) the introduction of a MC acceptance criterion that prohibits the formation of states containing more than one proper micelle, and (2) the introduction of a method of computing the micelle free energies from the probabilities of observing clusters of different sizes. The prohibition on states with more than one micelle improves the reliability of sampling large micelles that are thermodynamically unstable to fission into two smaller micelles. Computation of cluster formation free energies from formation probabilities, rather than from the dependence of the total system free energy on the number of surfactant molecules, is necesary to obtain accurate free energies for small clusters, particularly in systems of more soluble surfactants in which the simulation cell contains many free molecules.

Results were obtained for a highly asymmetric copolymers over a range of values of the α parameter (or the effective χ parameter), for which the critical micelle mole fraction ranged from 1.6% to 0.07%. Over this range, the barrier to stepwise dissociation was found to increase from (3-14)k_BT. In an accompanying paper [1], this is shown to cause the rate of stepwise micelle dissociation to decrease by many orders of magnitude over the same parameter range. The relative length of the A and B blocks used in this study (1/8 of the beads are of type B) was chosen so as to obtain spherical micelles in equilibrium, and is the same as the composition used by Cavallo *et al.* A study of the dependence of micelle shape on aggregation number shows, however, that there is a crossover from spherical to rodlike morpology beginning at an aggregation number that is approximately 50% larger than the most probable aggregation number. This change in shape is relevant to the interpretation of results for fission rates discussed in the accompanying paper, where we find that fission occurs most frequently in equilibrium for relatively rare micelles that are significantly larger than the most probable size.

The main limitation of the semi-grand simulation method used here is the fact that it is directly applicable only to a very special type of system, i.e., copolymer/homopolymer mixtures with chains of equal length. We chose to study both equilibrium and dynamical properties of a system of this type in order to take advantage of the efficiency of this method. There is, however, a clear need for more general methods of efficiently computing absolute formation free energies for micelles in othe types of system.

ACKNOWLEDGMENTS

This work was supported primarily by NSF grant DMR-1310436, with partial support from the NP and MP programs of the University of Minnesota Industrial Partnership for Interfacial and Materials Engineering (PRIME) center. We acknowledge the Minnesota Supercomputing Institute (MSI) at the University of Minnesota for providing computational resources for the work reported here.

Appendix A: Free Molecule Concentrations

In states of the simulation in which N, the number of copolymers, is greater than or comparable to the value of N_{cr} at the CMC, the simulation cell generally contains a micelle coexisting with a few free copolymer molecules. In such a state, the micelle and the surrounding solution of free molecules coexist at an exchange chemical potential $\Delta\tilde{\mu}(N)$ that depends slightly on N, and that is given

by

$$\Delta\tilde{\mu}(N) = \frac{\partial G(N)}{\partial N} \quad . \tag{A1}$$

The chemical potential $\Delta\tilde{\mu}(N)$ for a system with a known number of copolymers is generally not equal to the macroscopic chemical potential $\Delta\mu$ that we used to define $\Phi(N, \Delta\mu)$ in our analysis of a semi-grand ensemble. Observe that

$$\begin{aligned} \frac{\partial\Phi(N,\Delta\mu)}{\partial N} &= \frac{\partial(G(N) - \Delta\mu N)}{\partial N} \\ &= \Delta\tilde{\mu}(N) - \Delta\mu \quad . \end{aligned} \tag{A2}$$

The derivative $\partial\Phi(N, \Delta\mu)/\partial N$ vanishes only at the extrema of $\Phi(N, \Delta\mu)$, corrresponding to the two local minima of $\Phi(N, \Delta\mu)$ (or maxima of $P(N)$), and to the maxima at $N_{\rm cr}$ (or the minimum of $P(N)$). Eq. (A2) thus implies that the N-dependent exchange chemical potential $\Delta\tilde{\mu}(N)$ of a system with a known value of N is equal to the macroscopic exchange chemical potential $\Delta\mu$ used to define a semi-grand ensemble only at values of N corresponding to these extrema. Specifically, this means that $\mu(N) = \Delta\mu$ at the most probable micelle size $n_{\rm e}$, and at the transition state $n_{\rm t}$, but not at values of n intermediate between these values.

