Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

Temperature-Dependent Atomic Models of Detergent Micelles Refined against Small-Angle X-Ray Scattering Data

Miloš T. Ivanović, Linda K. Bruetzel, Jan Lipfert, and Jochen S. Hub**

anie_201713303_sm_miscellaneous_information.pdf

Author Contributions

M.I. Data curation: Lead; Investigation: Lead; Writing – original draft: Equal
L.B. Formal analysis: Supporting; Investigation: Supporting
J.L. Conceptualization: Equal; Funding acquisition: Equal; Methodology: Equal; Supervision: Equal; Writing – review & editing: Supporting
J.H. Conceptualization: Equal; Funding acquisition: Equal; Software: Lead; Supervision: Equal; Writing – original draft: Lead.

METHODS

Analytic model for the aggregation number

Maibaum, Dinner, and Chandler have proposed an analytic model for the aggregation number of spherical detergent micelles,[1]

$$N_{\mathrm{agg}}^{\mathrm{ana}} = (49\pi/48)\beta\gamma\delta^2, \tag{S1}$$

where $\beta = (k_B T)^{-1}$ denotes the inverse temperature, and γ is the interfacial oil-water surface tension. Values for γ for DDM and DM at the different temperatures were taken from Ref. 2. δ is a length parameter of the alkyl chain of n_c carbon atoms. Because the exact value of δ for a given type of detergent is not obvious, we adjusted δ such that $N_{\mathrm{agg}}^{\mathrm{ana}}$ matched N_{agg} estimated by comparing the MD simulations with the experimental SAXS curves at 25°C (Fig. 2, black circles). This procedure yields $\delta = 17.8\,\text{Å}$ for DDM and $\delta = 14.5\,\text{Å}$ for DM, in reasonable agreement with the maximum length of the tails given by Tanford's formula,[3] $\delta = (0.15 + 0.1265 n_c)\,\mathrm{nm}$, where n_c denotes the number of carbon atoms in the tail (10 for DM and 12 for DDM), suggesting that the parameter δ was adjusted to physically reasonable values (see also Fig. S10 below).

SAXS experiments

Experimental data were collected at beam line 12ID of the Advanced Photon Source[4] (APS), essentially as described previously (Table S1).[5–7] Measurements used custom-made sample

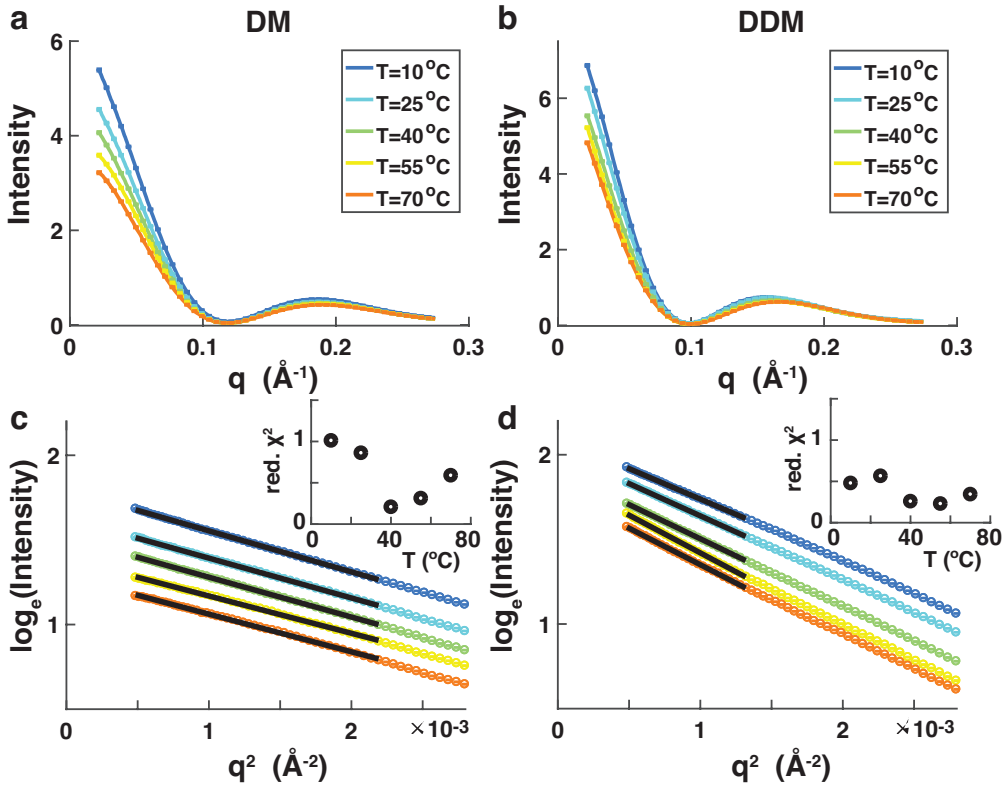


Figure S1: SAXS data and Guinier analysis for DM and DDM micelles as a function of temperature. a), b) Scattering profiles recorded for 45 mM DM (a) and 45 mM DDM (b) in the temperature range from 10°C to 70°C (see legends in panels a and b). Error bars are shown only for every 10$^{\text{th}}$ point for clarity. c), d) Guinier analysis of the same data as show in panels a and b, using the same color code. The black lines indicate typical fitting ranges, which where chosen such that $R_g \cdot q_{\max} \leq 1.3$ where $q_{\max}$ is the largest q-value included in the fit. Error bars are smaller than symbols. The insets show the reduced-χ^2 values (computed as squared difference between data and fit, divided by the variance, and normalized to the number of points included in the fit) for the Guinier fits presented in the main panel. Reduced-χ^2 values ≤ 1 indicate an excellent fit. The good linearity of the data on the Guinier range is apparent, suggesting monodisperse samples and the absence of bias in the data by micelle-micelle correlations.

cells and a cell holder that was temperature controlled in the range from 10°C to 70° with a circulating water bath.[8] Data were collected at an X-ray energy of 12 keV (corresponding to a wavelength of $\lambda = 1$ Å) using a sample-to-detector distance of 1.8 m, resulting in a useable q-range of 0.02 Å^{-1} to 0.275 Å^{-1}. Scattering angles were calibrated using a silver behenate standard sample. DM and DDM were purchased from Anatrace and measured in 20 mM phosphate buffer, pH 6.2, with 150 mM NaCl added at a detergent concentration

Table S1: Data-collection parameters. Data read out, normalization and circular averaging were performed using custom routines at beam line 12ID, APS, IL, USA. Buffer subtraction and Guinier analysis were performed using custom routines in Matlab (Mathworks) that are available from the authors upon request. Software for SAXS predictions and model refinement are described in detail in the SI text below.

Beam line	Beam line 12ID Advanced Photon Source, IL, USA
Wavelength (Å)	1.0
Useable q-range (Å^{-1})	0.02 - 0.275
Exposure time (s)	3.0
Monomer concentration (mM)	45
Temperature range (K)	283 - 343

of 45 mM. It has been previously shown that under these conditions, both DM and DDM micelles are monodisperse and interparticle interference (finite concentration) effects are negligible.[7,9] For each condition, three exposures of 1.0 s each were taken, image corrected, and circularly averaged. The three resulting profiles for each condition were compared to confirm the absence of radiation damage and averaged to improve signal. Matching buffer profiles were collected with identical procedures and subtracted for background correction. We employed 8 mg/ml horse heart cytochrome c (Sigma), in 100 mM acetate buffer, pH 4.6, with 0.5 M guanidinium hydrochloride as a scattering standard. All samples were centrifuged at $11000 \cdot g$ for 10 min prior to data collection.

Model-free determination of micelle aggregation number from forward scattering

The aggregation number N of micelles can be determined from the forward scattering intensity via the relationship:[7]

$$N = \frac{I(0)_{\mathrm{exp}}}{I(0)_{\mathrm{mon}}} = \frac{I(0)_{\mathrm{exp}}}{K\left[c(T) - cmc(T)\right]\left[\rho_{\mathrm{det}}(T) - \rho_{\mathrm{sol}}(T)\right]^2 V_{\mathrm{mon}}^2(T)} \tag{S2}$$

Table S2: Detergent parameters used in the calculation of the aggregation number from the forward scattering intensity.

