Implicit Solvent Simulations of DPC Micelle Formation

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · SEPTEMBER 2005

Impact Factor: 3.3 · DOI: 10.1021/jp0516801 · Source: PubMed

3 AUTHORS, INCLUDING:

Themis Lazaridis
City College of New York
98 PUBLICATIONS **7,850** CITATIONS

SEE PROFILE

Implicit Solvent Simulations of DPC Micelle Formation

Themis Lazaridis,* Buddhadeb Mallik, and Yong Chen

Department of Chemistry, City College of New York/CUNY, 138th Street and Convent Avenue, New York, New York 10031

Received: April 2, 2005; In Final Form: May 28, 2005

The formation of micelles by dodecylphosphocholine (DPC) is modeled by treating the surfactants in atomic detail and the solvent implicitly, in the spirit of the EEF1 solvation model for proteins. The solvation parameters of the DPC atoms are carried over from those of similar atoms in proteins. A slight adjustment of the parameters for the headgroup was found necessary for obtaining an aggregation number consistent with experiment. Molecular dynamics simulations of 960 DPC molecules at different concentrations are used to obtain the aggregation number, the micelle size distribution, and the CMC. At 20 mM concentration we obtain an aggregation number of 53−56 and a CMC of 1.25 mM, values close to the experimental ones. At 100 mM the aggregation number increases to 90. Simulations of individual micelles of varying size show that the effective energy per surfactant molecule is initially a decreasing function of aggregation number but stabilizes at about 60 molecules. The van der Waals term and the desolvation of nonpolar groups contribute to micellization, whereas the desolvation of polar groups opposes it. From the difference between the effective energy and the free energy (calculated from the CMC), the translational and rotational entropy contributions to the free energy are estimated at about 7 kcal/mol per monomer. The micelles obtained here are more irregular than those obtained in explicit water simulations. This modeling approach allows the study of larger surfactant aggregates for longer times and the extraction of thermodynamic in addition to structural information.

Introduction

Surfactants are amphipathic molecules that lower the surface tension of water or the interfacial tension of oil−water systems by positioning at the interface. At higher concentrations they form self-assembled structures in solution, usually small spherical aggregates (micelles), in which the polar headgroups are exposed to water and the hydrophobic chains are buried in the interior. Micelle formation is driven by the effectively unfavorable interaction of nonpolar groups with water (hydrophobic effect).[1] Many important uses for surfactants can be found in technological areas such as coatings, emulsions, detergency, pharmacology, enhanced oil recovery, cosmetics, etc..[2,3] Surfactants are also important in structural biology and biophysics, where they are used for solubilization of membrane proteins or as membrane mimics. The fast tumbling of micelles allows the application of solution NMR to the structure determination of hydrophobic peptides and proteins solubilized in them.[4,5]

Micelles have been studied extensively.[6] Quasielastic light scattering[7] and small angle neutron scattering[8,9] have provided information on their size, shape, aggregation number, and polydispersity. Fluorescence quenching has also been used to determine the aggregation number.[10,11] FTIR provides information on the conformation of the hydrocarbon chains.[12] Calorimetry provides the enthalpy of micelle formation.[13] NMR relaxation provides information on the dynamics of the surfactant molecules in the micelle.[14,15] However, because micelles are fluid and do not possess a well-defined structure, no detailed, high-resolution picture is available.

Numerous theoretical treatments of micelle formation have been developed. Thermodynamic analysis shows that to obtain a finite micelle size, the standard free energy per molecule must exhibit a minimum as a function of aggregation number.[1,16] The minimum can be explained by the principle of "opposing forces", according to which the attractive forces arising from the hydrophobic interaction compete with the repulsive forces arising from electrostatic repulsion between the headgroups.[1,17] The minimum is more difficult to explain for nonionic surfactants.[18] Israelachvili et al. illustrated how the molecular geometry of the surfactant molecules dictates what kind of aggregate will form.[16,19] The organization of surfactants in micelles and vesicles has been studied by a statistical treatment on[20,21] and off [22,23] a lattice. Models that can predict micellar properties, such as the critical micelle concentration (CMC) and average aggregation number, based on the molecular properties of the constituent surfactants have been put forth.[24−27]

The above analytical theories rely on many assumptions about the structure of a micelle and the interactions between its constituents. Computer simulations of simple model surfactants provided a means to test the assumptions of the theoretical models. A large body of literature exists on lattice and off-lattice simulations of simple, idealized surfactant systems with reasonable but arbitrary interaction potentials [28−33] or restraints to mimic surfactant−solvent interactions.[34,35]

The first realistic simulations in explicit water were of sodium octanoate micelles.[36,37] Since then realistic, all-atom simulations have been published of a lysophosphatidylethanolamine micelle,[38] a *n*-decyltrimethylammonium chloride micelle,[39] SDS micelles,[40,41] dodecylphosphocholine micelles,[42,43] octyl glucoside micelles,[44,45] mixed micelles modeling human bile,[46] and octyl pentaoxyethylene glycol (C8E5) micelles.[47] Some simulations have been performed of peptides[48−50] and a glycolipid[51] in micelles. Spontaneous micelle formation has also been simulated in solution[52] and around a protein.[53]

* Address correspondence to this author. Phone: (212) 650-8364. Fax: (212) 650-6107. E-mail: tlazaridis@ccny.cuny.edu.

10.1021/jp0516801 CCC: $30.25
Published on Web 07/15/2005

The high computational cost of realistic, all-atom explicit solvent simulations hinders the study of large surfactant aggregates and the calculation of thermodynamic properties and ensemble averages, such as the CMC and the aggregation number. As a result, the above explicit-water MD simulations of micelles focused on the description of micelle structure. One direction for alleviating this problem is the development of coarse-grained models for the lipids and solvent. For example, Marrink et al.[54] describe DPC using five interaction sites, two for the headgroup and three for the alkyl tail, which interact with Lennard-Jones potentials of varying strength. The solvent is also treated using Lennard-Jones particles. Here we present an alternative approach, where the surfactant molecules are treated in atomic detail but water is treated implicitly. This is a level of representation not yet adequately explored. In addition to computational savings, this approach offers the possibility of obtaining thermodynamic information, such as the CMC and the micelle size distribution. Thus, this approach combines the strengths of the simplified model systems with the realism of the more recent studies by using physics-based interaction potentials. A recent report on cationic surfactants[55] is similar in spirit to the present work.

We illustrate this approach with dodecylphosphocholine (DPC) micelles. DPC is one of the two most widely used surfactants in structural biology studies (the other being SDS). The phosphocholine group is identical with that in phospholipids but the single hydrophobic tail of DPC leads to formation of micelles rather than bilayers. The CMC and the aggregation number at 20−22 °C have been found to be about 1.1 mM and 50−60, respectively.[56] Information on the dynamics of the phosphocholine groups from NMR is also available.[14] Several of the explicit water simulations cited above have been on DPC micelles.[42,43,49−51]

Thermodynamic Framework. Micelle formation is most adequately analyzed in terms of the multiple equilibrium formalism where each aggregate is considered a distinct species.[57,58] Here we follow the work of Tanford with one difference: instead of mole fractions as concentration units we use molarities because molarities make better connection to statistical thermodynamics[59] and allow a clear calculation of the translational entropy contributions.