In a simulation volume that contain a micelle in coexistence with free chains, the concentration of free molecules in regions of the simulation cell well outside the micelle corona is equal to a concentration c_1 of free chains in a system with chemical potential $\Delta\tilde{\mu}(N)$. Figure 8 shows the corresponding total average number of free molecules $N_{\rm f}$ in our simulation cell for systems with different values of N and α. At the lowest values of $\alpha = 10$ and 12, the box contains a substantial number of free molecules (e.g., 8-25) free molecules, but $N_{\rm f}$ changes very little with changes in N. At the highest value of $\alpha = 16$, fractional variations in $N_{\rm f}(n)$ are larger, but absolute values $N_{\rm f}$ are much smaller. As a result, absolute values of $N_{\rm f}$ actually never varies with changes in N by more than about one molecule over the range of values of N and α studied here.

Appendix B: Micelle vs. System Free Energy

In this work, we present results for several closely related semi-grand canonical free energies. The quantity $\Phi(N)$ is the semi-grand canonical free energy of an entire simulation cell with a constrained number N of copolymer molecules, defined to within an arbitrary constant. The quantity W_n is instead the free energy required to form a cluster of specified aggregation number n. The quantity $\Phi_n^{\rm ex}$ differs from W_n by an n-independent constant, and is thus essentially equivalent to W_n. For values of values of n greater than the critical value of $n_{\rm t}$, the dependence of W_n on n qualitatively resembles the dependence of $\Phi(N)$ on N. The most obvious difference

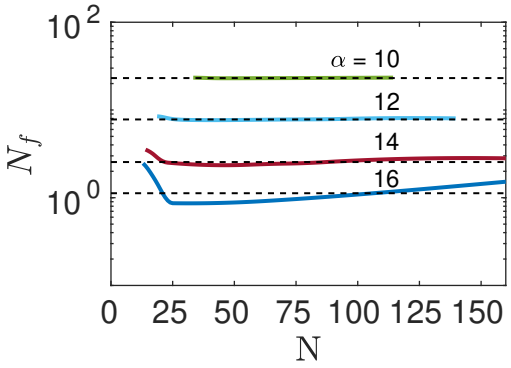

FIG. 8. The average number of free molecules in the simulation unit cell, denoted by $N_{\rm f}$, versus the total number of molecules in the cell, denoted by N. The difference in these two numbers is the micelle excess aggregation number $N_{\rm ex}$

is that points on a graph of $\Phi(N)$ occurs at N that exceeds the value of n at corresponding points on a plot of W_n simply because N includes free molecules, in addition to molecules that are within the micelle. This suggests that, under favorable circumstances, one might be able to obtain a close approximation for W_n, to within a constant, by simply plotting $\Phi(N)$ as a function of an approximation for the micelle aggregation number that we obtain simply by subtracting the average number of free molecules in the simulation cell from N.

In what follows we define the average excess aggregation number, denoted by $N_{\rm ex}$, to be equal to the difference

$$N_{\rm ex}(N) = N - N_{\rm f}(N) \tag{B1}$$

between the actual total number of copolymer molecules N in the simulation cell and the average number of free molecules $N_{\rm f}$ that we would expect in the simulation unit cell in the absence of a micelle. In this definition, $N_{\rm f}$ is computed at the exchange chemical potential $\Delta\tilde{\mu}(N)$ characteristic of a system with N copolymers, as determined in appendix A by using the relevant value of $\Delta\mu(N)$ in Eq. (A2). Because $\Delta\tilde{\mu}(N)$ can be expressed as a function N alone, this yields a definition of $N_{\rm ex}$ as a function of N.

In order to compare $\Phi(N)$ to W_n we use $N_{\rm ex}(N)$ as an approximation for the average aggregation number in the micelle in a system containing a total of N copolymer molecules. Performing this approximation allows us to write Φ as a function of $N_{\rm ex}$; however this form of approximation results in a discrete function $\Phi(N_{\rm ex})$ but the values of $N_{\rm ex}$ are generally noninteger unlike the true free energy surface which is only well defined for integer aggregation numbers. To make a proper comparison we construct a cubic spline approximation for $\Phi(N_{\rm ex})$ which allows us to interpolate to the proper integer values. In Figure 9, we compare plots of W_n vs. n to corresponding plots of Φ plotted as a function of $N_{\rm ex}(N)$, at several values of α. The plots of Φ vs. n shown here where