Symbol	Quantity	Value at 25°C	
		DM	DDM
c	Detergent concentration	45 mM	45 mM
cmc	Critical micelle concentration	1.8 mM	0.17 mM
ρ_{det}	Electron density of the detergent	0.407 e Å^{-3}	0.398 e Å^{-3}
ρ_{sol}	Electron density of the solvent	0.340 e Å^{-3}	0.340 e Å^{-3}
V_{mon}	Volume of a detergent monomer	644.0 Å^3	697.8 Å^3

$I(0)_{\mathrm{exp}}$ is the experimentally determined forward scattering intensity of a detergent solution above the critical micelle concentration cmc that is obtained from Guinier analysis of the SAXS data (Figure S1). The denominator is the expected forward scattering signal from a detergent monomer (which is too weak to be measured directly). The proportionality constant K is setup specific and was determined from measurements of the scattering standard cytochrome c. The remaining parameters in the denominator are properties of the detergent and the solvent: ρ_{det} and ρ_{sol} are the electron densities of the solvent and detergent, respectively, c is the detergent concentration, cmc the critical micelle concentration, and V_{mon} the volume of a detergent monomer. An overview of the parameters and their known room temperature values from the literature are given in Table S2. The SAXS data at 25°C obtained in this study were analyzed using the parameter values from Table S2. The solvent and detergent parameters are, however, temperature dependent, as discussed in the next paragraph.

Size monodispersity. In line with previous findings,[10] our SAXS data suggests that DDM and DM micelles are reasonably monodisperse in size (or in aggregation number N_{agg}), based on two lines of arguments:

Firstly, if the micelle ensembles would exhibit a wide size distribution, the distribution would likely shift to larger sizes when increasing the detergent concentration, thereby leading to non-trivial changes in the scattering profiles. However, for concentrations low enough such that interparticle interference is negligible, the scattering profiles are superimposable after

rescaling by concentration,[7] suggesting that the micelles are reasonably monodisperse.

Secondly, polydisperse systems typically exhibit deviations from linearity in the Guinier region. However, we observe good linearity of the experimental data in the Guinier plots (Fig. S1c,d), providing additional indication that the micelles are reasonably monodisperse.

Moreover, we found that modelling the micelles as heterogeneous ensembles over various N_{agg} does not improve the agreement between experimental SAXS data and the SAXS data computed from free, unbiased simulations (Fig. S9). This suggests that (i) residuals between calculated and experimental SAXS curves are not caused by polydipersity in N_{agg} but instead reflect differences in micelle shape, and (ii) that explicit modelling of polydispersity in N_{agg} would mainly increase the risk of overfitting. Hence, we did not consider polydispersity in N_{agg} during SAXS-driven simulations.

Temperature dependence of the concentration. The concentration depends on temperature since the volume of the solvent changes with temperature. Stock solutions were prepared at room temperature ($\sim$25°C) and aliquots from the same stock solution were measured at different temperatures. We approximate the volume expansion of the solvent by the tabulated values for the density of water (Figure S2). Since the density of water only changes by about 2 % in the temperature range investigated, the temperature dependent change in concentration is a small correction.

Temperature dependence of the electron density of the solvent. The electron density of our aqueous buffer with 150 mM NaCl (Table S2) was used previously[7] and taken from Ref. 12. Its temperature dependence was again approximated using the temperature dependence of the density of water (Fig. S2). Even though the density of the solvent changes only by about 2 % over the temperature range studied, the fact that the two densities in the $(\rho_{\mathrm{det}} - \rho_{\mathrm{sol}})^2$ term in the denominator of Eq. S2 are similar means that even small changes are significant.

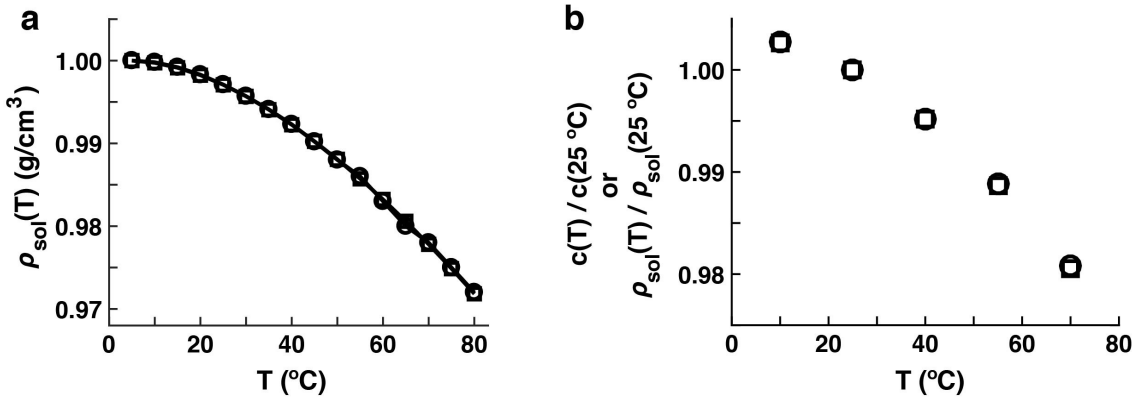


Figure S2: Temperature dependence of solvent density. a) Density of water as a function of temperature. Circles are data from http://www.engineeringtoolbox.com/water-thermal-properties-d_162.html; squares are data from Ref. 11. b) Relative temperature dependence of the density of the solvent and of the concentration, normalized to the values at 25°C. The values were computed from the water densities in panel a).

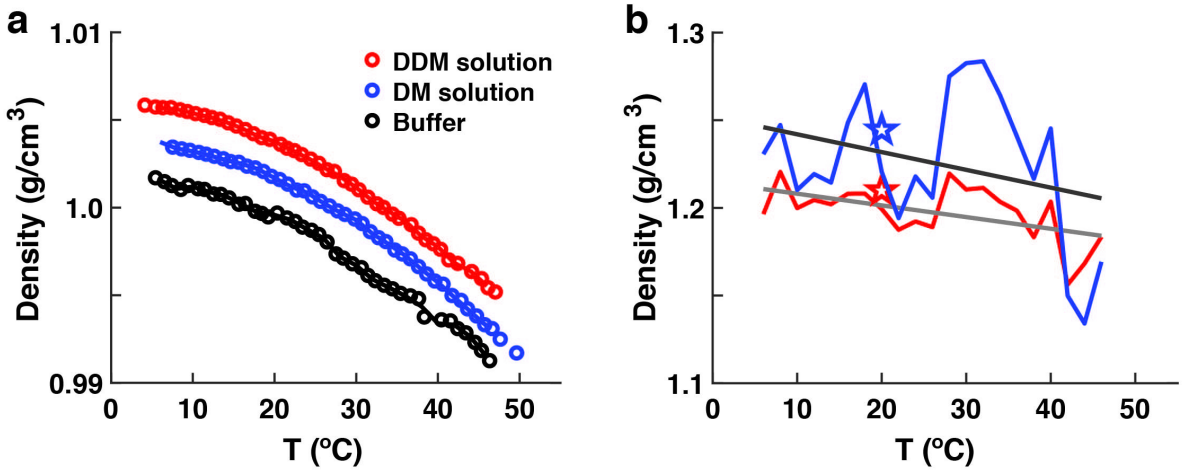


Figure S3: Temperature dependence of DM and DDM density. a) Density of buffer, DM, and DDM solutions as a function of temperature. Circles are from measurements with a Gay-Lussac pycnometer (see Methods). Lines are spline interpolated values. b) Density of DM and DDM as a function of temperature computed from the interpolated data in panel a) (colored lines; same color code as in panel a). The black and grey lines are linear fits to the DM and DDM data, respectively. The density values at room temperature obtained from the literature[7] are shown as stars for comparison.