The chemical potential of a surfactant molecule in an aggregate of s molecules is[60]

$$\mu_s = \mu_s^o + (RT/s)\ln(\rho_s) \tag{1}$$

where μ_s^o is the standard chemical potential and ρ_s is the molarity of s-aggregates. μ_s^o in eq 1 is the chemical potential at 1 M standard state. Here we neglect intermicellar interactions and thus omit any activity coefficients from eq 1.

At equilibrium the chemical potentials of a monomer in all aggregates are equal to each other and to the chemical potential of the monomer:

$$\mu_s = \mu_1 \tag{2}$$

From eqs 1 and 2 one can obtain the concentration of each s-aggregate (size distribution):

$$\rho_s = \rho_1^s \exp(-s(\mu_s^o - \mu_1^o)/RT) \tag{3}$$

where ρ_1 is the molarity of the monomers. For the concentration of different aggregates of size s and m one obtains:

$$\frac{\rho_m^s}{\rho_s^m} = \exp(-s{\cdot}m\,(\mu_m^o - \mu_s^o)/RT) \tag{4}$$

The standard free energy per monomer of forming an s-micelle is:

$$\Delta G_s^o = \mu_s^o - \mu_1^o \tag{5}$$

This standard free energy includes the change in interactions upon micelle formation (hydrophobic effect, headgroup repulsions, desolvation, etc.), the loss of translational and rotational entropy,[61] and possible changes in conformational entropy upon micelle formation. In an implicit solvent representation, the changes in interactions are represented by the change in effective energy[62]

$$\Delta G_s^o = \langle W\rangle_s - \langle W\rangle_1 - T\Delta S^{\text{trans/rot/conf}} \tag{6}$$

where $\langle W\rangle_s$ is the average effective energy of a detergent molecule in an s-micelle and $\langle W\rangle_1$ that of a monomeric detergent molecule. The change in effective energy is obtained directly from the simulations. The last term in eq 6 is only the part of the entropy due to surfactant degrees of freedom. The *solvent* entropy is included in the effective energy contribution. The standard state is implicit in eq 1. If molarities are used in eq 1, the standard state for eq 5 is 1 M and the translational entropy in eq 6 refers to a free monomer at 1 M concentration. A very small variation in ΔG_s^o is sufficient to generate a narrow size distribution.[63] For example, using eq 4, a 0.2 kcal/mol difference in energy per lipid favoring a 60mer over a 40mer would make the 60mer concentration 4 orders of magnitude higher.

The CMC does not have a universal definition. Roughly, it is the concentration at which micelles start forming. Above the CMC the monomer concentration does not vary much, i.e., at any point where micelles are present, the monomer concentration is approximately equal to the CMC.[1] A useful relationship between the CMC and $\Delta G°$ can be derived if it is assumed that the micelles are monodisperse:

$$s\mathrm{D} \rightarrow \mathrm{D}_s$$

i.e., s detergent monomers aggregate to form s-micelles. For polydisperse micelles s can be thought of as an average micelle size. The equilibrium constant for the micellization reaction is

$$K = \rho_s/\rho_1^s$$

The standard free energy of the above reaction is

$$s\Delta G^o = -RT\ln K = -RT(\ln \rho_s - s\ln \rho_1) \tag{7}$$

Let σ be the fraction of monomer in aggregate form: $\sigma = s\rho_s/\rho_{\text{tot}} \Rightarrow \rho_1 = (1-\sigma)\rho_{\text{tot}}$. Substituting into eq 7:

$$\frac{\Delta G^o}{RT} = \frac{s-1}{s}\ln \rho_{\text{tot}} + \ln(1-\sigma) + 1/s\,\ln(s/\sigma) \tag{8}$$

Below the CMC σ is essentially 0. At the CMC ($\rho_{\text{tot}} = \text{CMC}$) the concentration of micelles and σ start increasing. Thus, at the CMC we can set σ equal to some small value, e.g., 0.1. For large s and small σ, the last two terms in eq 8 are close to zero and

$$\Delta G^o \approx RT\ln(\text{CMC}) \tag{9}$$

The experimental CMC for DPC is 0.0011 M,[56] which gives $\Delta G°$(1 M standard state) = −4.14 kcal/mol at room temperature.

The optimal size can be determined from maximizing eq 3 (d ln ρ_s/ds = 0).[58] The result is:

$$s\frac{\mathrm{d}\Delta G_s^{\mathrm{o}}}{\mathrm{ds}} = RT \ln \rho_1 - \Delta G_s^{\mathrm{o}} \qquad (10)$$

To a first approximation ρ_1 = CMC and $\Delta G_s^{\mathrm{o}} = \Delta G° = RT$ ln(CMC), so the LHS of eq 10 is approximately zero, and ΔG_s^{o} is a minimum at the optimal size.

Methods

Potentials. All simulations in this study were performed using the CHARMM program[64] and the CHARMM27 all-hydrogen lipid parameters,[65] which have been successful in a number of studies of micelles and lipid bilayers.[40,43,66,67] The model for DPC resulted from a combination of POPC and SDS. The DPC potential could probably be simplified by elimination of the explicit hydrogens in the hydrophobic tail, but that would require reparametrization of the dihedral terms, which was beyond the scope of this work.

The solvation treatment is an extension of the EEF1 energy function for proteins.[62] EEF1 adds to the CHARMM polar hydrogen force field[68] an implicit solvation term that describes the "self-energy" of all atoms (the interaction of each atom with the solvent). In addition, it uses a distance dependent dielectric constant ($\epsilon = r$ in Å) for the electrostatic interactions and neutralizes the ionic side chains. The implicit solvation term has the form:

$$\Delta G^{\mathrm{slv}} = \sum_i \Delta G_i^{\mathrm{slv}} = \sum_i \Delta G_i^{\mathrm{ref}} - \sum_i \sum_{j \neq i} f_i(r_{ij}) V_j \qquad (11)$$

where $\Delta G_i^{\mathrm{slv}}$ is the solvation free energy of atom i, and r_{ij} is the distance between i and j. Equation 11 says that the solvation free energy of atom i is that in a small model system where the atom is fully exposed to solvent ($\Delta G_i^{\mathrm{ref}}$) minus the solvation free energy it loses due to the presence of surrounding atoms. The solvation free energy density is modeled as a Gaussian function

$$f_i(r) 4\pi r^2 = \alpha_i \exp(-x_i^2), \qquad x_i = \frac{r - R_i}{\lambda_i} \qquad (12)$$

where R_i is the van der Waals radius of i (1/2 of the distance to the energy minimum in the Lennard-Jones potential), λ_i is a correlation length (3.5 Å for most atoms), and α_i is a proportionality coefficient given by

$$\alpha_i = 2\Delta G_i^{\mathrm{free}}/\sqrt{\pi}\lambda_i \qquad (13)$$

where $\Delta G_i^{\mathrm{free}}$ is the solvation free energy of the free (isolated) atom i; $\Delta G_i^{\mathrm{free}}$ is close but not identical to $\Delta G_i^{\mathrm{ref}}$ and is determined by requiring that the solvation free energy of deeply buried atoms be zero. EEF1 uses a nonbonded cutoff of 9 Å with a switching function between 7 and 9 Å.