obtained by simply replotting $\Phi(N)$ for each value of N using the corresponding value of $N_{\rm ex}(N)$ as an absicca. Results for Φ vs. $N_{\rm ex}$ are not shown for values of $N_{\rm ex}$ much lower than $n_{\rm t}$ where this approximation expected to fail, since the behavior of $\Phi(N)$ and W_n is qualitatively different for very small values of N or n. The two functions agree for all $n > n_{\rm t}$ for the highest value of $\alpha = 16$, but differ significantly for $\alpha = 10$. There is a particularly large discrepancy near the local maximum $n \sim n_{\rm t}$ that increases with decreasing α. The most important difference between these two cases is the average number of free molecules in the simulation cell, which is approximately 23 for $\alpha = 10$ but approximately 1 for $\alpha = 16$. The large discrepancy obtained at $n \sim n_{\rm t}$ at $\alpha = 10$ appears to be a result of the fact that, at $\alpha = 10$, a system with N such that $N_{\rm ex}(N) \sim n_{\rm t}$, the system may exist in a "supersaturated" state with no large cluster but an unusually large number of free molecules rather than in a state with a cluster of aggregation number similar to $n_{\rm t}$. This reflects the fact that, in this case, the free energy cost of $W_{n_{\rm t}} \sim 10 k_B T$ of forming such a cluster is actually somewhat less than the free energy cost of a corresponding fluctuation in the number of free molecules in such a system. This comparison indicates that this simple method of shifting $\Phi(N)$ to construct an approximation for W_n is reliable only for values of $n \geq n_{\rm t}$ corresponding to proper micelles only in conditions in which there are very few free molecules in the simulation cell, so that the relationship between the two functions is not complicated by fluctuations in the number of free molecules.

Appendix C: Effect of Constraints on M

In each of the biased MC simulations described here, the value of N (the number of copolymers in the box) was constrained to remain with some specified range. This constraint is imposed by simply prohibiting transformation moves that, if accepted, would take N outside the allowed range. Introduction of such a constraint introduces an error in the values obtained for the cluster free energy $\Phi_n^{\rm ex}$ for values of the cluster aggregation number n very near the maximum allowed value of N. An example of this behavior is shown by the dashed line in Fig. 10, which shows the measured free energy W_n obtained in a simulation in which the total number of copolymers in the box was constrained to $N \leq 140$. Note the rapid increase in the value of W_n as n approaches 140. A much more accurate estimate of the free energy in this range is shown by the dotted line, which shows results obtained with a constraint that $N \leq 160$, for which this artifact does not occur within the range shown in this graph.

This error in our measurement of $\Phi_n^{\rm ex}$ is a result of an artificial constraint on the number of free unimers that is imposed by the imposition of a constraint on N. Consider a simulation in which N is constrained to a range $N \leq N$. Let Ξ_n denote the constrained semi-grand canonical partition function associated with the set of all

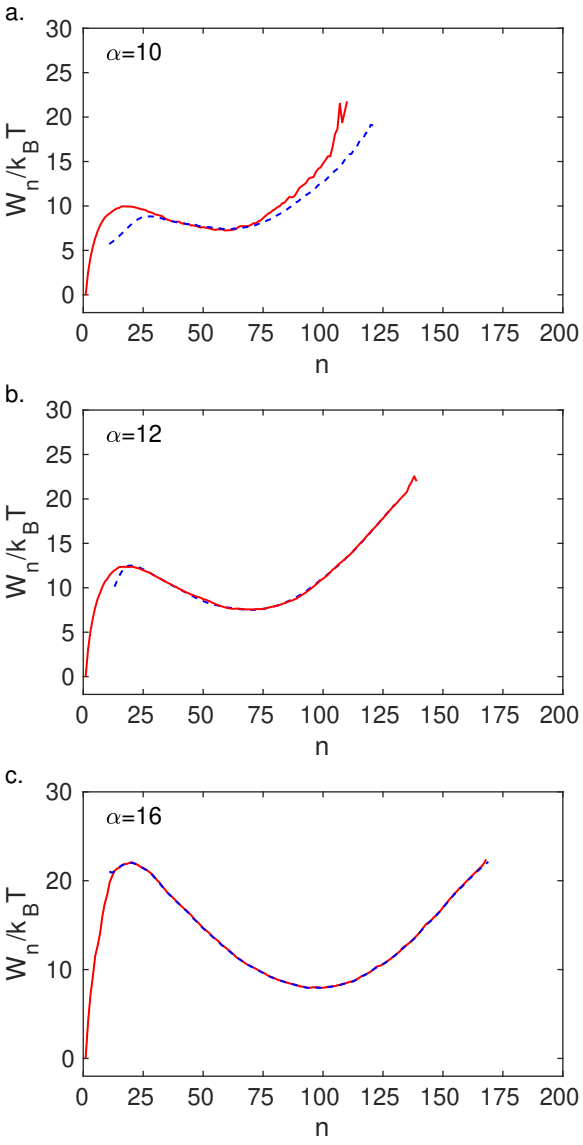

FIG. 9. Comparisons of the micelle formation free energy W_n plotted vs n to the simulation free energy Φ plotted vs. the excess aggregation $N_{\rm ex}$, for $\alpha = 10$, 12, and 16, top to bottom. The solid red line shows W_n vs. n, while the dashed blue line shows $\Phi(N)$ vs. $N_{\rm ex}(N)$. Results for $\Phi(N)$ have been shifted vertically by a constant amount to match the value of W_n at its minimum.