Temperature dependence of the electron densities and monomeric volumes of the detergents. The monomeric volumes V_{mon} and corresponding electron densities ρ_{det} at room temperature were previously computed[7] from the specific densities,[13] using the Tanford formula for alkyl chain volumes to adjust for different chain lengths.[3] formula[3] for

the alkyl chain volume. To determine the temperature dependence of the density of DM and DDM, we used a Gay-Lussac pycnometer (Brand, Cat. No. 43305). The masses of the pycnometer filled with DM and DDM solutions in buffer and with buffer only were measured as a function of temperature in the range 10°C - 50°C. The buffer was 20 mM phosphate buffer, pH 6.2, with 150 mM NaCl added, i.e. the same buffer that was used for SAXS measurements. Temperature control was achieved through a water bath-thermostat (Biosan, WB-4MS). We limited the measurements to temperatures ≤ 50°C, as for higher temperatures the accuracy decreases according to vendor specifications and bubble formation in the pycnometer made accurate sample handling challenging. From the measured masses $m(T)$, the temperature-dependent densities of the buffer and the DM and DDM solutions were computed by taking into account the (temperature-independent) volume V and mass m_{pyc} of the pycnometer as $\rho(T) = (m(T) - m_{\mathrm{pyc}})V^{-1}$ (Figure S3a). From the temperature-dependent densities of the buffer and DM solutions, ρ_{buf} and $\rho_{\mathrm{DM,sol}}$, the temperature-dependent densities of DM ρ_{DM} was computed using the relationship:

$$\rho_{\mathrm{DM}}(T) = \frac{\rho_{\mathrm{buf}}(T)}{\frac{\rho_{\mathrm{buf}}(T)}{\rho_{\mathrm{DM,sol}}(T)}\left(1 + \frac{1}{\chi}\frac{FW_{H_2O}}{FW_{\mathrm{DM}}}\right) - \frac{1}{\chi}\frac{FW_{H_2O}}{FW_{\mathrm{DM}}}} \tag{S3}$$

where FW_{DM} is the formula weight (i.e. the molecular mass) of DM, FW_{H_2O} the formula weight of water, and χ the molality of the DM solutions, in number of DM molecules per solvent molecule. An analogous expression was used for DDM. We used 18.02, 482.56, and 510.62 Da for the FW of water, DM and DDM, respectively. The resulting densities of DM and DDM are shown in Figure S3. The data are relatively noisy, due to the difficulty of measuring very small changes in density with the pycnometer. Nonetheless, the values are consistent, within experimental error, with the density values at room temperature from the literature[7,13] (Figure S3, stars). To obtain a robust estimate of the temperature dependence, and to extrapolate the data beyond the measured temperature range, we applied a linear fit to the measured data (Figure S3, black and grey lines). The fitted linear temperature

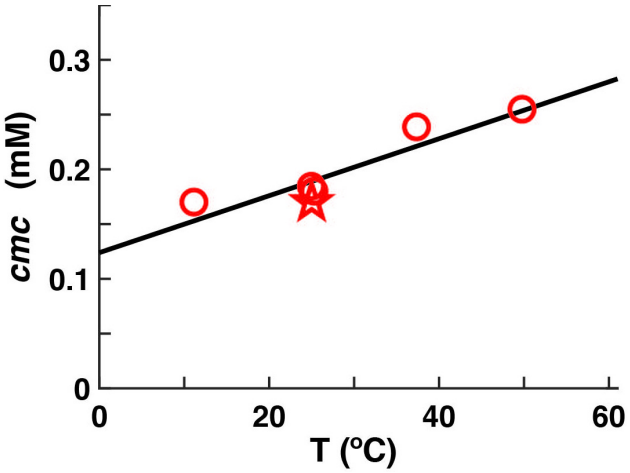


Figure S4: Temperature dependence of the critical micelle concentration. Critical micelle concentration of DDM as a function of temperature. The circles show the data from Ref. 14. The star is the room temperature value used previously.[7] The black solid line shows the linear fit to the data that was used to evaluate the temperature-dependence of the cmc for both DM and DDM.

dependence was used to evaluate Eq. S2.

Temperature dependence of the critical micelle concentrations. The critical micelle concentrations in Table S2 are the values used previously[7] and were taken from the Anatrace catalogue (https://www.anatrace.com/). The temperature dependence of the *cmc* for DDM was measured spectrofluorometrically by Aoudia *et al.*[14] The temperature-dependent data by Aoudia *et al.* are well described by a simple linear relationship (Figure S4) and we used this linear dependence on temperature to compute the aggregation number by evaluating Equation S2 for both DM and DDM.

MD simulations and SAXS calculations

MD setup and simulation parameters. Structures of single detergent molecules of the n-dodecyl-β-D-maltoside (DDM) and n-decyl-β-D-maltoside (DM) were taken from the CHARMM-GUI web site.[15] To build the micelle, initial coordinates were generated by placing the detergent molecules in a spherical and uniformly distributed arrangement. The structures were placed into a simulation box of a dodecahedron, keeping a distance of at least

3 nm to the box boundary. The simulation boxes were filled by CHARMM-modified TIP3P water.[16,17] Water molecules were removed from the hydrophobic cores of the micelles. The energy of each system was minimized with the steepest-decent algorithm. Micelles of different aggregation numbers ($N_{\mathrm{agg}}^{\mathrm{sim}}$) were setup and initially equilibrated at 30°C. For DDM, we set up micelles of 70 to 210 detergent molecules in steps of 5. For DM we we set up micelles of 60 to 110 detergent molecules in steps of 5. The final structures from these equilibration simulations were used as a starting structures for the free production simulations. These simulations were run for another 100 ns, if not stated otherwise, at the temperatures of 10°C, 25°C, 40°C, 55°C, and 70°C. SAXS curves were calculated from 1000 snapshots of the last 50 ns of the simulations, using the explicit-solvent SAXS calculations described previously (see below for details).[18]

Unbiased, free simulations were carried out using the GROMACS simulation software, version 5.0.4.[19] SAXS curve predictions and SAXS-driven MD simulations (with experiment-derived energetic restraints) were conducted with an in-house modification of GROMACS 4.6. Detergent interactions were modeled with the CHARMM36 lipid forcefield,[20] version of March 2014, translated into GROMACS.[21] The temperature was controlled at the desired value using velocity rescaling[22] during free simulations ($\tau = 1\,\mathrm{ps}$), and using a stochastic dynamics integrator during SAXS-driven simulations ($\tau = 0.2\,\mathrm{ps}$). The pressure was kept at 1 bar using the Berendsen barostat[23] ($\tau = 5\,\mathrm{ps}$). Long-range electrostatic interactions were calculated using the particle-mesh Ewald method.[24,25] Dispersive interactions and short-range repulsion were described together by a Lennard-Jones potential with a cutoff at 1.2 nm.

