

A Surface Site Interaction Point Method for Dissipative Particle Dynamics Parametrization: Application to Alkyl Ethoxylate Surfactant Self-Assembly.

Ennio Lavagnini, Joanne L. Cook, Patrick B. Warren, Mark J. Williamson and Christopher A. Hunter

ABSTRACT: Dissipative particle dynamics (DPD) is a coarse-grained approach to the simulation of large supramolecular systems, but one limitation has been that the parameters required to describe the non-covalent interactions between beads are not readily accessible. A first principles computational method has been developed so that bead interaction parameters can be calculated directly from ab initio gas phase molecular electrostatic potential surfaces of the molecular fragments that represent the beads. A footprinting algorithm converts the molecular electrostatic potential surfaces into a discrete set of surface site interaction points (SSIPs), and these SSIPs are used in the SSIMPLE (surface site interaction model for the properties of liquids at equilibrium) algorithm to calculate the free energies of transfer of one bead into a solution of any other bead. The bead transfer free energies are then converted into the required DPD interaction parameters for all pairwise combinations of different beads. The reliability of the parameters was demonstrated using DPD simulations of a range of alkyl ethoxylate surfactants. The simulations reproduce the experimentally determined values of the critical micelle concentration and mean aggregation number well for all 22 surfactants studied.

INTRODUCTION

The formation of supramolecular structures such as micelles, vesicles and bilayer membranes is a fundamentally important process in biology and in industry, with many applications in health and personal care products. The self-assembly of surfactants is a complicated process, and despite the development of simple tools that can be used to predict some aspects of surfactant behavior based on chemical structure (e.g. the critical packing parameter and the hydrophilic lipophilic balance (HLB)), the development of new surfactant systems still relies on experimental screening. The key parameters are the critical micelle concentration (CMC), which is the concentration at which surfactants start to aggregate into supramolecular structures (micelles), and the mean aggregation number ($N_{agg}$), which is the average number of surfactants in a micelle. Common techniques used to measure the CMC are surface tension, dynamic light scattering (DLS), fluorescence, UV-Vis and NMR spectroscopy, and in the case of charged species, conductometry and capillary electrophoresis. Methods for determining the $N_{agg}$ include DLS, small angle neutron diffraction, and time-resolved fluorescence quenching.

Many factors affect these measurements: for example, temperature, the presence of electrolytes, and organic impurities in the solution. The method of analysis also plays a role, because different methods are sensitive to different aspects of the self-assembled supramolecular structure. As a result, experimental screening of new formulations is time-consuming and expensive, and in silico prediction of surfactant self-assembly processes would be an attractive alternative. One computational approach is all-atom molecular dynamics (MD), but due to the long equilibration times and the large numbers of molecules required, coarse graining (CG) approaches are the method of choice. Coarse-graining combines several atoms or molecules into single CG beads, reducing the number of force centres and the associated computational cost.

Various CG methods have been used for the prediction of surfactant behaviour in solution, including Monte Carlo (MC) and lattice Boltzmann methods. Although recent progress in coarse grained molecular dynamics (CG-MD), using force-fields such as MARTINI shows promise, the size of the simulated systems using Lennard-Jones potentials with hard-core repulsions is still limited in length and timescale. One of the possibilities for further increasing the computational efficiency is the use of methods that consider only soft-core interactions, such as dissipative particle dynamics (DPD). Originally developed by Hoogerbrugge and Koelman and extended by Español and Warren, DPD uses soft-core beads, which move according to Newton’s equations of motion.

The total force acting on a DPD bead is the sum of a conservative force, a drag force, and a random force. The drag and random forces are used solely for thermostating (NVT ensemble). The conservative force derives from a short-range soft pair potential of the form

$$U_{ij} = \frac{1}{2} a_{ij} \left(1 - \frac{r_{ij}}{R_{ij}}\right)^2 \quad \text{for} \quad r_{ij} < R_{ij}, \quad \text{Eq. 1}$$

$R_{ij}$ is the minimum of the two radii of the potential spheres.
where \( a_d \) is the repulsion amplitude describing the interaction between bead \( i \) and bead \( j \), \( r_{ij} \) is the distance between the two beads, and \( R_s \) is the cut-off distance beyond which there is no interaction between the beads.

Since the conservative force accounts for the difference in non-covalent interactions between different beads, determination of the parameters in Equation 1 is critical for accurately reproducing the relationship between chemical structure and surfactant behavior.

**Approach**

Surface site interaction points (SSIPs) provide a quantitative description of all intermolecular interactions that a molecule can make with its environment. Molecules are represented as discrete sets of SSIPs of surface area 10 Å\(^2\) and volume 5 Å\(^3\) as illustrated in Figure 2. The number and properties of the SSIPs required to represent a specific molecule are calculated using the \textit{ab initio} molecular electrostatic potential surface and a footprinting algorithm.\(^{39}\) SSIPs can be used in the surface site interaction model for the properties of liquids at equilibrium (SSIMPLE) algorithm described in reference 38 to calculate solvation free energies and partition coefficients, which as we have outlined above can be used for estimating the \( a_d \) parameters required for DPD.

In SSIMPLE, the solvent and solute molecules are each described as an ensemble of SSIPs. The equilibrium constant for the pairwise interaction of any two SSIPs is given by

\[
K_{ij} = \frac{1}{2} e^{-\left(\epsilon_i + \epsilon_j + E_{vdw}\right)/RT} \tag{Eq. 2}
\]

where \( \epsilon_i \) and \( \epsilon_j \) describe the interaction properties of two SSIPs, and \( E_{vdw} \) is a constant that was estimated to be \(-5.6 \text{ kJ mol}^{-1}\) based on experimental data on the enthalpy change for the vapour-liquid equilibria of non-polar liquids as described in references 38 and 40.

The values of \( K_i \) can be used to determine the speciation of all possible SSIP contacts in a liquid phase. For a solute SSIP dissolved in a solvent S1, solving the set of simultaneous equations allows calculation of

**Figure 1.** Coarse grained representation of alkyl ethoxylate (C\(_n\)E\(_m\)) surfactants. The hydrocarbon chain is represented by a bead for the CH\(_3\) terminal group (purple) and beads for ethylene groups (red). The glycol chain is represented by a bead for the terminal alcohol moiety (blue) and beads for the CH\(_2\)OCH\(_2\) groups (green).