EEF1 is adapted to DPC molecules in the following way. The solvation parameters for methyl and methylene groups are carried over from the protein model. The presence of explicit hydrogens in these groups does not affect the solvation calculation. The solvation parameters are assigned to the carbon; the hydrogens neither have any solvation free energy nor affect the solvation free energy of surrounding atoms.

For the zwitterionic headgroup of DPC two parameter sets were tested. The first used the standard charge distribution from the CHARMM27 force field and assigned reference solvation free energies of −80 kcal/mol to the phosphate and choline groups. This set was found to give large favorable electrostatic interaction changes upon micelle formation and changes in effective energy that are too negative compared to what one expects from the CMC (see below). This parameter set was not pursued further. The second set follows the empirical procedure adopted by the protein EEF1 energy function and makes the phosphate and choline groups net neutral by adding or subtracting a full charge from the P atom or the N atom, respectively. Thus, the charge of N was changed from −0.6 to −1.6 and the charge of P from +1.5 to +2.5. The solvation parameters assigned to these groups were of the order of −20 kcal/mol, as in proteins (for example, $\Delta G_i^{\mathrm{ref}}$ and $\Delta G_i^{\mathrm{free}}$ for the lysine NH_3 group are equal to −20 kcal/mol). The exact values for $\Delta G_i^{\mathrm{ref}}$ and $\Delta G_i^{\mathrm{free}}$ of N and P were adjusted to (a) obtain zero solvation free energy for buried groups and (b) give an aggregation number close to experiment (see Results). Therefore, these parameters should be viewed as adjustable. As expected, the results are sensitive to the value of the headgroup solvation parameters. $\Delta G_i^{\mathrm{free}}$ of −20 kcal/mol gave aggregation numbers that were too large, and $\Delta G_i^{\mathrm{free}}$ of −30 kcal/mol gave aggregation numbers that were too small. The final parameters are $\Delta G_i^{\mathrm{ref}} = -13.0$ kcal/ mol for both N and P, $\Delta G_i^{\mathrm{free}} = -26.0$ for N and −24.0 kcal/mol for P. The phosphate oxygens were assigned zero solvation parameters.

Solution Simulations. To simulate finite concentrations there has to be a limited volume available to the molecules. We used the MMFP facility of CHARMM to constrain the molecules within spheres of various radii with force constant 5 kcal/mol/Å^2. All solution simulations contained 960 DPC molecules (58 560 atoms). Concentrations of 1, 10, 20, and 100 mM correspond to sphere radii of 725, 335, 266, and 155 Å, respectively. The simulations were carried out using the Verlet integrator, a 2 fs time step, and SHAKE to constrain the bond lengths involving hydrogen. The temperature was set at 298 K and the velocities were scaled if the average temperature deviated from 298 K by more than 5 K. Two sets of simulations were run. The first (R) started from "random" initial configurations, which, for convenience, were obtained by simulating a 120mer micelle at 800 K for 100 ps and replicating the system 8 times. This simulation time was sufficient to produce a uniform distribution of detergent molecules throughout the available volume. The second (M) started from eight 120mer micelles. The purpose of running two simulations with different initial conditions was to estimate the extent of equilibration in the simulations. The CPU time for these simulations was about 170 h per ns on a single 2 GHz AMD Athlon processor (the 1mM R simulation took less because there were fewer interactions within the cutoff distance).

In determining micelle statistics one needs a definition of what constitutes a micelle. We define a micelle as a cluster of surfactants each of which has its center of mass within Rc Å from another member of the cluster. We found that the value Rc = 15 Å gave results that agreed well with visual identification of micelles. Still, the program can be "fooled" if, for example, one surfactant sits between two micelles. In that case the entire complex may be considered one micelle. Such an occurrence can be spotted by considering results over a period of time.

For every snapshot we compute the number of monomers, the weight average aggregation number (N_w), and the

number average aggregation number (N_n). These are defined as follows:

$$N_n = \frac{\sum_{i>1} iH_i}{\sum_{i>1} H_i} \text{ and } N_w = \frac{\sum_{i>1} i^2 H_i}{\sum_{i>1} iH_i} \quad (14)$$

where H_i is the number of micelles of size i. The more these two aggregation numbers differ, the higher the polydispersity of the sample is.[57,69] Most experimental methods (such as sedimentation equilibrium) give the weight-average aggregation number. The micelle size distribution is given by the formula $P_i = \langle iH_i \rangle$.

Individual Micelle Simulations. We constructed individual micelles as suggested by Tieleman et al.[42] The first DPC molecule was built in an all-trans conformation and aligned with the *X* axis. Then a coordinate translation of 5 Å was performed along the *X* axis and the whole structure was copied and rotated around the *Z* axis. Next, a rotation was done around the *Y* axis and that molecule was replicated around the *Z* axis as before. The number of molecules decreases with θ until $\theta = 90$, where we have only one DPC molecule coinciding with the *Z* axis. This way a spherical micelle was created with an internal cavity radius of 5 Å. This internal cavity was necessary to avoid overlap van der Waals radii at the center of the micelle. Micelles were constructed with 60, 80, 100, and 120 molecules. The constructed micelles were subjected to a series of minimizations and constrained MD simulations with progressively softer constraints.[40] Simulations were also performed starting from models (40, 54, and 65 molecules) produced by explicit solvent simulations[42] (available at http://moose.bio.ucalgary.ca/People/Peter). The initial aggregates of 2−20 monomers were obtained by selecting neighboring monomers from larger equilibrated micelles.

Simulation of a micelle without any boundary conditions corresponds to infinite dilution (zero detergent concentration). In such simulations detergent molecules can detach and move away from the micelle. In the long time limit, all detergent molecules will fly away (no micelle is stable at zero concentration). To prevent this we constrained all molecules to lie within a sphere of 40 Å radius (using the GEO SPHERE command from the MMFP module of CHARMM). The dynamics simulations were performed using the Nose-Hoover thermostat set to 298 K and lasted 0.5 ns. No heating was done. The final 0.4 ns were used to calculate the average effective energies. The principal moments of inertia were calculated by applying the CHARMM command COOR ORIE twice. The CPU time for 0.5 ns of the 120mer simulation was 16.2 h on a single 3 GHz Xeon processor.

Results

Simulations of Detergent Solutions. We first simulated solutions of 960 detergent molecules. A finite concentration was specified by constraining the molecules in a sphere of the appropriate size. For example, to simulate a 20 mM detergent solution, we constrained the 960 molecules in a sphere of radius 266 Å. To estimate the extent of equilibration, two sets of simulations were run, one (R) starting from "random" initial configurations, and another (M) starting from eight 120mer micelles.