SAXS curve predictions and SAXS-driven MD simulations. SAXS curves were computed using the explicit-solvent calculations described previously.[18] Explicit water molecules that contributed to the SAXS calculations were defined by a spatial envelope that enclosed the micelle and the hydration layer (Fig. S5, blue surface). The envelope was constructed

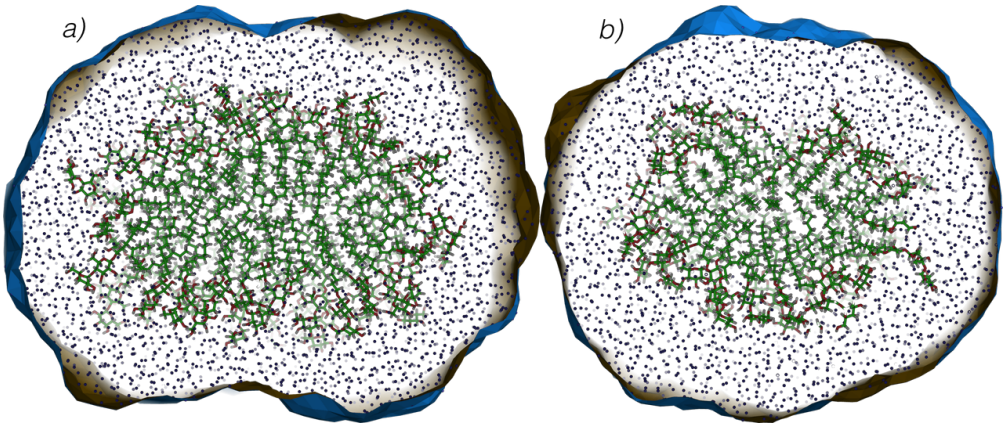


Figure S5: Spatial envelope around micelles of DDM (a) and DM (b), separating the micelles and the solvation layer from the bulk water. The envelope was constructed at the ∼1.5 nm distance from the micelles surface. Explicit water inside the envelope contributed to the SAXS calculations, thereby accurately accounting for scattering contributions from the hydration layer.

such that the vertices of the envelope had a distance of at least 0.6 nm from all micelle atoms during the simulation. Due to substantial fluctuation of the micelle, this procedure led to a distance between micelle and envelope of ∼1.5 nm in most simulation frames. To carry out the orientational average (or spherical quadrature), scattering amplitudes were computed for 1000 **q**-vectors per absolute value of the momentum transfer q, which were distributed by the spiral method. Because the density of the applied TIP3P water slightly differs from the experimental density, we corrected the solvent density to $334\,\mathrm{e\,nm^{-3}}$ as described previously.[18]

Extracting the aggregation number from SAXS data using MD simulations

Figure S6b shows SAXS curves of DDM micelle computed with the analytic model by Lipfert *et al.*,[7] which models a micelle as two-component density (for head groups and tails) in the shape of a spheroid, with one semi-axis a and two semi-axes b. SAXS curves were computed for various prolate ($a > b$) and oblate ($a < b$) micelles with different semi-axes a and b (see Fig. S6b, legend), but with constant overall volume (and hence modelling a constant

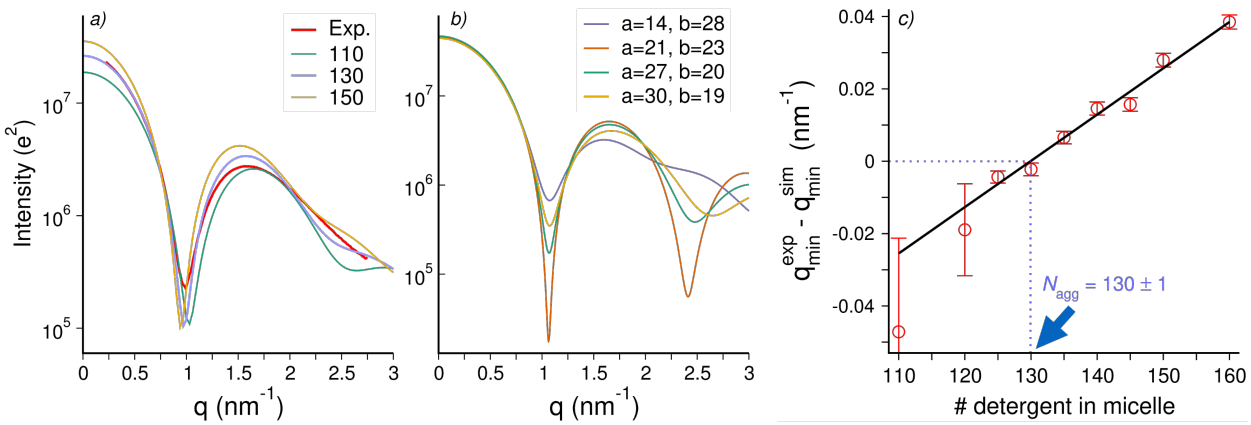


Figure S6: (a) SAXS curves computed from the simulations of micelles composed of 110, 130 and 150 DDM molecules (for color coding, see legend). Red curve: experimental SAXS curve for DDM micelles at 25°C. Evidently, the location of the minimum at $q \approx 1\,\mathrm{nm}^{-1}$ is shifted to lower q with increasing aggregation number. (b) SAXS curves of DDM micelle computed with the analytic model by Lipfert *et al.*[7] (c) Linear fit of the difference between the q-positions of the experimental and the calculated SAXS curve minima, here shown for DDM at 25°C. The blue arrow indicates the simulated detergent number that leads to the best match of the minimum between simulation and experiment.

aggregation number) as apparent from the constant forward intensity at $q = 0$. For the analytic model, the head group thicknesses were taken as $t_\mathrm{a} = t_\mathrm{a} = t_\mathrm{c} = 6.06\,\text{Å}$, while the densities of the core, the head groups, and the solvent were taken as $0.227\,\mathrm{e\,Å}^{-3}$, $0.520\,\mathrm{e\,Å}^{-3}$, and $0.334\,\mathrm{e\,Å}^{-3}$, respectively.[7] Evidently, although a and b strongly influenced the SAXS curves, the position of the first minimum at $q \approx 1\,\mathrm{nm}^{-1}$ is well conserved, suggesting that the position of the minimum encodes mainly the volume and much less the shape of the micelle. Indeed, SAXS curves computed from free micelle simulations with increasing number of detergent molecules reveal a systematic left-shift of the minimum with increasing detergent number (Fig. S6a). Consequently, we estimated the aggregation number by comparing the position of the SAXS curve minimum between (i) the experimental SAXS curves, $q_\mathrm{min}^\mathrm{exp}$, and (ii) the SAXS curves computed from unbiased MD simulations, $q_\mathrm{min}^\mathrm{sim}$, with different number of detergent molecules in the simulated micelle.

The q-positions of the SAXS curve minima q_min were extracted by fitting a parabola to the minimia, $I_\mathrm{fit}(q) = a(q - q_\mathrm{min})^2 + c$, within a small q-range around the minima, using the

Table S3: Aggregation numbers (N_{agg}) obtained by comparing the position of the SAXS curve minimum between experimental curves and MD-computed curves, statistical error (δN_{agg}), and aggregation number used for production SAXS-driven simulations (Used) at different temperatures.

Temperature	**DDM**			**DM**		
(°C)	$N_{\mathbf{agg}}$	$\delta N_{\mathbf{agg}}$	**Used**	$N_{\mathbf{agg}}$	$\delta N_{\mathbf{agg}}$	**Used**
10°	146.8	2.5	145	91.7	2.4	90
25°	129.9	1	130	84.8	0.2	85
40°	115.8	2.4	115	80	0.7	80
55°	105.7	1.2	105	76.6	2	75
70°	97.8	2.6	100	70.4	1.5	70

Levenberg-Marquardt algorithm. The fitted q-range was $0.1\,\mathrm{nm}^{-1}$ for the experimental SAXS curves, and $0.3\,\mathrm{nm}^{-1}$ and $0.4\,\mathrm{nm}^{-1}$ for the calculated curves for DDM and DM, respectively. All fitted parabolas closely matched the data in the fitted range. This procedure was repeated for SAXS curves computed from simulations with various detergent numbers $N_{\mathrm{agg}}^{\mathrm{sim}}$, yielding a series of computed minima positions, $q_{\mathrm{min}}^{\mathrm{sim}}(N_{\mathrm{agg}}^{\mathrm{sim}})$. Finally, we fitted a straight line to $\Delta q(N_{\mathrm{agg}}^{\mathrm{sim}}) = q_{\mathrm{min}}^{\mathrm{exp}} - q_{\mathrm{min}}^{\mathrm{sim}}(N_{\mathrm{agg}}^{\mathrm{sim}})$ (Fig. S6c, black line), and we obtained the experimental aggregation number N_{agg} by extrapolating to $\Delta q = 0$ (Fig. S6c, blue arrow). Statistical errors of $\Delta q(N_{\mathrm{agg}}^{\mathrm{sim}})$ were obtained from the Levenberg-Marquardt algorithm and using error propagation. The error of N_{agg} was taken from the Levenberg-Marquardt algorithm. We stress that these errors represent purely statistical errors due to the extraction of the SAXS curve minimum. Putative systematic errors are not included, which could, for instance, appear in case of imperfections of the detergent densities in the simulations; however, because the estimated N_{agg} well agrees with N_{agg} obtained with the model-free approach (see above, and Fig. 2), such systematic errors are probably small. The estimated N_{agg} are summarized in Table S3.