Much effort has gone into the development of systematic parametrization methods to obtain the DPD repulsion amplitudes \( a_d \). An early approach by Groot and Warren links \( a_d \) with Flory-Huggins parameters.\(^{25}\) This correlation has been used to obtain \( a_d \) from experimental solubility data.\(^{26}\) Hildebrand cohesive energy density parameters,\(^{27-28}\) and infinite dilution activity coefficients.\(^{29}\) Alternatively, \( a_d \) can be obtained by using experimental data on the mutual solubilities of small molecules. For example, using the CG approach for the alkyl ethoxylate surfactants shown in Figure 1, the interaction between the terminal alcohol bead and a bead in the ethylene glycol chain could be obtained using the mutual solubilities of methanol and methoxymethane. Computational approaches have also been investigated as an alternative to empirical parameterisation.\(^{30-32}\) A combination of MC and MD simulations was used to derive relationships between Flory-Huggins parameters and \( a_d \) values for interactions between different sized beads.\(^{33}\) COSMO-RS was used to calculate infinity dilution activity coefficients and 1-octanol/water partition coefficients in order to obtain DPD repulsion parameters.\(^{34-36}\)

However, estimation of repulsion amplitudes based on experimental or calculated molecular properties has some limitations. Since DPD beads typically represent a fragment of a molecule, in reality a portion of the molecular surface is buried by the overlap between covalently connected beads; this area should not be included when computing the bead interactions. Computational approaches provide an opportunity to remove the regions of bead overlap from a molecular surface. Saathoff used COSMO-SAC to delete parts of the molecular surface in the calculation of solvation free energies.\(^{37}\) An alternative approach to avoid this problem was adopted by Anderson \textit{et al}. who used DPD simulations of the partition coefficients of complete molecules to optimize the set of \( a_d \) values, fitting to experimentally data.\(^{38}\)

Here, we propose a new approach to the calculation of DPD bead repulsion amplitudes \( a_d \) using surface site interaction points (SSIPs).\(^{39}\) We develop the method in the context of CG models of alkyl ethoxylate surfactants shown in Figure 1. In this approach, different beads are used to represent chemical sub-groups containing between one and three heavy atoms as in Anderson \textit{et al}. This provides flexibility and allows straightforward extension to more complicated systems. We show by simulation that the \( a_d \) values thus obtained accurately reproduce the experimental CMC and \( N_{agg} \) values for this class of non-ionic surfactants.
where $[x_{\text{bound}}]$ is the concentration of SSIP $x$ that is bound to a solvent SSIP, and $[x_{\text{free}}]$ is the concentration of SSIP $x$ that is free. This represents the overall equilibrium constant for the interaction of SSIP $x$ with the solvent. The SSIP description of non-covalent interactions was parameterized using equilibrium constants for H-bonded complexes, which were measured at room temperature, and the approach has not yet been generalized to different temperatures, so this paper will focus on the room temperature behavior of surfactants.

Equating the concentrations of free SSIPs in two different phases allows calculation of the change in free energy ($\Delta G_i$) for moving an SSIP from solvent 1 to solvent 2, according to

$$\Delta G_{12}(x) = RT \ln \left( \frac{1 + K_{12}(x)}{1 + K_{21}(x)} \right)$$  \hspace{1cm} \text{Eq. 4}

where $\theta_1$ and $\theta_2$ are the fractional occupancies of the two phases, equal to the total SSIP concentration relative to the maximum possible SSIP concentration (300 M), $R$ is the gas constant and $T$ is the absolute temperature.

Equation 4 represents $\Delta G_i$ as the sum of a binding free energy, which accounts for the interactions the SSIP makes with the solvent SSIPs (first term), and a confinement free energy (second term). The confinement energy term is obtained by using an equilibrium constant of unity for all pairwise interactions in a phase and corrects for the difference in the probability of interaction associated with constraining the SSIPs to phases with different overall SSIP concentrations. For a solute molecule which is represented by multiple SSIPs, the free energy of transfer between two phases is calculated by summing the values of $\Delta G_i$ over all SSIPs. Similarly, it is possible to represent a DPD bead as a set of SSIPs and to calculate the free energies of transfer of beads between different liquid phases as the sum of the free energies of transfer of the individual SSIPs that represent the bead. These free energies of transfer are then used to obtain DPD repulsion amplitudes $a_{ij}$ as follows.

First, the Flory-Huggins $\chi$ parameter is calculated from the bead transfer free energy using Equation 5, which corrects for differences in the volumes of different beads.\(^{41}\)

$$\chi = \frac{1}{R} \frac{V_i}{V} \left( \frac{\Delta G_{12}}{V_1} + \frac{\Delta G_{21}}{V_2} \right)^{\frac{1}{2}}$$  \hspace{1cm} \text{Eq. 5}

where $V_1$, $V_2$, and $V$ are the van der Waals volumes calculated using the 0.002 electron/bohr\(^3\) electron density isosurface for two molecules of water, and the fragments that represent beads 1 and 2 respectively, and $\Delta G_i$ is the calculated change in free energy of transfer of one mole of bead 1 from a pure liquid composed of bead 1 to a dilute solution in a liquid composed of bead 2.