Figure 1 shows the evolution of N_w, N_n, and $N_{monomer}$ as function of time for the R run at 20 mM. The aggregation numbers and number of monomers seem to converge near the end of the 6.5 ns run. These final values are $N_w \sim 53$, $N_n \sim 35$, and $N_{monomer} \sim 60$. The M run lasted 15 ns. The results over the last 7.5 ns are shown in Figure 2. The average values of N_w, N_n, and $N_{monomer}$ over this period are approximately 56, 44, and 60. These values compare well with those from run R, although N_n is somewhat larger in the M run. The deviation of N_w from N_n shows that there is considerable polydispersity in the sample, a little more in the R run. This also can be seen in Figure 3, which shows the last snapshot from run R.

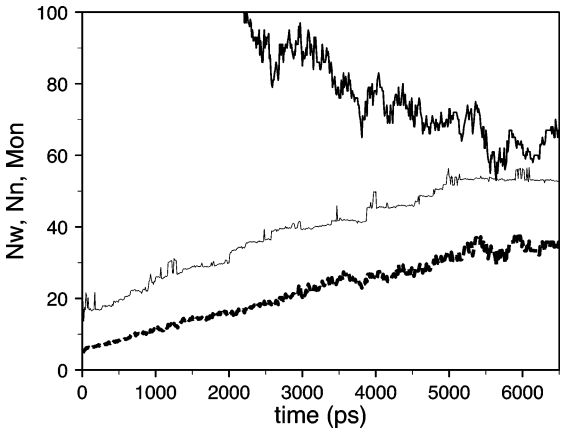

Figure 1. Aggregation numbers (N_n dashed line, N_w thin solid line) and number of monomers (thick solid line) as a function of time for the 20 mM R simulation.

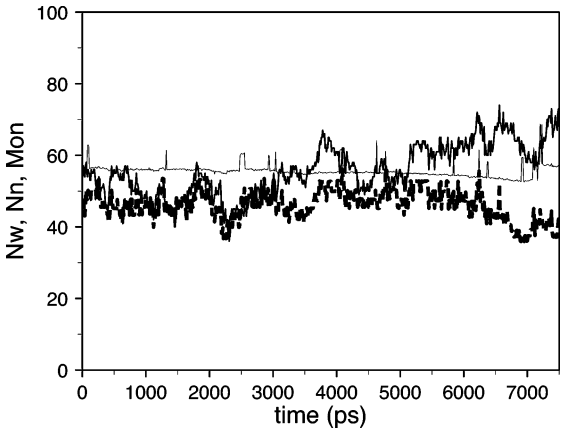

Figure 2. Aggregation numbers (N_n dashed line, N_w thin solid line) and number of monomers (thick solid line) as a function of time for the last 7.5 ns of the 20 mM M simulation.

It is usually observed that the monomer concentration does not vary much above the CMC. Therefore, at any point where micelles are observed, the monomer concentration should be close to the CMC. Using this, the predicted value of the CMC is about 1.25 mM, not far from the experimental value.[56]

Figure 4 shows the micelle size distribution obtained from the last 7.5 ns of the M run. The maximum at about 60 seems reliable but the fine details of the distribution seem to not have converged. For example, there is no justification for zero probability around 70. There just were not any 70mers in the simulation box during this period. Clearly, the size distribution is much harder to converge than the aggregation number. To sample all possible micelle sizes one would need either a larger system or a longer simulation.

Three other concentrations were simulated: 1, 10, and 100mM. At 1 mM, 22 ns of simulation starting from a random configuration produced no self-assembly. 14 ns of simulation starting from eight micelles led to dissolution of much of the

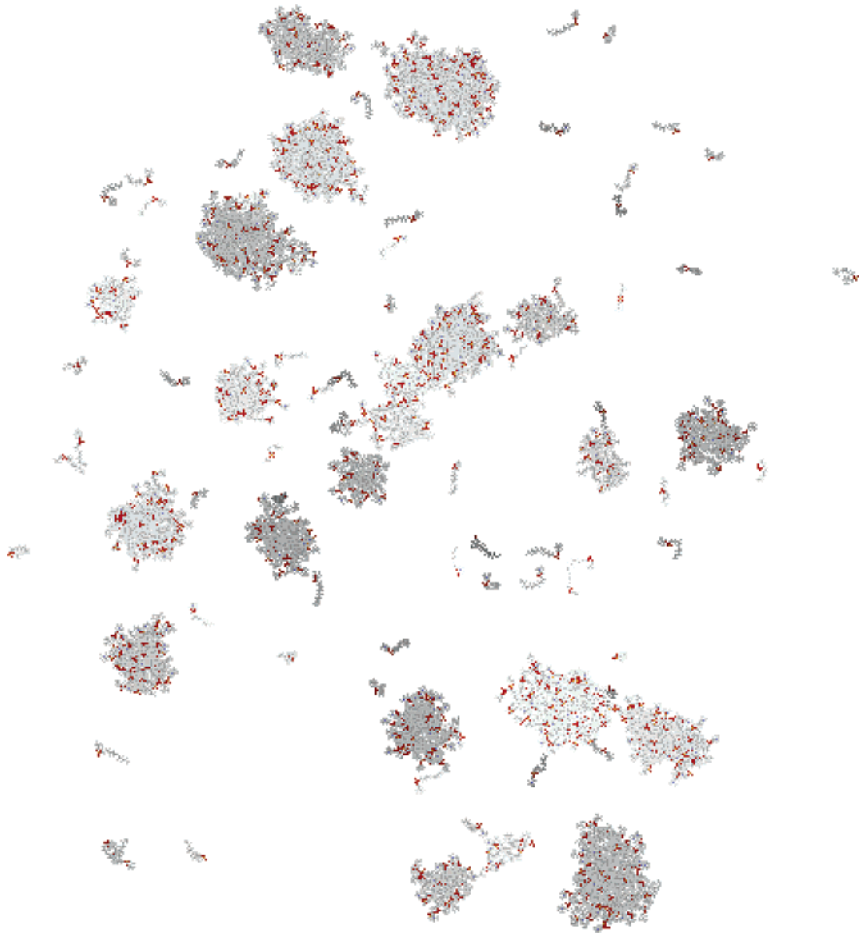

Figure 3. Last snapshot from the 20 mM R simulation.

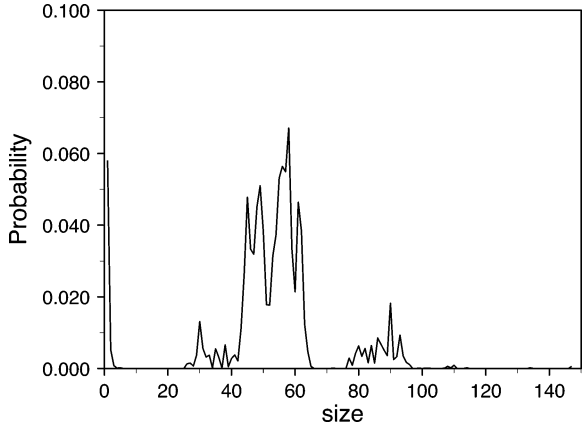

Figure 4. Micelle size distribution over the last 7.5 ns of the 20 mM M simulation.