SAXS-driven MD simulations. Final structures obtained by free simulations were used as starting structures for the SAXS-driven simulations. These simulations were run for 300 ns at 25°C, and between 40 ns and 60 ns at all other temperatures. SAXS-derived forces applied

to the detergent molecules were calculated from the SAXS-derived potential E_{SAXS}:[26]

$$E_{\mathrm{SAXS}}(\mathbf{R},t) = \alpha(t) k_c \frac{k_B T}{n_q} \sum_{i=1}^{n_q} \frac{[\langle I_c(\mathbf{R},q_i)\rangle_{t;\tau} - f I_e(q_i)]^2}{\sigma^2(q_i)} \tag{S4}$$

where $I_{\mathrm{e}}(q_i)$ denotes the experimental SAXS intensity. f is a fitting constant for the absolute intensity scale, which was adjusted at every step such that E_{SAXS} is minimized. In contrast to previous work,[26] no fitting constant for a constant intensity offset was applied because adjusting an offset was not required to obtain quantitative agreement between experiment and simulation. $\langle I_c(\mathbf{R},q_i)\rangle_{t;\tau}$ is the SAXS intensity computed on-the-fly from the simulation coordinates $\mathbf{R}$. The symbol $\langle\cdot\rangle_{t;\tau}$ denotes the running average at time t, using weights that decay exponentially into the past with a memory time τ. In this work, we used $\tau = 300\,\mathrm{ps}$. As such, the time-averaged $\langle I_c(\mathbf{R},q_i)\rangle_{t;\tau}$ represents an average over fluctuations that occur on a time scale of a few hundred picoseconds, implying that a fluctuation-averaged SAXS curve is compared with the experimental curve in Eq. S4. The overall uncertainty $\sigma(q_i)$ accounts for experimental and statistical calculated errors, as well as for a systematic error that originates from an uncertainty of the buffer density.[26] To estimate the latter, we assumed a relative uncertainty of the solvent density of 0.1%. The experimental errors were modelled as 1% of the experimental intensity. The symbol k_c is a force constant set to 1 in this study, n_{q} is the number of intensity q-points, $k_B T$ the thermal energy, and $\alpha(t)$ is a time-dependent function that allows a gradual introduction of the SAXS-derived potential at the beginning of the simulation ($0 < \alpha(t) \leq 1$). The first 8 ns of SAXS-driven simulations were not used for analysis in order to account for equilibration.

Figure S7 presents average E_{SAXS} values obtained from SAXS-driven simulations of DDM and DM micelles with various aggregation numbers at 25°C. Evidently, E_{SAXS} takes small values in DDM and DM simulations if N_{agg} is close to 130 and 85 for DDM and DM, respectively, corresponding to the N_{agg} estimates that we obtained by matching the position of the SAXS curve minimum at $q \approx 1\,\mathrm{nm}^{-1}$ between experiment and unbiased simulation

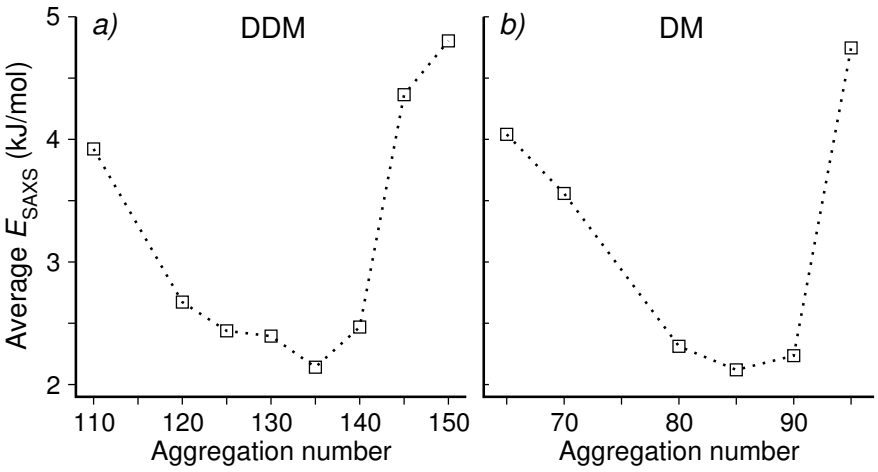


Figure S7: Average SAXS-derived potential E_{SAXS} (Eq. S4) during SAXS-driven simulations of DDM (a) and DM (b). E_{SAXS} micelles at different aggregation number at 25°C, calculated from the last 10 ns of the simulations.

(see Fig. S6). In contrast, E_{SAXS} is increased in SAXS-driven simulation with N_{agg} values that strongly deviate from the optimal values, reflecting that micelles with an incorrect N_{agg} cannot be refined to shapes that accurately agree with the experimental data.

To exclude that the SAXS-driven simulations are biased by the initial conformation, we started SAXS-driven simulations for DDM at 25°C from multiple initial frames picked from a free simulation, such that the micelle exhibited different shapes in the initial frames (some more prolate-, some more oblate-like). These SAXS-driven simulations consistently led to rather prolate-like shapes, suggesting that (i) the SAXS-driven simulations were not biased by the initial frames, and (ii) that the simulations do not suffer from sampling problems owing to multiple energetic minima.

The modified Gromacs source code used for SAXS predictions and SAXS-driven MD simulations is available on the authors' website (http://cmb.bio.uni-goettingen.de/).

Electron density calculations. Electron densities along the principal axes were computed as an average over SAXS-driven simulations, as follows (Fig. 4a-f). For each simulation frame, the mass-weighted principal axes were computed from all micelle atoms. Subsequently, a cylinder of 0.5 nm was aligned along each axis. The electron densities were computed from

the atoms of the respective atom type (representing tails, head groups, water, or all atoms) within the cylinder along the respective axis (minor, middle, or major). Here, the hydrocarbon chain was defined as "tail", and all other atoms (including the oxygen bound to the hydrocarbon chain) as "head group".

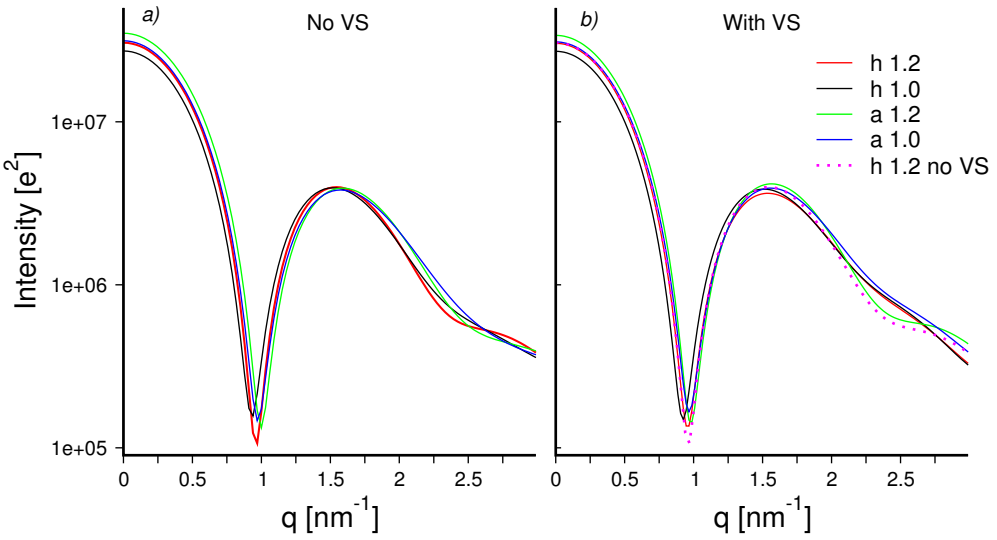
The length of the three full semi-axes, $a + t_a$, $b + t_b$, and $c + t_c$ (Fig. 4g/h), were defined as the distance from the micelle center of mass, where the density of detergent dropped below $120\,\mathrm{e\,nm^{-3}}$, corresponding to approximately half the density of micelle core. Likewise, the semi-axes of the hydrophobic core (Fig. S12a/b) were defined by the distance where the density of the tails dropped below $120\,\mathrm{e\,nm^{-3}}$. The error bars were computed by block averaging, using blocks of 4 ns.