Second, the linear correlation proposed by Groot and Warren is used to obtain the actual DPD repulsion parameters using Equations 6 and 7.\(^{25}\)

$$\Delta a_{ij} = \frac{\chi}{c_\rho}$$  \hspace{1cm} \text{Eq. 6}

$$a_{ij} = \frac{a_0 + a_{ij}}{2} + \Delta a_{ij}$$  \hspace{1cm} \text{Eq. 7}

where $a_0$ and $a_{ij}$ are the self-interaction parameter for beads $i$ and $j$. One of the limitations of this method is that the interaction between beads of the same type, $a_0$, cannot be calculated. One approach is to use the same self-interaction parameter for all beads, and the water $a_0$ parameter derived from compressibility data is commonly used (25 kBT).\(^{27}\) However, this approximation assumes that all beads have the same volume,\(^{27}\) so we use self-interaction parameters reported by Anderson et al., which were tuned to match the experimental densities of selected molecular liquids.\(^{27}\)

**RESULTS AND DISCUSSION**

**DPD PARAMETERS**

The SSIP description of the four beads required for DPD simulations of the surfactants in Figure 1 is shown in Figure 3. SSIPs were calculated using methoxymethane for EO, ethane for C2, methane for T and ethanol for OH. To convert these molecular descriptions to bead descriptions, the SSIPs associated with the hydrogen atoms located at the points of covalent connectivity between beads were removed, i.e. the points indicated by dotted lines in Figure 3. The SSIP description of water has been reported previously,\(^{19}\) and the SSIP values used for all of the beads are summarised in Table 1. These values were used in SSIMPLE to calculate free energies of transfer for all pairwise combinations of beads (Table 1). A concentration of 1 mM was used for the solute beads to make sure that there are no solute-solute interactions, so the results are equivalent to the infinite dilution values required for Equation 5. The concentrations of the solvent beads were estimated based on structurally related liquids: the concentration of methanol was used as the bead concentration for a liquid composed of OH beads, half of the concentration of dimethoxyethane was used as the bead concentration for a liquid.
composed of EO beads, one quarter of the concentration of \( n \)-octane was used as the bead concentration for a liquid composed of C2 beads, and one eighth of the concentration of \( n \)-octane was used as the bead concentration for a liquid composed of T beads.

Table 1: SSIP representation of DPD beads and calculated free energies of transfer (\( \Delta G_{i,j} \) in kJ mol\(^{-1} \)).

| Bead | \( n(\text{Å}^3) \) | SSIP \( \epsilon_i \) | \( \Delta G_{T} \) | \( \Delta G_{C2} \) | \( \Delta G_{EO} \) | \( \Delta G_{OH} \) | \( \Delta G_{EO} \) |
|------|-----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| T    | 25.9            | 0.4, 0.4, 0.4, -0.3 | 0.00           | 0.40           | 0.00           | 2.40           | 18.38          |
| C2   | 38.9            | 0.4, 0.4, 0.4, -0.3 | -0.60          | 0.00           | 0.10           | 2.30           | 17.89          |
| EO   | 48.7            | 0.4, 0.4, 0.4, -5.3 | -1.14          | -0.04          | 0.00           | -1.52          | -6.17          |
| OH   | 34.7            | 2.7, 0.4, -5.3, -5.3 | 14.58          | 15.95          | 1.27           | 0.00           | -3.16          |
| W    | 42.0            | 2.8, 2.8, -4.5, -4.5 | 25.33          | 26.01          | -0.71          | 0.50           | 0.00           |

Figure 3. SSIP representations of the DPD beads used in this work. Positive and negative SSIPs are shown as red and blue dots respectively. The dotted lines indicate the point of covalent attachment to adjacent beads.

Table 2: DPD parameters for all pairwise bead interactions.

| Bead \( j \) | Bead \( i \) | \( a_{ij}^{36} \) | \( a_i \) | \( \Delta a_i \) | \( R_{ij}^{36} \) |
|-------------|-------------|----------------|--------|----------------|----------------|
| C2          | C2          | 22.0          | /      | /              | 1.074          |
| EO          | EO          | 25.5          | /      | /              | 1.116          |
| OH          | OH          | 14.0          | /      | /              | 0.980          |
| W           | W           | 25.0          | /      | /              | 1.000          |
| T           | T           | 24.0          | /      | /              | 0.955          |
| C2          | EO          | 23.8          | 0.03   | 1.095          |
| C2          | OH          | 27.1          | 9.13   | 1.027          |
| C2          | W           | 45.5          | 21.95  | 1.037          |
| EO          | OH          | 19.6          | -0.13  | 1.048          |
| EO          | W           | 21.8          | -3.44  | 1.058          |
| OH          | W           | 18.2          | -1.33  | 0.990          |
| T           | W           | 46.2          | 21.85  | 0.978          |
| T           | OH          | 27.5          | 8.49   | 0.968          |
| T           | C2          | 22.9          | -0.08  | 1.015          |
| T           | EO          | 24.2          | -1.05  | 1.036          |

The transfer free energies in Table 1 were used in Equations 5-7 to obtain the repulsion parameters listed in Table 2. For beads of different size, a volume correction is used in Equation 5. The van der Waals volume of the terminal beads T and OH was calculated by using half the calculated volume of ethane and 1,2-ethanediol, respectively. For the inner beads C2 and EO, the van der Waals volumes were obtained by calculating the volumes of homologous series of alkanes or ethylene glycols and taking the slope of a plot of volume versus number of bead repeats in the chain (see ESI). The van der Waals volumes and \( a_v \) values used in Equations 5-7 are reported in Tables 1 and 2.

The cross-interaction parameters \( \Delta a_{ij} \) in Table 2 show that the strongest repulsion, about 20 kJ mol\(^{-1} \), occurs between water and the two hydrocarbon beads C2 and T as expected. The values compare well with those reported in literature.\(^{25,28,37,44-46} \) The values of \( \Delta a_{ij} \) for the interactions of the terminal alcohol bead with the hydrocarbon beads, OH-C2 and OH-T, are approximately 9 kJ mol\(^{-1} \), which represents a significant repulsion. The OH bead is well solvated by itself due to H-bonding interactions between the hydroxyl groups, and these interactions are lost when this bead is transferred to a hydrocarbon solvent. The values are similar to those reported in literature: values of \( \Delta a_{ij} \) of 6 kJ mol\(^{-1} \) and 7 kJ mol\(^{-1} \) were used previously for interactions between OH-C2 and OH-T.\(^{37} \)