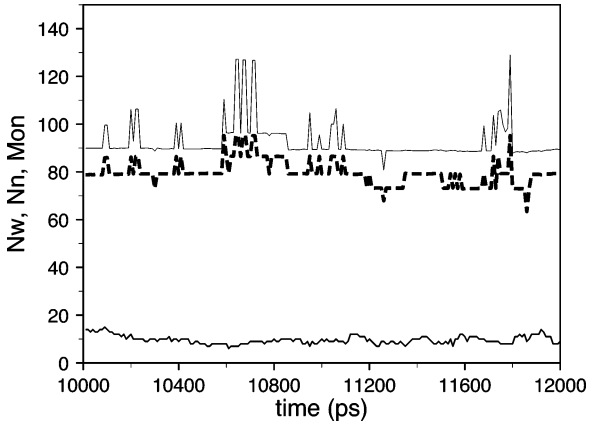

Figure 5. Aggregation numbers (N_n dashed line, N_w thin solid line) and number of monomers (thick solid line) as a function of time in the last 2 ns of the 100 mM R simulation.

micelles, but even after 14 ns some micellar remnants were still present. The trend in this simulation is toward increasing monomer concentration and decreasing N_w. Therefore, the tentative conclusion is that this concentration is below the CMC.

At 10 mM there is also not enough convergence between the M and R runs. After 12 ns the R run gave N_w around 35 with a slight upward trend. After 10 ns the M run gave $N_w = 104$ with a downward trend. This concentration is clearly above the CMC, but longer simulations would be needed to obtain reliable averages.

At 100 mM the R run after 12 ns gives $N_w \sim 90$, $N_n \sim 80$, and $N_{monomer} \sim 10$. Figure 5 shows the evolution of the average properties over the last 2 ns. The spikes in aggregation number occur when two micelles come close enough and are considered one micelle by the clustering algorithm. Figure 6 shows the last snapshot of this simulation. One can see the presence of more elongated micelles. The M run after 12 ns gave a larger aggregation number (118). The monomer concentration is 1.1 mM in the R run and 1.7 mM in the M run. The R run results are probably more reliable. Thus, the aggregation number increases substantially with overall detergent concentration.

Structure. Micelles are usually depicted as spherical but experimental information on their structure is limited. Quasielastic light scattering and sedimentation provide the diffusion coefficient and the molecular weight, from which the hydrodynamic radius and the aggregation number can be obtained. Such measurements for DPC provided an aggregation number

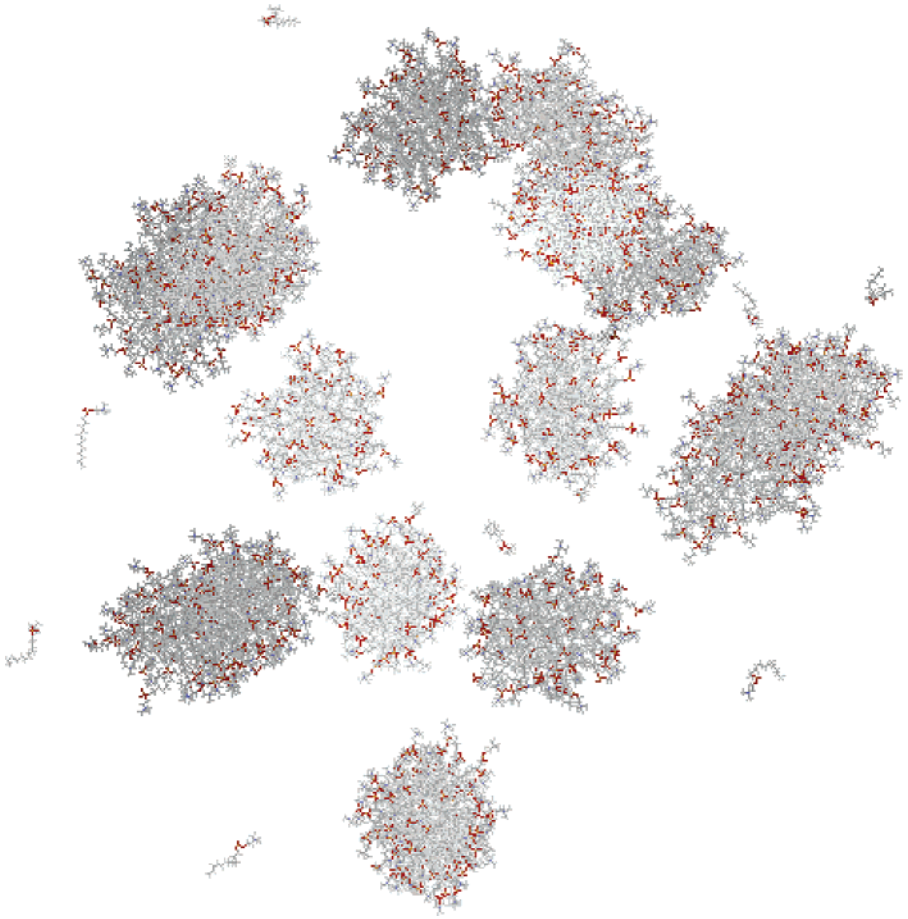

Figure 6. Last snapshot from the 100 mM R simulation.

of 56 and a radius of about 19.5 Å.[56] These data also gave a maximum asymmetry *a*/*b* of prolate ellipsoid of 6. However, the exact shape and regularity of the micellar surface are unknown.

Figure 7 shows snapshots of four micelles from the 20 mM solution simulation. Their aggregation numbers range from 49 to 86 and the radii of gyration from 15.7 to 20.5 Å. The ratios of the moments of inertia vary from 1.17 to 3.0 and show that the shape of these micelles deviates significantly from spherical. Some can be described as oblate ellipsoids, while others have more complex shapes. The micelle structures obtained here seem to be more irregular than those obtained using explicit solvent simulations starting with regular, spherical micelles; in those studies the maximum ratio of moments of inertia observed was 1.4.[42,43] Perhaps the solvent friction present in explicit solvent simulations does not allow large configurational changes to take place during the limited simulation time.

Thermodynamics. To obtain information on the thermodynamics of micelle formation, we performed additional simulations of individual micellar aggregates. In these simulations the appropriate number of detergent molecules is constrained in a sphere just large enough to contain them. For most micelles the sphere had radius 40 Å, for the 5mer 20 Å, and for the 2mer 10 Å.