The electron densities of lipid membranes of DMPC and POPC were computed from 20 ns and 40 ns of equilibrium simulations, respectively. The membrane simulations contained 128 lipids plus 40 water molecules per lipid. Interactions were described by the CHARMM36 force field and the CHARMM-modified TIP3P model.[20,27] The simulation of DMPC was taken from a recent study.[28] The POPC setup and MD parameters were identical to the DMPC simulation described previously.[28] The tail densities of the membranes were computed purely from the hydrocarbon lipid tails, starting with the first carbon atom below the ester groups, i.e., the ester and glycerol groups were not considered as part of the "tails".

Computational tests

Influence of MD parameters and force field details on calculated SAXS curves. Before running production simulations, we have carefully evaluated the influence of various MD parameters on free micelle simulations: (i) the influence of cutoff distances for non-bonded interactions; (ii) the effect of constraining all bonds instead of purely bonds of hydrogen atoms; (iii) the effect of modelling hydrogen atoms as virtual sites (v-sites), allowing one to increase the integration time step from 2 to 4 fs; (iv) influence of the water model.

All test simulations of a DDM micelle were conducted under the same conditions (aggre-



[h!]

Figure S8: SAXS curves computed from free simulations of a DDM micelle composed of 140 detergent molecules at 25°C, used to test the influence of various MD parameters on the calculated SAXS curves. (a) Simulations modelling hydrogen atoms as normal atoms, and (b) modelling hydrogen atoms at virtual sites (VS). Numbers 1.2 and 1.0 represents the cutoff distance for Lennard-Jones and short-range Coulomb interactions. Letters "a" and "h" indicate bond constraints of all atoms or purely involving hydrogen atoms, respectively.

gation number 140, temperature 25°C). SAXS curves calculated from these test simulations are shown in the Figure S8. We found that the cutoff distance as well as the constraints settings may influence the SAXS curves. In contrast, modelling hydrogen atoms as virtual sites had only a small effect on the SAXS curves. Likewise, using the standard TIP3P water model[27] instead of the CHARMM-modified TIP3P model (with Lennard-Jones interactions of hydrogen atoms) did not influence the SAXS curves. We decided to follow the settings that closely resemble the default settings for the CHARMM36 force field: cutoff at 1.2 nm, bond constraints applied purely to hydrogen atoms, and hydrogen atoms *not* modelled as v-sites.

Test of convergence. To exclude that the calculated SAXS curves were biased by sampling problems, we conducted two independent simulations of the DDM micelle using increased temperatures and simulated annealing, as follows: (i) 50 ns of simulation at either 370 K or 420 K; (ii) annealing down to 300 K within 20 ns; and (iii) 100 ns at 300 K. SAXS

curves were calculated from the last 50 ns of the two simulations and compared with the previously calculated curves for the same system simulated purely at 300 K. All three calculated SAXS curves were nearly identical, suggesting that our simulations and SAXS calculations were not biased by sampling problems.

In addition, to exclude that the initial detergent conformation influences the computed SAXS curve, the system of 140 DDM detergent molecules was set up following three different procedures: (i) placing the detergent molecules in a spherical and uniformly distributed arrangement; (ii) by building a preassembled micelle with the CHARMM-GUI server;[15] and (iii) via a simulation of micelle aggregation, starting from a random distribution of 140 detergent molecules in a water box. Here, the micelle formed within 200 ns. SAXS curves calculated from the three different setup procedures were nearly identical, suggesting that our approach of generating the initial micelle coordinates does not bias the results.

Additional supporting figures

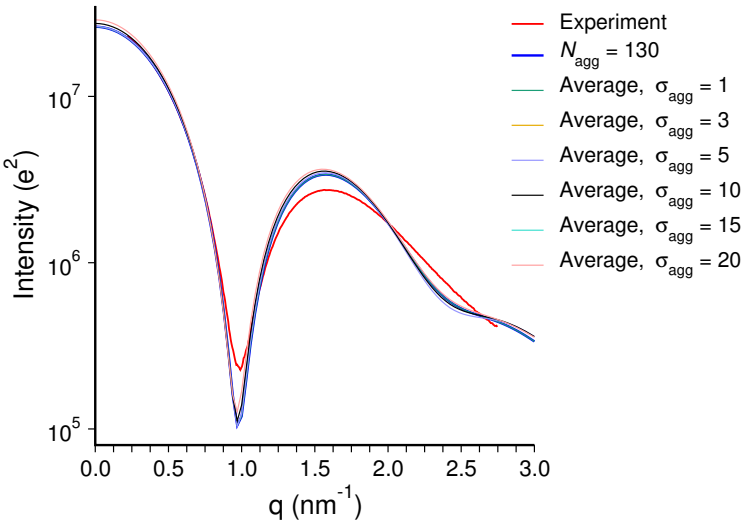


Figure S9: SAXS curves for DDM at 25°C, demonstrating that modelling a heterogeneous ensemble over various aggregation numbers (polydispersity in micelle size) does *not* improve the agreement between the experimental curve (red) and the curve calculated from free MD simulations. Blue: SAXS curve from free simulation with 130 DDM molecules. All other curves: Average SAXS curves $I_{\mathrm{av}}(q;\sigma_{\mathrm{agg}}) = W^{-1}\sum_n w_n(\sigma_{\mathrm{agg}})I(q;n)$ computed as a weighted average over curves $I(q;n)$ computed from free MD simulations with $n = 110, 120, 125, 130, 135, 140, 145, 150$ or 160 detergent molecules. Here, w_n is the weight taken from a Gaussian distribution with mean 130 and width σ_{agg}, and $W = \sum_n w_n(\sigma_{\mathrm{agg}})$ is the normalization constant.