The values of \( \Delta a_{ij} \) for the interactions of the ethylene glycol bead with the hydrocarbon beads, EO-C2 and EO-T, are approximately zero. The major difference between the EO bead and the hydrocarbon beads is the presence of two H-bond acceptor sites on the EO oxygen (Table 1). However, none of these beads have H-bond donor sites, so there are no significant differences in the interactions the beads make with each other. These EO-hydrocarbon cross-interaction parameters in Table 1 differ from those reported in the literature. Because of the limited experimental data on the mutual solubilities of alkanes and oligoethylene glycols, \( \Delta a_{ii} \) was originally estimated to be intermediate in value between the hydrocarbon-hydrocarbon and hydrocarbon-water parameters, and a value of 6.5 kJ mol\(^{-1} \) was used in the literature.\(^{26} \) Attempts to calculate this parameter using COSMO-RS were inconclusive, due to a strong dependence on conformation.\(^{39} \) Anderson et al. used experimental water-octanol partition coefficients to obtain a value of 3.1 kJ mol\(^{-1} \) for the EO-hydrocarbon cross-interaction parameter, which is closer to the values in Table 2.\(^{37} \)

The values of \( \Delta a_{ij} \) for the interactions of the terminal alcohol and ethylene oxide beads with water, W-OH and W-EO, are both negative. The miscibility of water with alcohols and with oligoethylene oxides makes it impossible to determine the values of \( \Delta a_{ij} \) from ex-
perimental solubility measurements. Extrapolation from high temperature measurements on polyethylene oxide-water mixtures gave a value of $\Delta a_{EO}$ of 1.0–1.5 kT for the W-EO interaction. Anderson et al. used experimental logP measurements to obtain a value of $-1.25$ kT for the W-EO interaction, which is consistent with the negative value we calculate. The origin of the negative values of $\Delta a_{EO}$ in the SSIMPLE calculation is due to the fact that both alcohols and ethylene glycols are H-bond acceptors, and so the H-bonds formed with the water H-bond donors are favorable in the mixtures.

**DPD SIMULATIONS**

To test the parametrisation, DPD simulations were performed on linear, non-ionic surfactants belonging to the alkyl ethoxylate family shown in Figure 1; these are generically denoted as $C_mE_n$, where $m$ is the number of carbon atoms in the hydrocarbon chain and $n$ is the number of oxygens in the ethylene glycol chain. In the literature, these compounds are generally described using two types of bead,26,29,43–47 one for the hydrophobic tail and the other one for the hydrophilic head group, but here we use the more detailed CG description shown in Figure 1.37 The solvent beads used in the simulations were each composed of two molecules of water. The value of $R_b$ for the interaction of two water beads was set equal to $r_b$, the length unit in DPD, which is in turn set using the mapping number identity proposed by Groot and Rabone:26 $\rho N_o v_m = 1$, where $p$ is the bead density, $N_o$ (the ‘mapping number’) is the number of water molecules in a water bead, and $v_m = 30 \, \text{Å}^3$ is the molecular volume of water. With $N_o = 2$ and dimensionless bead density $\rho r_b^3 = 3$, this means that $r_b = 5.64$ Å. The correlated cut off distances for all other bead-bead interactions ($R_b$) were taken from Anderson et al.37,42

In our CG representation of these surfactants (Figure 1), the DPD molecules are linear chains with a stiff harmonic spring potential,

$$U(r) = \left(\frac{1}{2}\right) k_b (r_{ij} - r_0)^2$$

between connected pairs of beads. We set $k_b = 150$ kT and choose $n$ to match (approximately) the physical bond lengths as described below. To provide additional chain rigidity we supplement this two-body spring potential with a three-body angular potential,

$$U(\theta) = \left(\frac{1}{2}\right) k_\theta (\theta_{ij} - \theta_0)^2$$

where $k_\theta = 5$ kT and $\theta_0 = 180^\circ$.

For two covalently bonded C2 beads, the equilibrium distance $n$ was set at 0.39$n_0$ which gave an average distance during the simulation of 0.45$n_0$, equivalent to a bond length of 2.55 Å.37 The other equilibrium distances ($n_b$) between covalently bonded beads were modified by 0.1$r_b$ for each heavy atom added or deleted relative to C2, which gives 0.29$r_b$ for C2-T, 0.49$r_b$ for C2-EO, 0.59$r_b$ for EO-EO and 0.49$r_b$ for EO-OH as in Anderson et al.37,42

All the simulations were performed in a cubic box of side 40$n_0$, containing in total 192,000 beads. Box sizes from 20$n_0$ to 40$n_0$ were investigated, but no effect on the value of the calculated CMC was observed (see ESI). Simulations were run for $(2–4)\times 10^7$ timesteps with a timestep of 0.01 in DPD time units. In the literature, the DPD timescale in this kind of CG representation has been estimated using the diffusion of small molecules,39 so that one DPD timestep corresponds approximately to 0.5 ps, and our simulation runs correspond roughly to 1–2 μs. Simulations were performed using the DL_MESO DPD package (version 2.7),48 and analyzed using a combination of the UHMAP analysis tool49 and purpose written scripts. Trajectory files were collected every 500 timesteps, and the NVT ensemble was ensured by a commonly used DPD thermostat based on the standard velocity Verlet integration.50

DPD simulations were run at 4, 5 and 6% wt in water for all the $C_mE_n$ surfactants shown in Figure 1. Figure 4 shows examples of the results. Two molecules were considered part of the same supramolecular cluster, if they were closer than the cut-off distance $R_b$ and this criterion was used to calculate the number of ‘free’ surfactants as monomers or in submicellar aggregates for each timestep in a simulation. Figure 4a shows the free surfactant concentration for four different surfactants plotted against the number of timesteps. In all cases equilibrium was established between 1.0×10$^6$ and 2.5×10$^6$ timesteps. The cut-off distance criterion was also used to calculate the total number of molecules present within each supramolecular aggregate in each timestep of the simulation. The distribution of supramolecular assemblies was calculated for all timesteps greater than 5×10$^6$ and the populations were summed to obtain the aggregation number distribution (P(N)) for each simulation. Figure 4b shows an example of an aggregation number distribution for simulation of $C_{10}E_3$. There are clearly two populations. There are a large number of monomers that appear close to the origin, and then there is a clear gap in the trimer to 10-mer region, before a bell-shaped distribution appears with a maximum population around 20. The bell-shaped distribution corresponds to micellar aggregates of various sizes. The results from single simulations can be rather noisy, because the molecules tend to become kinetically trapped in the micellar assemblies. Figure 4b shows that the values of P(N) obtained from a simulation of $C_{10}E_3$ are not a simple function of N. It is possible to obtain a smoother aggregation number distribution by combining multiple simulations, and Figure 4c illustrates the result for $C_{10}E_3$.