The average effective energy per monomer in the micelle simulations is reported in Table 1 and plotted as a function of aggregation number in Figure 8. As expected, this quantity decreases as aggregate size increases up to about 65 monomers. Beyond that point, the energy stays approximately constant. The van der Waals energy is the main contributor to the decrease in effective energy. The electrostatic energy decreases slightly, and

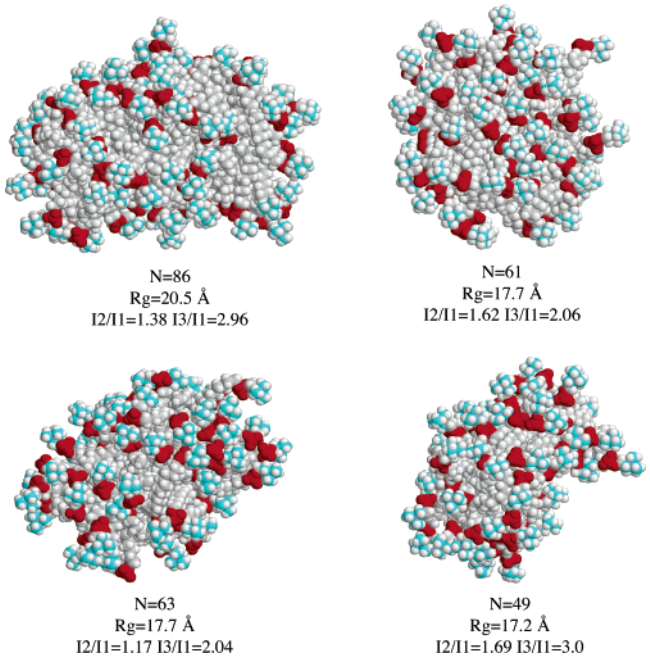

Figure 7. Individual micelles from a snapshot of the 20 mM M simulation. Red is the phosphate group and cyan the choline group. Aggregation numbers, radii of gyration, and ratios of moments of inertia are given beneath each picture.

the solvation free energy increases with size. The solvation free energy contains contributions from both hydrophobic and polar groups but the latter dominate. For example, the −8.14 kcal/mol per lipid for the 120mer is the sum of about −13.3 kcal/mol for the polar atoms and +5.2 for the nonpolar atoms. The

TABLE 1: Average Effective Energy Per Lipid (*W*) and Its van der Waals, Electrostatic, and Solvation Contributions in DPC Micelles of Different Size (kcal/mol)[a]

size	*W*	van der Waals	electrostatic	solvation
1	59.0 ± 0.2	−1.7 ± 0.1	12.3 ± 0.2	−13.8 ± 0.1
2	55.3 ± 0.2	−6.1 ± 0.2	12.1 ± 0.2	−12.2 ± 0.2
5	53.4 ± 0.2	−8.5 ±0.2	11.9 ± 0.1	−11.4 ± 0.2
10	51.6 ± 0.1	−10.8 ±0.1	11.9 ± 0.1	−10.9 ± 0.1
20	50.1 ± 0.1	−12.7 ±0.1	11.6 ± 0.05	−10.3 ± 0.1
40	48.8 ± 0.07	−14.6 ±0.07	11.27 ± 0.04	−9.43 ± 0.05
54	48.6 ± 0.06	−14.99 ± 0.06	11.25 ± 0.03	−9.23 ± 0.05
60	48.36 ± 0.06	−15.38 ± 0.07	11.14 ± 0.03	−8.99 ± 0.05
65	48.00 ± 0.04	−15.84 ±0.05	10.94 ± 0.03	−8.88 ± 0.04
80	48.19 ± 0.04	−15.78 ±0.04	11.13 ± 0.02	−8.66 ± 0.03
100	47.63 ± 0.04	−16.53 ± 0.04	10.82 ± 0.03	−8.33 ± 0.04
120	47.94 ± 0.04	−16.47 ± 0.04	10.87 ± 0.02	−8.14 ± 0.03

[a] Error bars are twice the standard deviation of the mean.

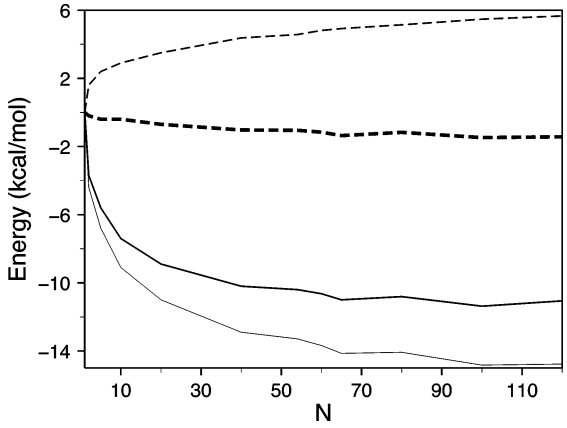

Figure 8. Average effective energy (thick solid line) and its components (thin solid: van der Waals; thick dashed: electrostatic; thin dashed: solvation) vs aggregation number for individual micelle simulations.

monomer has about −21 kcal/mol solvation free energy for the polar atoms and +7.2 kcal/mol for the nonpolar atoms. It is clear that the desolvation of the nonpolar groups favors micellization and that of the polar groups strongly disfavors it. The increase in solvation free energy is due to the higher desolvation of the polar headgroups as micelle size increases. The hydrophobic interaction in this breakup of the energetics is a combination of the van der Waals interactions of the nonpolar groups and the favorable desolvation of these groups. Clearly, the hydrophobic interaction is the driving force for micellization.

The effective energy of micelle formation is about −11 kcal/mol for micelles over ∼60mers. The value for the free energy of micelle formation inferred from the calculated CMC of ∼1.25 mM (eq 9) is −4 kcal/mol at 1 M standard state. This suggests that the entropic contributions (translational, rotational, and conformational entropy of the detergent molecules) are about +7 kcal/mol per detergent molecule at 1 M standard state.

Discussion

Large-scale simulations of a DPC solution were carried out to investigate micelle structure and the energetics of micelle formation. It is the first time that statistical properties, like the aggregation number, the size distribution, and the CMC, are obtained for realistic, all-atom models of surfactants. Self-assembly is readily observed at all concentrations except for 1 mM. The longest simulations were at 20 mM, the concentration used in the classic experiments of Lauterwein et al.[56] The simulations appear sufficient for convergence of the aggregation number, although still not long enough for full convergence of the size distribution. The aggregation number and CMC are in good agreement with experiment, although it is noted that the solvation parameters of the headgroups were slightly adjusted to achieve this agreement. Nevertheless, it is gratifying that both properties are reproduced simultaneously.

The polydispersity of DPC micelles, which is thermodynamically linked to the concentration dependence of the aggregation number,[57,69] seems to be insufficiently characterized experimentally. Based on the good agreement between the diffusion constants obtained by quasielastic light scattering and by ultracentrifugation, Lauterwein et al. concluded that the size distribution is narrow and independent of concentration between 5 and 100 mM.[56] However, they did not provide data at different concentrations. Another NMR study found diffusion coefficients, and thus aggregation numbers, similar to those of Lauterwein et al. at 98 mM concentration and 25 °C.[70] Another NMR study at 37 °C found an aggregation number of 44 at 228 mM pure DPC solution and that in the presence of peptide the aggregation number increases from 50 at 55 mM to 83 at 400 mM.[71] Clearly, more systematic studies are needed for both pure DPC and DPC−peptide solutions. In our calculations the "polydispersity index", defined as the ratio N_w/N_n, is about 1.27−1.5 at 20 mM and about 1.13 at 100 mM; increase in surfactant concentration leads to some increase in aggregation number. This makes physical sense, because the energy per lipid is approximately independent of size above ∼60. Increase in concentration reduces the translational entropy cost of adding monomers to micelles, and hence leads to larger micelles.