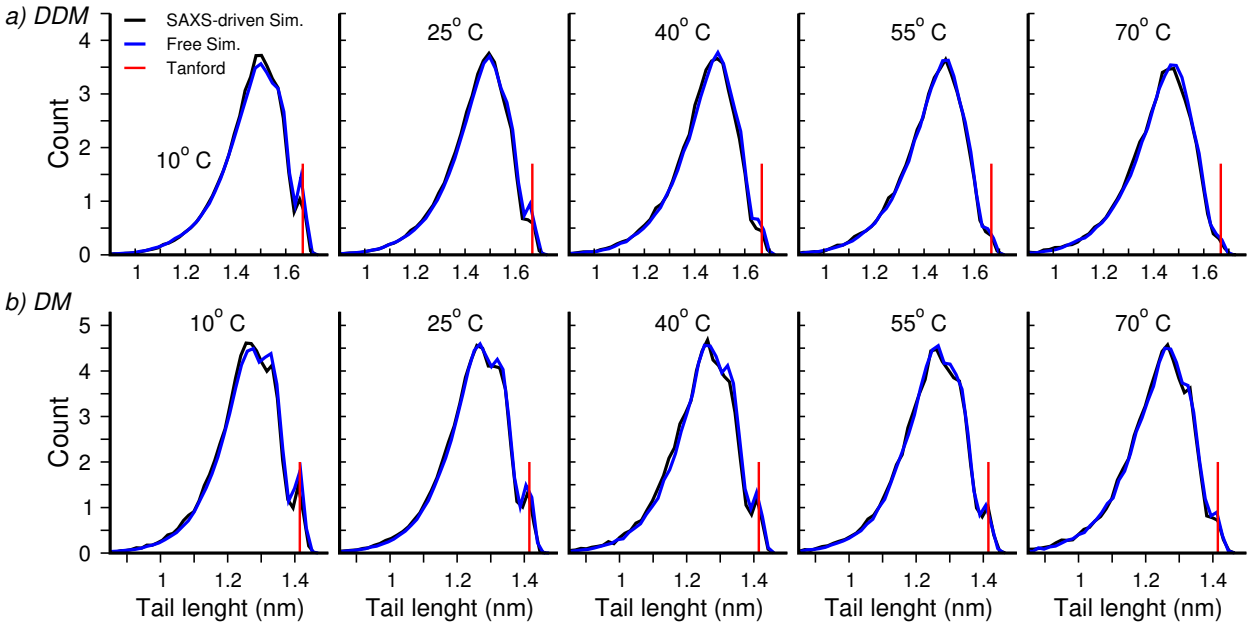


Figure S10: Distributions of the detergent tail length for DDM (a) and DM (b), computed as the distance between the the first carbon atom (C_1) of the tail (ii) the terminal carbon atom (C_t) of the tail, corrected by 0.21 nm due to (i) the Van-der-Waals radius of the terminal methyl group (0.15 nm) and (ii) half of the bond length between C_1 and the neighboring oxygen atom (0.06 nm), thus following Tanford's defintion.[3] Blue curves: free simulations; black curves: SAXS-driven simulations. Results are shown for simulations at temperatures between 10°C and 70°C, see labels. Red vertical marks indicate the maximum tail length estimated by Tanford's equation, 1.668 nm for DDM tail and 1.415 nm for DM tail. The curves demonstrate that the SAXS-derived restraints hardly influence tail length distribution and, hence, hardly influence the tail structure.

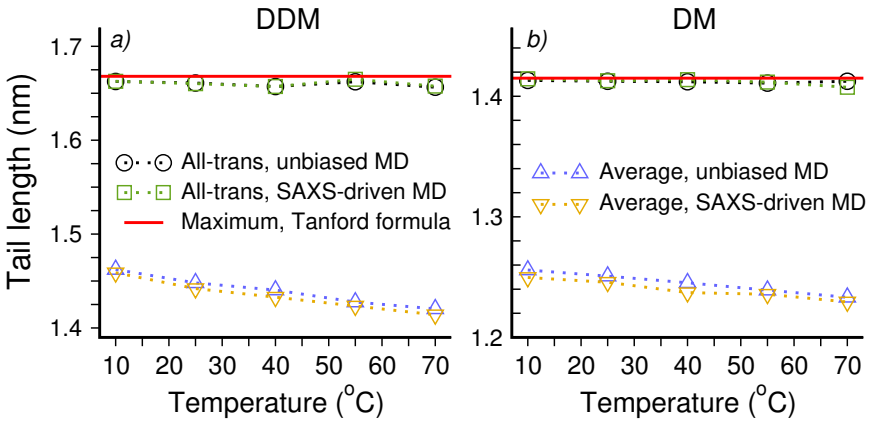


Figure S11: Average detergent tail length of DDM (a) and DM (b) at various temperatures, taken from the last 50 ns of free, unbiased simulations (blue) or last 10 ns of SAXS-driven simulations (orange). The tail length was computed as the distance between the the first carbon atom (C_1) of the tail (ii) the terminal carbon atom (C_t) of the tail, corrected by 0.21 nm due to (i) the Van-der-Waals radius of the terminal methyl group (0.15 nm) and (ii) half of the bond length between C_1 and the neighboring oxygen atom (0.06 nm), following Tanford's defintion.[3] The data demonstrate that the SAXS-derived restraints influence the average tail length only marginally. Further, the average tail length slightly decreases with increasing temperature, as expected since disordered tails are favoured by entropy. For comparison, black and green symbols show the tail length averaged purely over all-trans configurations of detergent molecules, revealing close agreement with the maximum extension estimated by Tanford's equation (red lines; DDM: 1.668 nm; DM: 1.415 nm).

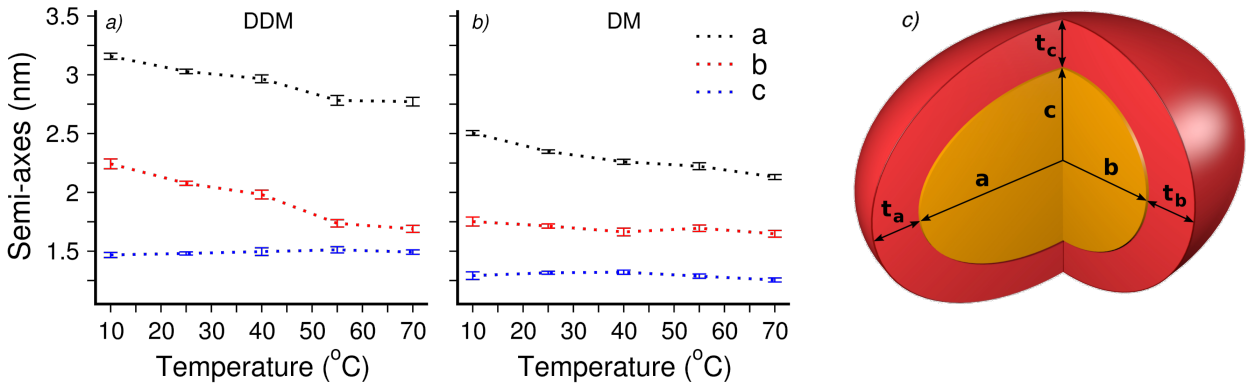


Figure S12: Length of the semi-axes a, b, and c of the hydrophobic cores along the three principal axes of refined micelles of DDM (a) and DM (b). a, b, and c are plotted versus temperature. Errors were computed by binning analysis. The major and middle semi-axes (b and c) shrink with increasing temperature, whereas the minor semi-axes c are approximately temperature-invariant in both DDM and DM. Error bars denote 1 SEM computed from block averaging. (c) Schematic model of a micelle, illustrating the hydrophobic core with semi-axes a, b, and c (orange), and the head groups with thicknesses t_a, t_b, and t_c (red). The lengths of the full semi-axes including the head groups, $a + t_a$, $b + t_b$, and $c + t_c$, are shown in main text Fig. 4.

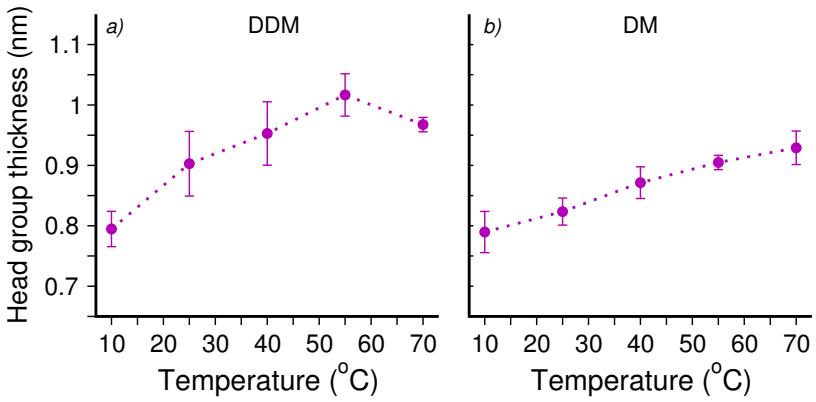


Figure S13: Thicknesses of head groups in a DDM and DM micelle plotted as a function of temperature. Because the head group thicknesses along the three principal axes were identical within statistical errors, we here averaged the thicknesses over the three principal axes. The thicknesses were computed as FWHM of the Gaussian-like head group electron density distributions (Fig. 4b-g, magenta lines). Statistical errors (1 SEM) are slightly increased owing to occasional long-living head group/head group contacts, leading to slower sampling of head group conformation as compared to tail conformations. Overall, the head group thickness slightly increases with temperature, rationalized by increased fluctuations and disorder at higher temperatures.