**CRITICAL MICELLE CONCENTRATIONS**

The aggregation number distributions were used to calculate CMC values. The minimum number of molecules required for a cluster to be considered a micelle must first be assigned. A variable $N_{cm}$ is introduced such that clusters with aggregation numbers $N > N_{cm}$ are identified as micelles, and those with $N < N_{cm}$ as monomers and submicellar aggregates; the totality of the latter were used to define the free surfactant concentration. For the example shown in Figure 4b, there is a clear gap between the two populations, so choosing a value of $N_{cm}$ anywhere in the $5–10$ range will give the same result. In some simulations, there was overlap between the submicellar and micellar populations, but in these cases, it was still possible to identify a sparsely populated intermediate region, where the precise value of $N_{cm}$ had minimal effect on the calculated CMC values (see ESI). The values of CMC obtained from different simulations of the same surfactant were similar, leading to relatively well-defined values with small error bars. The surfactant concentration used in a DPD simulation can affect the calculated CMC value,35 but even for simulations run over a wider concentration range (1 to 15% wt of surfactant, see ESI), no significant variations in CMC were found.35
to the hydrocarbon chain

order of magnitude drop in CMC for every two CH
chains. The results are also quantitatively accurate reproducing the
tenations, but in all cases, there is excellent agreement between calcula-
tions, due to the small numbers of free surfactants in individual simula-
tions, 4%, 5% and 6% wt. The results are compared with the corre-
spanding experimentally determined values in Figure 5. The error
bars in the calculated CMC values are larger for the C12 surfactants,
due to the small numbers of free surfactants in individual simula-
tions, but in all cases, there is excellent agreement between calcula-
tion and experiment. The simulations clearly reproduce the trend that CMC is independent of \( n \), the length of the ethylene glycol
chain, but highly dependent on \( m \), the length of the hydrocarbon
chain. The results are also quantitatively accurate reproducing the
order of magnitude drop in CMC for every two CH\(_2\) groups added
to the hydrocarbon chain (Stauff-Kleevens rule).

Figure 4. a) The concentration of free surfactants plotted as a function of timestep for C\(_{6}\)E\(_{4}\) (green), C\(_{6}\)E\(_{8}\) (violet), C\(_{10}\)E\(_{8}\) (blue) and C\(_{12}\)E\(_{8}\) (yellow). The red dotted line shows the timestep at which the simulations were considered to have reached equilibrium. All simulations were run at 5%wt. b) The aggregation number distribution for a single simulation of C\(_{10}\)E\(_{8}\) at 5%wt. c) The aggregation number distribution for C\(_{10}\)E\(_{8}\) calculated by summing the results from nine simulations run at concentrations of 4%, 4.25%, 4.5%, 4.75%, 5.0%, 5.25%, 5.5%, 5.75%, 6.0% wt.

CMC values were calculated for all 22 surfactants by averaging the results from the three different concentrations used in the simulations, 4%, 5% and 6% wt. The results are compared with the corresponding experimentally determined values in Figure 5. The error bars in the calculated CMC values are larger for the C12 surfactants, due to the small numbers of free surfactants in individual simulations, but in all cases, there is excellent agreement between calculation and experiment. The simulations clearly reproduce the trend that CMC is independent of \( n \), the length of the ethylene glycol chain, but highly dependent on \( m \), the length of the hydrocarbon chain. The results are also quantitatively accurate reproducing the order of magnitude drop in CMC for every two CH\(_2\) groups added to the hydrocarbon chain (Stauff-Klevens rule).

Figure 5. CMC values for C\(_{m}\)E\(_{n}\) surfactants plotted as a function of \( n \). The lines represent the experimental values for each family of surfactants with the same value of \( m \) (blue \( m = 6 \), red \( m = 8 \), green \( m = 10 \), purple \( m = 12 \)) and the empty diamonds are the values calculated from DPD sim-
ulations.

MEAN AGGREGATION NUMBERS

For each timestep in a DPD simulation, the mean aggregation number (\( N_{agg} \)) can be calculated using

\[
N_{agg} = \frac{\sum_{N>N_{cut}} N^2 P(N)}{\sum_{N>N_{cut}} NP(N)}
\]

Eq. 10

Some authors use different \( N_{agg} \) numbers to determine the values of \( N_{agg} \) and the CMC from the same \( P(N) \) distribution, but here we used the same \( N_{agg} \) for calculation of both values (Table 3).

Consistent values of \( N_{agg} \) were obtained for simulations at different surfactant concentrations between 2 and 10% wt, and the results in Table 3 are quoted as the range of \( N_{agg} \) values obtained from individual simulations. All simulations performed in this work gave spherical micelles with no significant population of other morphologies in agreement with simulations and experimental reports in the litera-
ture.

Figures 6a-d and 6f-i show snapshots illustrating the evolution of the micelle structure with time for C\(_{6}\)E\(_{4}\) and C\(_{8}\)E\(_{4}\). The timescale for equilibration of the \( N_{agg} \) shown in Figures 6e and 6j is much longer than the timescale for equilibration of the monomer concentration shown in Figure 6a. The reason is that the equilibration of micelle size must take place via exchange of molecules between micelles via monomers in solution, and both the number of micelles and the dissociation rate from a micelle are low. Figures 6e and 6j illustrates the effect of surfactant concentration on \( N_{agg} \) equilibration time. For C\(_{8}\)E\(_{4}\), all of the simulations converge to the same value of \( N_{agg} \) (about 50) after \( 3 \times 10^6 \) timesteps, but the more dilute solutions equilibrate more slowly, because there are fewer micelles. For C\(_{8}\)E\(_{8}\), equilibration is significantly slower, and after \( 4 \times 10^6 \) timesteps, the values of \( N_{agg} \) for different concentrations have not converged. This behaviour is illustrated in Table 3 where \( N_{agg} \) values are reported after \( 2 \times 10^6 \) timesteps for all simulations and after \( 4 \times 10^6 \) timesteps for selected examples. The \( N_{agg} \) values for C\(_{8}\)E\(_{8}\) surfactants show no substantial variation between \( 2 \times 10^6 \) and \( 4 \times 10^6 \) timesteps, confirming that equi-
libration occurs rapidly. For C\(_{8}\)E\(_{8}\), the \( N_{agg} \) values increase slightly between \( 2 \times 10^6 \) and \( 4 \times 10^6 \) steps, but as shown in Figures 6e, 6f, the equi-
libration is reached by \( 4 \times 10^6 \) timesteps simulations. For C\(_{8}\)E\(_{8}\) and C\(_{12}\)E\(_{8}\),
there is a significant difference between the values of \( N_{agg} \) obtained at \(2 \times 10^6\) and \(4 \times 10^6\) timesteps. Figure 6j shows that the slow equilibration of the more hydrophobic surfactants means that the calculated values of \( N_{agg} \) are likely to be underestimated.\(^{45}\)