The physical forces behind micelle formation have been debated for a long time. A popular idea has been the principle of "opposing forces", whereby one force (the hydrophobic interaction) drives micelle growth and another force limits its growth to finite sizes. For ionic detergents such as SDS the opposing force was identified as the electrostatic repulsion between the headgroups. For nonionic detergents no such repulsion exists, and thus it has been more difficult to apply the above principle to this case. Although the desolvation of polar groups opposes micelle formation, other forces make up for the difference and the data in Table 1 fail to show a clear minimum in effective energy as micelle size grows. A similar observation was made on model nonionic surfactants.[33] The true opposing force must then be sought in the entropic terms. Larger micelle assemblies have lower entropy than small micelle assemblies. If there is no energetic advantage in larger micelles, they will not be formed. The entropic terms at 1 M standard state were estimated as 7 kcal/mol.

Although simulations in implicit solvent without friction overestimate the kinetics of diffusional processes, some information on micelle formation or dissolution dynamics can be obtained from such simulations. Micelle dissolution was observed to be quite slow. Even at 1 mM concentration, the initial micelles in the M run failed to completely dissolve in 14 ns of simulation. The barrier to this is the energy needed to detach one molecule from the micelle, which can be up to 11 kcal/mol. More realistic kinetic information can be easily obtained by using Langevin dynamics. This was not done here because the main focus in this work was on thermodynamics, i.e., time-independent properties.

The reliability of simulation results ultimately depends on the quality of the force field. In the explicit lipid-implicit water approach presented here, the intra- and interlipid interactions are treated with a state of the art nonpolarizable force field. Solvation is treated using an approximate, solvent-exclusion

model that seems to work well in proteins.[62] The largest uncertainty in the present approach is in the treatment of the zwitterionic headgroups. As in the protein EEF1, the phosphate and choline groups were given a zero net charge and solvation parameters that are much smaller than what is appropriate for charged groups. Alternatives would be the Generalized Born approximation[72,73] or the Screened Coulomb approach,[74] which, however, are also approximate. More rigorous treatments of charged group solvation such as the Poisson−Boltzmann equation[75] or the Protein Dipole Langevin Dipole method[76] do not easily lend themselves to MD simulations. The treatment of hydrophobic interactions is also approximate, but no practical alternatives seem to exist at present.

Compared to coarse-grained models of micelles and bilayers,[54,77] the present approach has a number of advantages: (a) the lipid representation and the interactions are more physically realistic, (b) the temperature dependence of the solvation potential also can be easily included, as in the protein EEF1, (c) the lack of explicit solvent is computationally expedient, and (d) extension to protein−detergent systems is straightforward. On the other hand, the coarse grained description of lipids allows the study of larger aggregates, lipid phase transitions, and nonlamellar phases. Application of the CG model to DPC micelle formation gave a good aggregation number, although the lack of any visible monomers in the simulation box (see Figure 12 of ref 54) suggests that the CMC may be too low.

A major motivation for this work has been the study of peptide and protein interactions with surfactants. Preliminary simulations of two membrane proteins, the transmembrane β barrel OmpA and the helical transmembrane dimer of glycophorin A, in a spontaneously assembled DPC micelle gave stable trajectories. It is feasible now to study mixed peptide−surfactant solutions at finite concentrations corresponding to experimental conditions and examine, for example, the effect of peptides on the CMC, or whether the aggregation number changes upon peptide solubilization, or upon association of two transmembrane helices. It is also possible to extend this approach to bilayer-forming lipids, such as DMPC, and model peptide interactions with bicelles and the effect of lipid phase curvature on peptide structure. That will allow us to address the long-standing question of how relevant micelles are as a mimic of biological membranes and the possible caveats of structural and thermodynamic studies of membrane proteins in micelles. Another possible application could be on the mechanism of protein denaturation by surfactants. It will be interesting to see whether this approach can be applied to ionic surfactants, such as SDS.

Acknowledgment. This work was supported by the National Science Foundation (MCB-0316667). Computational resources were provided by an RCMI grant from NIH (5G12RR003060) and by an NSF cooperative agreement (HRD-0206162) to the CREST Center for Mesoscopic Modeling and Simulation.

References and Notes

(1) Tanford, C. *The hydrophobic effect*: *Formation of micelles and biological membranes*, 2nd ed.; Wiley-Interscience: New York, 1980.

(2) *Industrial applications of surfactants*; Karsa, D. R., Ed.; Royal Society of Chemistry: Cambridge, UK, 1987.

(3) Attwood, D.; Florence, A. T. *Surfactant systems. Their chemistry, pharmacy, and biology*; Chapman and Hall: London, UK, 1983.

(4) Opella, S. J. *Nat. Struct. Biol.* **1997**, *4*, 845.

(5) Sanders, C. R.; Landis, G. C. *Biochemistry* **1995**, *34*, 4030.

(6) Moroi, Y. *Micelles. Theoretical and applied aspects*; Plenum Press: New York, 1992.

(7) Mazer, N. A.; Benedek, G. B.; Carey, M. C. *J. Phys. Chem.* **1976**, *80*, 1075.

(8) Bendedouch, D.; Chen, S.-H.; Koehler, W. C. *J. Phys. Chem.* **1983**, *87*, 153.

(9) Berr, S.; Jones, R. R. M.; Johnson, J. S. *J. Phys. Chem.* **1992**, *96*, 5611.

(10) Lianos, P.; Zana, R. *J. Phys. Chem.* **1980**, *84*, 3339.

(11) Croonen, Y.; Gelade, E.; van der Zegel, M.; van der Auweraer, M.; Vandendriessche, H.; De Schryver, F. C.; Almgren, M. *J. Phys. Chem.* **1983**, *87*, 1426.

(12) Holler, F.; Callis, J. B. *J. Phys. Chem.* **1989**, *93*, 2053.

(13) Sarmiento, F.; del Rio, J. M.; Prieto, G.; Attwood, D.; Jones, M. N.; Mosquera, V. *J. Phys. Chem.* **1995**, *99*, 17628.

(14) Beswick, V.; Guerois, R.; Cordier-Ochsenbein, F.; Coic, Y.-M.; Huynh-Dinh, T.; Tostain, J.; Noel, J.-P.; Sanson, A.; Neumann, J.-M. *Eur. Biophys. J.* **1998**, *28*, 48.

(15) Soderman, O.; Walderhaug, H.; Henriksson, U.; Stilbs, P. *J. Phys. Chem.* **1985**, *89*, 3693.

(16) Israelachvili, J. N. *Intermolecular and Surface Forces*; Academic Press: London, UK, 1985.

(17) Debye, P. *J. Phys. Chem.* **1949**, *53*, 1.

(18) Poland, D. C.; Scheraga, H. A. *J. Phys. Chem.* **1965**, *69*, 2431.

(19) Israelachvili, J. N.; Mitchell, D. J.; Ninham, B. W. *J. Chem. Soc., Faraday Trans. 2* **1976**, *72*, 1525.

(20) Dill, K. A.; Flory, P. J. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 676.