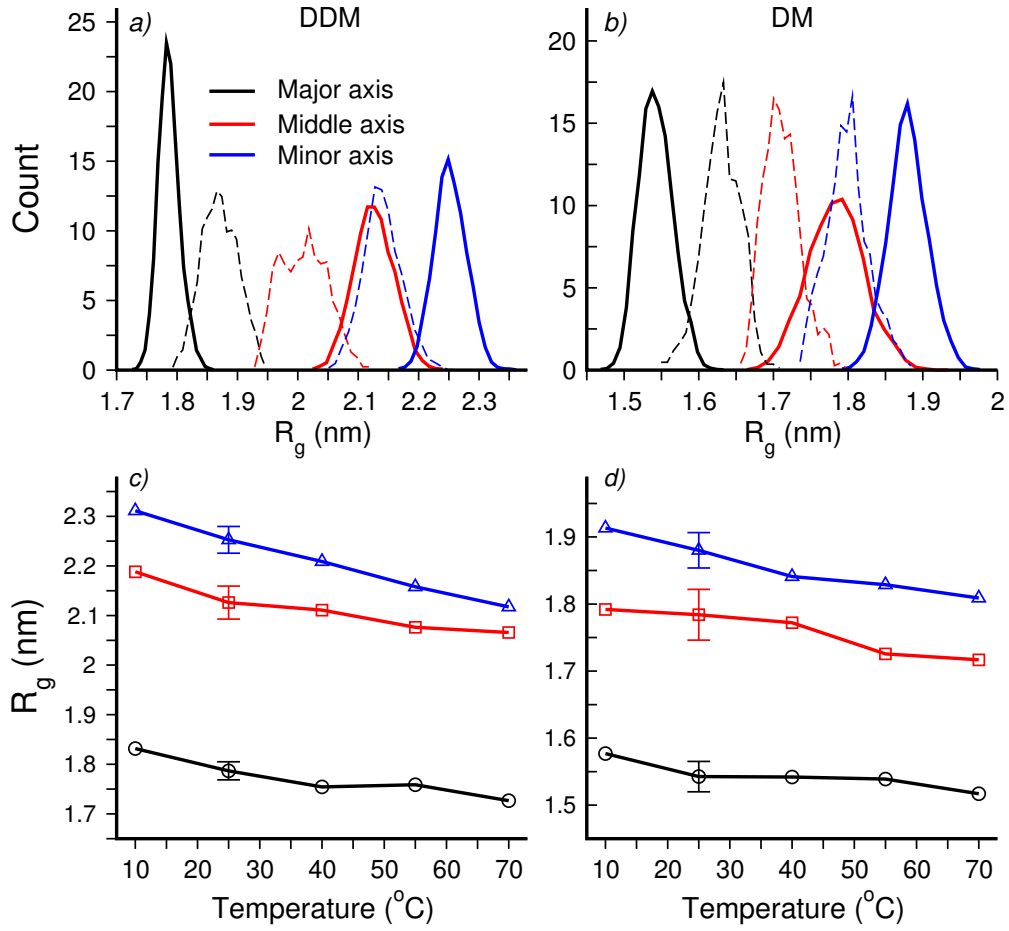


Figure S14: Distributions of radii of gyration R_g around three principal axes at the 25°C for DDM (a) and DM (b). Radii of gyration are related to the moments of inertia (MOI) $I^{(\mathrm{m})}$ via $R_{\mathrm{g}}^{(\mathrm{m})} = (I^{(\mathrm{m})}/M)^{1/2}$ where $m = 1, 2, 3$ indicates the major, middle or minor principal axis, and M is mass of the micelle. Solid lines represent the results from 300 ns of SAXS-driven simulations. For comparison, dashed lines show distributions from free simulations, demonstrating that micelles in free simulations, which agree with the SAXS data only approximately, were too spherical. (c,d) Temperature dependence of average radii of gyration around the three principal axes, $R_{\mathrm{g}}^{(\mathrm{m})}$, for refined DDM (c) and DM (d) micelles. The error bars at the 25°C indicate the standard deviations of the R_g distributions. Two large and one small $R_{\mathrm{g}}^{(\mathrm{m})}$ (see blue and red versus black line) indicate rather prolate-like than oblate-like micellar shapes for all temperatures.

References

[1] Maibaum, L.; Dinner, A. R.; Chandler, D. *J. Phys. Chem. B* **2004**, *108*, 6778–6781.

[2] Zeppieri, S.; Rodríguez, J.; López de Ramos, A. *J. Chem. Eng. Data* **2001**, *46*, 1086–1088.

[3] Tanford, C. *The Hydrophobic Effect: Formation of Micelles and Biological Membranes 2d Ed*; J. Wiley., 1980.

[4] Seifert, S.; Winans, R.; Tiede, D.; Thiyagarajan, P. *J. Appl. Crystallogr.* **2000**, *33*, 782–784.

[5] Oliver, R. C.; Lipfert, J.; Fox, D. A.; Lo, R. H.; Kim, J. J.; Doniach, S.; Columbus, L. *Langmuir* **2014**, *30*, 13353–13361.

[6] Oliver, R. C.; Lipfert, J.; Fox, D. A.; Lo, R. H.; Doniach, S.; Columbus, L. *PLoS One* **2013**, *8*, e62488.

[7] Lipfert, J.; Columbus, L.; Chu, V. B.; Lesley, S. A.; Doniach, S. *J. Phys. Chem. B* **2007**, *111*, 12427–12438.

[8] Lipfert, J.; Millett, I. S.; Seifert, S.; Doniach, S. *Rev. Sci. Instrum.* **2006**, *77*, 046108.

[9] Bruetzel, L. K.; Fischer, S.; Salditt, A.; Sedlak, S. M.; Nickel, B.; Lipfert, J. *Rev. Sci. Instrum.* **2016**, *87*, 025103.

[10] Dupuy, C.; Auvray, X.; Petipas, C.; Rico-Lattes, I.; Lattes, A. *Langmuir* **1997**, *13*, 3965–3967.

[11] Kell, G. S. *Journal of Chemical and Engineering Data* **1975**, *20*, 97–105.

[12] Schiel, J. E.; Hage, D. S. *Talanta* **2005**, *65*, 495–500.

[13] Le Maire, M.; Champeil, P.; Møller, J. V. *BBA-Biomembranes* **2000**, *1508*, 86–111.

[14] Aoudia, M.; Zana, R. *Colloid Interface Sci.* **1998**, *206*, 158–167.

[15] Cheng, X.; Jo, S.; Lee, H. S.; Klauda, J. B.; Im, W. CHARMM-GUI micelle builder for pure/mixed micelle and protein/micelle complex systems. 2013.

[16] Brooks, B. R.; Brooks, C. L.; MacKerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S. *J. Comput. Chem.* **2009**, *30*, 1545–1614.

[17] Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E.; Mittal, J.; Feig, M.; MacKerell Jr, A. D. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.

[18] Chen, P.; Hub, J. S. *Biophys. J.* **2014**, *107*, 435–447.

[19] Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. *SoftwareX* **2015**, *1*, 19–25.

[20] Pastor, R.; MacKerell Jr, A. *J. Phys. Chem. Lett.* **2011**, *2*, 1526–1532.

[21] Bjelkmar, P.; Larsson, P.; Cuendet, M. A.; Hess, B.; Lindahl, E. *J. Chem. Theory Comput.* **2010**, *6*, 459–466.

[22] Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 014101.

[23] Berendsen, H. J. C.; Postma, J. P. M.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

[24] Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

[25] Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys* **1995**, *103*, 8577–8593.

[26] Chen, P.; Hub, J. S. *Biophys. J.* **2015**, *108*, 2573–2584.

[27] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.

[28] Hub, J. S.; Awasthi, N. *J. Chem. Theory Comput.* **2017**, *13*, 2352–2366.