We note that in principle the mean aggregation number is an increasing function of concentration.\(^{61}\) This could account for the behavior seen here. Additionally, there are slow processes on ms time scales which are likely to be inaccessible in our \( \mu \)s simulations.\(^{62}\) Thus the number of micelles may be poorly equilibrated even if the free surfactant concentration has reached a steady state. This obviously impacts the calculation of the mean aggregation number.

Table 3 compares the calculated values of \( N_{agg} \) with the corresponding experimentally determined values reported by Swope et al.\(^{63}\) Experimental measurement of \( N_{agg} \) is not straightforward, and the application of different techniques to very broad distributions of micelle size can lead to discrepancies of an order of magnitude in reported values (see \( C_{10}E_5 \) and \( C_{12}E_5 \) in Table 3). However, the values of \( N_{agg} \) calculated for the fully equilibrated \( C_{10}E_n \) and \( C_{12}E_n \) surfactants agree well with the experimental ranges. For the surfactants with longer hydrocarbon chains, \( C_{10}E_n \) and \( C_{12}E_n \), the simulations consistently underestimate the value of \( N_{agg} \).

**Figure 6.** Evolution of micelle structure in DPD simulations. a-d) Snapshots of \( C_8E_4 \) simulation (5% wt) taken after \(5 \times 10^3, 2 \times 10^5, 10^6\) and \(4 \times 10^6\) timesteps. e) \( N_{agg} \) calculated for \( C_8E_4 \) plotted as a function of timestep at six different concentrations. f-i) Snapshots of \( C_{10}E_6 \) (5% wt) taken after \(5 \times 10^3, 2 \times 10^5, 10^6\) and \(4 \times 10^6\) timesteps. j) \( N_{agg} \) calculated for \( C_{10}E_6 \) plotted as a function of timestep at six different concentrations.
CONCLUSION

A method for calculating the bead interaction parameters required for dissipative particle dynamics (DPD) simulations has been developed. The method is based on ab initio calculation of the gas phase molecular electrostatic potential surfaces of the molecular fragments that represent the beads, so the approach should be generally applicable to the coarse graining of any molecular system using DPD. A footprinting algorithm was used to convert the molecular electrostatic potential surfaces into a discrete set of surface site interaction points (SSIPs), and these SSIPs were used in the SSIMPLE algorithm to calculate the free energies of transfer of one bead into a solution of any other bead. The bead transfer free energies were used to obtain the required DPD interaction parameters for all pairwise combinations of different beads. The reliability of this computational approach to determination of accurate DPD parameters was demonstrated using DPD simulations of a range of alkyl ethoxylate surfactants. The simulations reproduce the experimentally determined values of critical micelle concentration and aggregation number well for all 22 surfactants studied. The approach provides a powerful new tool for first principles calculation of DPD parameters and for prediction of the surfactant properties of molecules for which experimental data is not available.

Table 3: N_mg values from simulation and experiment.

| Surfactant | 2 × 10^6 timesteps | 4 × 10^6 timesteps | Experiment^† |
|------------|---------------------|---------------------|---------------|
| C6E3       | 27-31               | 28-33               | 24-57         |
| C6E4       | 22-24               | 21-23               | 21-55         |
| C6E5       | 19-21               | 19-23               | 8-12^4        |
| C6E6       | 63-76               | 43-47               | 23-147        |
| C8E5       | 34-39               | 31-36               | 17-90         |
| C8E6       | 31-36               | 33-41               | 30-41         |
| C8E7       | 29-33               | 28-34               | 72 65         |
| C10E3      | 55-71               | 58-77               | 100 66        |
| C10E4      | 40-44               | 45-57               | 17-172        |
| C10E5      | 31-41               | 32-54               | 66-105        |
| C10E6      | 28-36               | 29-39               | 46-70         |
| C10E7      | 85-126              | 30 65               |
| C12E3      | 58-76               | 44-53               | 112-4460      |
| C12E4      | 42-46               | 45-57               | 100-555       |
| C12E5      | 36-38               | 31-37               | 39-159        |

ASSOCIATED CONTENT

Supporting Information
Details of the fragment volume calculation, calculated aggregation number distributions for all surfactants, the effect of box size on calculated CMC values, and computational details. The Supporting Information is available free of charge on the ACS Publications website.

AUTHOR INFORMATION

Corresponding Author
* E-mail: herchelsmith.orgchem@ch.cam.ac.uk.

Notes
PBW declares a substantive (> $10k) stock holding in Unilever PLC. All other authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors acknowledge valuable discussions with Ian Stott and Massimo Noro, and thank Richard Anderson, David Bray and Michael Saxton for support with the UMMAP and DL_MESO packages. We thank the Engineering and Physical Sciences Research Council for financial support.