microstates in which the system contains one large cluster of aggregation number n and n unimers. The constrained partition function is thus given to within a constant by a product

$$\Xi_n = e^{-\Phi_n^{\rm ex}/k_B T} S_{N-n} \quad , \tag{C1}$$

where we have defined

$$S_K \equiv \sum_{n=0}^{k} \frac{1}{n!} e^{-n\Phi_1^{\rm ex}/k_B T} \tag{C2}$$

as a factor that arises from a sum over values of $n =$

$0, \ldots, K$, which is evaluated here using an upper limit $K = N - n$. The probability of finding exactly one cluster of size n within the simulation is proportional to Ξ_n. In the absence of the constraint on N, the sum over n could be extended to infinity, replacing S_{N-n} by a factor S_∞. Because the value of S_∞ is independent n, and also independent of the presence or absence of any large cluster, the existence of such a common prefactor would have no effect on the normalized probabilities of finding cluster of various sizes. A similar situation is recovered in a system in which N is constrained for values of n sufficiently far below N, for which $S_{N-n} \simeq S_\infty$. A signficant error occurs only when $N - n$ is small enough so that S_{N-n} differs signficantly from S_∞. This error occur whenever $N - n$ is not much greater than $e^{-\Phi_1^{\rm ex}/k_B T}$, which is the average number of unimers that would be present in the the absence of a constraint on N. When this occurs, the constraint on the range of allowed values of n decreases Ξ_n, thus decreasing the probability of finding a cluster of size n, and thereby increasing the value of $\Phi_n^{\rm ex}$ that we infer from the simulation.

The above analysis suggests a simple method to correct this error. In Eq. (C1), the introduction of a constraint on N reduces Ξ_n by a factor S_{N-n}/S_∞, corresponding to an increase in the corresponding free energy by an amount

$$\Delta\Phi_n^{\rm ex} = -k_B T \ln\left(\frac{S_{N-n}}{S_\infty}\right) \quad . \tag{C3}$$

The error caused by proximity of n to the upper bound can thus be corrected simply by subtracting the above expression from $\Delta\Phi_n^{\rm ex}$ from results for $\Phi_n^{\rm ex}$ for n near the maximum allowed value of N. This correction has applied whenever needed to obtain accurate values of $\Phi_n^{\rm ex}$ and W_n in all plots of these quantities.

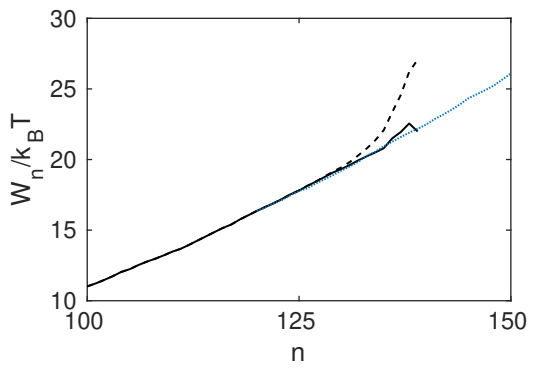

FIG. 10. Comparison of corrected and uncorrected results for the standard micelle formation free energy W_d vs. n for $\alpha = 12$. The dashed line shows uncorrected results obtained from simulations in which the simulation cell was constrained to have no more than $N = 140$ copolymer molecules, which show the appearance of an artifact for n near this limiting value. The solid line is the result obtained after applying the correction given in Eq. (C3). The lighter blue line is the result obtained from a simulation with higher limit of $N = 160$, for which no artifact appears in the range of n shown here.