(21) Leermakers, F. A. M.; Scheutjens, J. M. H. M. *J. Colloid Interface Sci.* **1990**, *136*, 231.

(22) Ben-Shaul, A.; Szleifer, I.; Gelbart, W. M. *J. Chem. Phys.* **1985**, *83*, 3597.

(23) Guerin, C. B. E.; Szleifer, I. *Langmuir* **1999**, *15*, 7901.

(24) Goldstein, R. E. *J. Chem. Phys.* **1986**, *84*, 3367.

(25) Nagarajan, R.; Ruckenstein, E. *Langmuir* **1991**, *7*, 2934.

(26) Puvvada, S.; Blankschtein, D. *J. Chem. Phys.* **1990**, *92*, 3710.

(27) Reif, I.; Mulqueen, M.; Blankschtein, D. *Langmuir* **2001**, *17*, 5801.

(28) Larson, R. G.; Scriven, L. E.; Davis, H. T. *J. Chem. Phys.* **1985**, *83*, 2411.

(29) Owenson, B.; Pratt, L. R. *J. Phys. Chem.* **1984**, *88*, 2905.

(30) Smit, B.; Esselink, K.; Hilbers, P. A. J.; van Os, N. M.; Rupert, L. A. M.; Szleifer, I. *Langmuir* **1993**, *9*, 9.

(31) Bedrov, D.; Smith, G. D.; Freed, K. F.; Dudowicz, J. *J. Chem. Phys.* **2002**, *116*, 4765.

(32) Floriano, M. A.; Caponetti, E.; Panagiotopoulos, A. Z. *Langmuir* **1999**, *15*, 3143.

(33) Wijmans, C. M.; Linse, P. *Langmuir* **1995**, *11*, 3748.

(34) Karaborni, S.; O'Connell, J. P. *J. Phys. Chem.* **1990**, *94*, 2624.

(35) Haile, J. M.; O'Connell, J. P. *J. Phys. Chem.* **1984**, *88*, 6363.

(36) Watanabe, K.; Ferrario, M.; Klein, M. L. *J. Phys. Chem.* **1988**, *92*, 819.

(37) Jonsson, B.; Edholm, O.; Teleman, O. *J. Chem. Phys.* **1986**, *85*, 2259.

(38) Wendoloski, J. J.; Kimatian, S. J.; Schutt, C. E.; Salemme, F. R. *Science* **1989**, *243*, 636.

(39) Bocker, J.; Brickmann, J.; Bopp, P. *J. Phys. Chem.* **1994**, *98*, 712.

(40) MacKerell, A. D. *J. Phys. Chem.* **1995**, *99*, 1846.

(41) Bruce, C. D.; Berkowitz, M. L.; Perera, L.; Forbes, M. D. E. *J. Phys. Chem. B* **2002**, *106*, 3788.

(42) Tieleman, D. P.; van der Spoel, D.; Berendsen, H. J. C. *J. Phys. Chem. B* **2000**, *104*, 6389.

(43) Wymore, T.; Gao, X. F.; Wong, T. C. *J. Mol. Struct.* **1999**, *485*, 195.

(44) Bogusz, S.; Venable, R. M.; Pastor, R. W. *J. Phys. Chem. B* **2000**, *104*, 5462.

(45) Bogusz, S.; Venable, R. M.; Pastor, R. W. *J. Phys. Chem. B* **2001**, *105*, 8312.

(46) Marrink, S. J.; Mark, A. E. *Biochemistry* **2002**, *41*, 5375.

(47) Garde, S.; Yang, L.; Dordick, J. S.; Paulaitis, M. E. *Mol. Phys.* **2002**, *100*, 2299.

(48) Dixon, A. M.; Venable, R. M.; Pastor, R. W.; Bull, T. E. *Biopolymers* **2002**, *65*, 284.

(49) Wymore, T.; Wong, T. C. *Biophys. J.* **1999**, *76*, 1213.

(50) Gao, X.; Wong, T. C. *Biopolymers* **2001**, *58*, 643.

(51) Vasudevan, S. V.; Balaji, P. V. *J. Phys. Chem. B* **2001**, *105*, 7033.

(52) Marrink, S. J.; Tieleman, D. P.; Mark, A. E. *J. Phys. Chem. B* **2000**, *104*, 12165.

(53) Braun, R.; Engelman, D. M.; Schulten, K. *Biophys. J.* **2004**, *87*, 754.

(54) Marrink, S. J.; de Vries, A. H.; Mark, A. E. *J. Phys. Chem. B* **2004**, *108*, 750.

(55) Shinto, H.; Morisada, S.; Miyahara, M.; Higashitani, K. *Langmuir* **2004**, *20*, 2017.

(56) Lauterwein, J.; Bosch, C.; Brown, L. R.; Wuthrich, K. *Biochim. Biophys. Acta* **1979**, *556*, 244.

(57) Mukerjee, P. *J. Phys. Chem.* **1972**, *76*, 565.

(58) Tanford, C. *J. Phys. Chem.* **1974**, *78*, 2469.

(59) Ben-Naim, A. *J. Phys. Chem.* **1978**, *82*, 792.
(60) Ben-Shaul, A.; Gelbart, W. M. *J. Phys. Chem.* **1982**, *86*, 316.
(61) McMullen, W. E., III; Gelbart, W. M.; Ben-Shaul, A. *J. Phys. Chem.* **1984**, *88*, 6649.
(62) Lazaridis, T.; Karplus, M. *Proteins* **1999**, *35*, 133.
(63) Mukerjee, P. *J. Pharm. Sci.* **1974**, *63*, 972.
(64) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(65) Feller, S.; MacKerell, A. D., Jr. *J. Phys. Chem. B* **2000**, *104*, 7510.
(66) Feller, S. E.; Venable, R. M.; Pastor, R. W. *Langmuir* **1997**, *13*, 6555.
(67) Woolf, T. B. *Biophys. J.* **1998**, *74*, 115.
(68) Neria, E.; Fischer, S.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 1902.
(69) Nagarajan, R. *Langmuir* **1994**, *10*, 2028.
(70) Gao, X. F.; Wong, T. C. *Biophys. J.* **1998**, *74*, 1871.
(71) Kallick, D. A.; Tessmer, M. R.; Watts, C. R.; Li, C. Y. *J. Magn. Reson. Ser. B* **1995**, *109*, 60.
(72) Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem. A* **1997**, *101*, 3005.
(73) Lee, M. S.; Feig, M.; Salsbury, F. R., Jr.; Brooks, C. L., III *J. Comput Chem.* **2003**, *24*, 1348.
(74) Hassan, S. A.; Guarnieri, F.; Mehler, E. L. *J. Phys. Chem. B* **2000**, *104*, 6478.
(75) Honig, B.; Nicholls, A. *Science* **1995**, *268*, 1144.
(76) Russell, S. T.; Warshel, A. *J. Mol. Biol.* **1985**, *185*, 389.
(77) Shelley, J. C.; Shelley, M. Y.; Reeder, R. C.; Bandyopadhyay, S.; Klein, M. L. *J. Phys. Chem. B* **2001**, *105*, 4464.