REFERENCES

(1) Kolesnichenko, I. V; Anslyn, E. V. Practical Applications of Supramolecular Chemistry. Chem. Soc. Rev. 2017, 46 (9), 2385–2390.
(2) Israelachvili, J. N.; Mitchell, D. J.; Ninham, B. W. Theory of Self-Assembly of Hydrocarbon Amphiphiles into Micelles and Bilayers. J. Chem. Soc. Faraday Trans. 2 Mol. Chem. Phys. 1976, 72 (0), 1525–1568.
(3) Griffin, W. C. Classification of Surface-Active Agents by “HLB.” J. Cosmet. Sci. 1949, 1 (5), 311–326.
(4) Eastoe, J.; Dalton, J. S.; Rogueda, P. G. A.; Crooks, E. R.; Pitt, A. R.; Simister, E. A. Dynamic Surface Tensions of Nonionic Surfactant Solutions. J. Colloid Interface Sci. 1997, 188 (2), 423–430.
(5) Zhao, H. W.; Huang, C. Z.; Li, Y. F. Validation of Micelle Formation of Proteins and Determination of Their Critical Micelle Concentrations by Measuring Synchronous Light Scattering Signals Using a Common Spectrofluorometer. Chem. Lett. 2006, 35 (4), 418–419.
(6) Zhang, X.; Jackson, K. J.; Burt, H. M. Determination of Surfactant Critical Micelle Concentration by a Novel Fluorescence Depolarization Technique. J. Biochem. Biophys. Methods 1996, 31 (3–4), 145–150.
(7) Hatt, S. K.; Moulik, S. P. Determination of Critical Micelle Concentration (CMC) of Nonionic Surfactants by Donor-Acceptor Interaction with Lixodine and Correlation of CMC with Hydrophile-Lipophile Balance and Other Parameters of the Surfactants. J. Surfactants Deterg. 2001, 4 (3), 303–309.
(8) Cerichelli, G.; Mancinit, G. NMR Techniques Applied to Micellar Systems. Curr. Opin. Colloid Interface Sci. 1997, 2 (6), 641–648.
(9) Scholz, N.; Behnke, T.; Resch-Genger, U. Determination of the Critical Micelle Concentration of Neutral and Ionic Surfactants with Fluorometry, Conductometry, and Surface Tension—A Method Comparison. J. Fluoresc. 2018, 28 (1), 465–476.
(10) Lin, C.-E. Determination of Critical Micelle Concentration of Surfactants by Capillary Electrophoresis. J. Chromatogr. A 2004, 1037 (1), 467–478.
(11) Sharma, K. S.; Patil, S. R.; Rakshit, A. K.; Glenn, K.; Doiron, M.; Palepu, R. M.; Hassan, P. A. Self-Aggregation of a Cationic–Nonionic Surfactant Mixture in Aqueous Media: Tensimetric, Conductometric, Density, Light Scattering, Potentiometric, and Fluorometric Studies. J. Phys. Chem. B 2004, 108 (34), 12804–12812.
(12) Oliver, R. C.; Lipfert, J.; Fox, D. A.; Lo, R. H.; Doniach, S.; Columbia, L. Dependence of Micelle Size and Shape on Detergent Alkyl Chain Length and Head Group. Proc. Natl. Acad. Sci. U. S. A. 2013, 110 (19), 7530–7534.
(32) Karac, G.; Peters, E. A. J. F.; de With, G. A Generalized Method for Parameterization of Dissipative Particle Dynamics for Variable Bead Volumes. EPL (Europhysics Lett.) 2013, 102 (4), 40009.
(33) Liyana-Arachchi, T. P.; Madamgini, S. N.; Eike, D.; Koenig, P. H.; Ilja Siepmann, J. Liquid-Liquid Equilibria for Soft-Repulsive Particles: Improved Equation of State and Methodology for Representing Molecules of Different Sizes and Chemistry in Dissipative Particle Dynamics. J. Chem. Phys. 2015, 142 (4).
(34) Vishnyakov, A.; Lee, M.-T.; Neimark, A. V. Prediction of the Critical Micelle Concentration of Nonionic Surfactants by Dissipative Particle Dynamics Simulations. J. Phys. Chem. Lett. 2013, 4 (5), 797–802.
(15) Ginde, R. M.; Myerson, A. S. Effect of Impurities on Cluster Growth and Nucleation. J. Cryst. Growth 1993, 126 (2), 216–222.
(16) Hansson, P.; Jönsson, B.; Ström, C.; Söderman, O. Determination of Micellar Aggregation Numbers in Dilute Surfactant Systems with the Fluorescence Quenching Method. J. Phys. Chem. B 2000, 104 (15), 3496–3506.
(17) Hadley, K. R.; McCabe, C. Coarse-Grained Molecular Models of Water: A Review. Mol. Simul. 2012, 38 (8–9), 671–681.
(18) Zehl, T.; Wahab, M.; Mögel, H.-J.; Schiller, P. Monte Carlo Simulations of Self-Assembled Surfactant Aggregates. Langmuir 2006, 22 (6), 2523–2527.
(19) Furtado, K.; Skartlien, R. Derivation and Thermodynamics of a Lattice Boltzmann Model with Soluble Amphiphilic Surfactant. Phys. Rev. E 2010, 81 (6), 66704.
(20) Taddei, M.; Carbone, P. Effect of Chain Length on the Partition Properties of Poly(Ethylene Oxide): Comparison between MARTINI Coarse-Grained and Atomistic Models. J. Phys. Chem. B 2017, 121 (7), 1601–1609.
(21) Velinova, M.; Sengupta, D.; Tadjev, A. V.; Marrink, S.-J. Sphere-to-Rod Transitions of Nonionic Surfactant Micelles in Aqueous Solution Modeled by Molecular Dynamics Simulations. Langmuir 2011, 27 (23), 14071–14077.
(22) Anogianakis, S. D.; Petris, P. C.; Theodorou, D. N. Promising Route for the Development of a Computational Framework for Self-Assembly and Phase Behavior Prediction of Ionic Surfactants Using MARTINI. J. Phys. Chem. B 2020, 124 (1), 556–67.
(23) Hooogenbrugge, P. J.; Koelman, J. M. V. A. Simulating Microscopic Hydrodynamic Phenomena with Dissipative Particle Dynamics. Europhys. Lett. 1992, 19 (3), 155–160.