[1] J. A. Mysona, A. V. McCormick, and D. C. Morse, Phys. Rev. E. **99** (2019).
[2] E. A. G. Aniansson and S. Wall, Journal of Physical Chemistry **78**, 1024 (1974).
[3] E. A. G. Aniansson and S. Wall, Journal of Physical Chemistry **79**, 857 (1975).
[4] E. Aniansson, S. Wall, M. Almgren, H. Hoffmann, I. Kielmann, W. Ulbricht, R. Zana, J. Lang, and C. Tondre, Journal of Physical Chemistry **80**, 905 (1976).
[5] I. Griffiths, C. Breward, D. Colegate, P. Dellar, P. Howell, and C. Bain, Soft Matter **9**, 853 (2013).
[6] I. Nyrkova and A. Semenov, Macromolecular Theory and Simulations **14**, 569 (2005).
[7] T. Nicolai, O. Colombani, and C. Chassenieux, Soft Matter **2010**, 3111 (2010).
[8] L. Meli, J. Santiago, and T. Lodge, Macromolecules **43**, 2018 (2010).
[9] L. Meli, J. M. Santiago, and T. P. Lodge, Macromolecules **43**, 2018 (2010).
[10] S. Choi, T. Lodge, and F. Bates, Physical Review Letters **104**, 047802 (2010).
[11] Z. Li and E. Dormidontova, Macromolecules **43**, 3521 (2010).
[12] Z. Li and E. Dormidontova, Soft Matter **7**, 4179 (2011).
[13] A. Halperin and S. Alexander, Macromolecules **22**, 2403 (1989).
[14] F. M. Kuni, A. I. Rusanov, A. P. Grinin, and A. K. Shchekin, Colloid Journal **63**, 197 (2001).
[15] F. M. Kuni, A. I. Rusanov, A. P. Grinin, and A. K. Shchekin, Colloid Journal **63**, 723 (2001).
[16] J. C. Shelley and M. Y. Shelley, Current Opinion in Colloid and Interface Science **5**, 101 (2000).
[17] M. Sammalkorpi, M. Karttunen, and M. Haataja, J. Phys. Chem. B **111**, 11722 (2007).
[18] M. Velinova, D. Sengupta, A. V. Tadjer, and S.-J. Marrink, Langmuir **27**, 14071 (2011).
[19] C. Bruce, M. Berkowitz, L. Perera, and M. Forbes, J. Phys. Chem. B **106**, 3788 (2002).
[20] B. Aoun, V. K. Sharma, E. Pellegrini, S. Mitra, M. Johnson, and R. Mukhopadhyay, J. Phy. Chem. B **119**, 5079 (2015).
[21] M. Johnston, W. Swope, K. Jordan, P. Warren, M. Noro, D. Bray, and R. Anderson, J. Phy. Chem. B **120**, 6337 (2016).
[22] R. Anderson, D. Bray, A. D. Regno, M. Seaton, A. Ferrante, and P. Warren, J. Chem. Theory Comput. **14**, 2633 (2018).
[23] A. Vishnayakov, M.-T. Lee, and A. V. Neimark, Journal of Physical Chemistry Letters **4**, 797 (2013).
[24] B. Levine, D. LeBard, R. DeVane, W. Shinoda, A. Kohlmeyer, and M. Klein, Journal of Chemical Theory and Computation **7**, 4135 (2011).
[25] M. Jorge, Langmuir **24**, 5714 (2008).
[26] D. I. Kopelevich, A. Panagiotopoulos, and I. G. Kevrekidis, Journal of Chemical Physics **122**, 044908 (2005).
[27] A. Santos and A. Panagiotopoulos, Journal of Chemical Physics **144**, 044709 (2016).
[28] S. Jiao, A. Santos, and A. Panagiotopoulos, Fluid Phase Equilibria **470**, 126 (2018).
[29] R. Pool and P. G. Bolhuis, J. Phys. Chem. B **109**, 6650 (2005).
[30] A. Cavallo, M. Mueller, and K. Binder, Macromolecules **39**, 9539 (2006).
[31] J. Glaser, J. Qin, P. Medapuram, M. Mueller, and D. C. Morse, Soft Matter **8**, 11310 (2012).
[32] J. Glaser, J. Qin, P. Medapuram, and D. C. Morse, Macromolecules **47**, 841 (2014).
[33] J. Glaser, P. Medapuram, T. M. Beardsley, M. W. Matsen, and D. C. Morse, Physical Review Letters **113**, 068302 (2014).
[34] P. Medapuram, J. Glaser, and D. C. Morse, Macromolecules **48**, 819 (2015).
[35] T. Ghasimakbari and D. C. Morse, Macromolecules **51**, 2335 (2018).
[36] B. Mehlig, D. W. Heermann, and B. M. Forrest, Physical Review B **45**, 679 (1992).
[37] G. Torrie and J. Valleau, Journal of Computational Physics **23**, 187 (1977).
[38] F. M. Kuni, A. I. Rusanov, A. K. Shchekin, and A. P. Grinin, Russian Journal of Physical Chemistry **79**, 833 (2005).