(24) Espanol P. Warren, P. Statistical Mechanics of Dissipative Particle Dynamics. Europhys. Lett. 1995, 30, 191.
(25) Warren, P. B.; Groot, R. D. Dissipative Particle Dynamics: Bridging the Gap between Atomistic and Mesoscopic Simulation. J. Phys. Chem. B 1997, 101 (11), 4423–4435.
(26) Groot, R. D.; Rabone, K. L. Mesoscopic Simulation of Cell Membrane Damage, Morphology Change and Rupture by Nonionic Surfactants. Biophys. J. 2001, 81 (2), 725–736.
(27) Maiti, A.; McGrother, S. bead-Bead Interaction Parameters in Dissipative Particle Dynamics: Relation to Bead-Size, Solubility Parameter, and Surface Tension. J. Chem. Phys. 2004, 120 (3), 1594–1601.
(28) Davis, K. P.; Bankhead, M.; Good, K.; Owens, S. L. New Parameterization Method for Dissipative Particle Dynamics. J. Chem. Phys. 2007, 127 (1).
(29) Lee, M. T.; Vishnyakov, A.; Neimark, A. V. Calculations of Critical Micelle Concentration by Dissipative Particle Dynamics Simulations: The Role of Chain Rigidity. J. Phys. Chem. B 2013.
(30) Doi, H.; Okuwaki, K.; Mochizuki, Y.; Ozawa, T.; Yasuoka, K. Dissipative Particle Dynamics (DPD) Simulations with Fragment Molecular Orbital (FMO) Based Effective Parameters for 1-Palmitoyl-2-Oleoyl Phosphatidyl Choline (POPC) Membrane. Chem. Phys. Lett. 2017, 684, 427–432.
(31) Saathoff, J. Effectively Parameterizing Dissipative Particle Dynamics Using COSMO-SAC: A Partition Coefficient Study. J. Chem. Phys. 2018, 148 (15), 154102.
(53) Mai, Z.; Couallier, E.; Rakib, M.; Rousseau, B. Parameterization of a Mesoscopic Model for the Self-Assembly of Linear Sodium Alkyl Sulfates. J. Chem. Phys. 2014, 140(20), 204902.
(54) Fournial, A.-G.; Zhu, Y.; Molinier, V.; Vermeersch, G.; Aubry, J.-M.; Azaroual, N. Aqueous Phase Behavior of Tetraethylene Glycol Decanoyl Ester (C9COE4) and Ether (C10E4) Investigated by Nuclear Magnetic Resonance Spectroscopic Techniques. Langmuir 2007, 23(23), 11443–11450.
(55) Nishikido, N. Micellar Growth of Poly(Oxyethylene) Nonionic Surfactants with Increasing Temperature: Deduction from Critical Micellization Concentration-Temperature Relationships. Langmuir 1990, 6(7), 1225–1228.
(56) Stubenrauch, C.; Nydén, M.; Findenegg, G. H.; Lindman, B. NMR Self-Diffusion Study of Aqueous Solutions of Tetraoxyethylene n-Octyl Ether (C8E4). J. Phys. Chem. 1996, 100(42), 17028–17033.
(57) Krister Holmberg. Handbook of Applied Surface and Colloid Chemistry; John Wiley & Sons Ltd, 2002; pp Vo1, Chapter 13, p 293.
(58) Nilsson, P. G.; Wennerstroem, H.; Lindman, B. Structure of Micellar Solutions of Nonionic Surfactants. Nuclear Magnetic Resonance Self-Diffusion and Proton Relaxation Studies of Poly(Ethylene Oxide) Alkyl Ethers. J. Phys. Chem. 1983, 87(8), 1377–1385.
(59) Kato, T.; Terao, T.; Tsukada, M.; Seimiya, T. Self-Diffusion Processes in Semidilute Solutions of Nonionic Surfactant (C16E7) Studied by Light Scattering and Pulsed Gradient Spin Echo Methods. J. Phys. Chem. 1993, 97(15), 3910–3917.
(60) Sterpone, F.; Briganti, G.; Pierleoni, C. Sphere versus Cylinder: The Effect of Packing on the Structure of Nonionic C12E6 Micelles. Langmuir 2009, 25(16), 8960–8967.
(61) Nagarajan, R. On Interpreting Fluorescence Measurements: What Does Thermodynamics Have To Say about Change in Micellar Aggregation Number versus Change in Size Distribution Induced by Increasing Concentration of the Surfactant in Solution? Langmuir 1994, 10(6), 2028-2034.
(62) Mysona, J. A.; McCormick, A. V.; Morse, D. C. Mechanism of Micelle Birth and Death. Phys. Rev. Lett. 2019, 123(038003).
(63) Swope, W. C.; Johnston, M. A.; Duff, A. I.; McDonagh, J. L; Anderson, R. L.; Alva, G.; Tek, A. T.; Maschino, A. P. Challenge to Reconcile Experimental Micellar Properties of the CnEm Nonionic Surfactant Family. J. Phys. Chem. B 2019, 123(7), 1696–1707.
(64) Corkill, J. M.; Goodman, J. F.; Tate, J. R. Calorimetric Determination of the Heats of Micelle Formation of Some Non-Ionic Detergents. Trans. Faraday Soc. 1964, 60(0), 996–1002.
(65) Frindi, M.; Michels, B.; Zana, R. Ultrasonic Absorption Studies of Surfactant Exchange between Micelles and Bulk Phase in Aqueous Micellar Solutions of Nonionic Surfactants with a Short Alkyl Chain. 2. C6E3, C6E5, C8E4, and C8E8. J. Phys. Chem. 1992, 96(14), 6095–6102.
(66) Telgmann, T.; Kaatze, U. Monomer Exchange and Concentration Fluctuations in Poly(Ethylene Glycol) Monoalkyl Ether/Water Mixtures. Dependence upon Nonionic Surfactant Composition. J. Phys. Chem. A 2000, 104(21), 4846–4856.
(67) Olofsson, G. Micelle Formation in Non-Aqueous Solvents: Calorimetric Study of the Association of Poly(Ethylene Oxide) Alkyl Ethers and Hexadecyltrimethylammonium Bromide in Formamide. J. Chem. Soc. Faraday Trans. 1991, 87(18), 3037–3042.
