Electronic Supplementary Information for Aggregation response of triglyceride hydrolysis products in cyclohexane and triolein

Sampsa Vierros[1], Monika Österberg[2], and Maria Sammalkorpi[1]

[1]Department of Chemistry and Materials Science, Aalto University, P.O. Box 16100, 00076 Aalto, Finland
[2]Department of Bioproducts and Biotechnology, Aalto University, P.O. Box 11000, 00076 Aalto, Finland
Email: maria.sammalkorpi@aalto.fi

The supporting information consists of detailed information of 1) the simulated system compositions (Table S1), 2) determination of the aggregate cluster cutoff distance, 3) equilibration and convergence assessment of the simulations, 4) the influence of the aggregate cluster cutoff distance on the aggregation energy, 5) photographs of the dye solubilization samples in the experiments, and 6) the aggregate size distributions resulting from the simulations.

1 Simulated system compositions

Table S1 specifies in detail the system compositions and simulation setup details for the simulation runs. In the table, the surfactant is either monopalmitin (MPG), dipalmitin (DPG), or palmitic acid (PA) while solvent is triolein (TOG) or cyclohexane (CHX).

2 Determination of cluster cutoff distance

As reverse micellar aggregation is driven by polar interactions occurring between the headgroups of different surfactant molecules, the surfactants were classified to aggregates based on their headgroups residing close to each other. More specifically, surfactant molecules were considered to be part of the same cluster if their headgroups were closer than the cluster cutoff distance r_c from each other. The distance between the headgroups of two surfactants was calculated as the shortest distance between any two headgroup atoms in the two surfactants. The headgroup atoms included the glyceryl moiety and the ester groups of monoglyceride and diglyceride and for the free fatty acid, the acid moiety.

The aggregate size distribution, particularly in highly concentrated systems, can be expected to depend on the cluster cutoff distance r_c. To obtain a reasonable value range for the cutoff distance, the size distribution corresponding to a handful of simulation configurations was classified manually based on visual analysis and the aggregates compared to those obtained using different cutoff distances. Cutoff distance ranges 0.30-0.35 nm and 0.29-0.36 nm were found to reproduce the reference distributions of monoglyceride and diglyceride, respectively. As the ranges are rather similar, the clustering of free fatty acid was not examined explicitly. A cutoff value of 0.33 nm, corresponding to the middle of the range, was used in the data analysis as the cluster cutoff distance for all examined surfactant species.

To estimate the uncertainty originating from the choice of cutoff in the simulations, the free surfactant concentration was calculated with different cutoffs, see Figure S1a. The figure shows that the maximum uncertainty resulting from cutoff is around 8-16 mM. Additionally, examination of the automatically generated size distributions with varying cutoff distance revealed that the free surfactant concentration was sensitive to the cutoff particularly in the concentrated systems, see Figure S1b.

Table S1: Detailed system compositions and simulation setup details in the simulation runs. The table specifies for each surfactant-solvent pair the employed simulation models, the total surfactant concentration, number of surfactant molecules N_s, number of solvent molecules N_{solv}, cubic simulation box edge length L, and total simulation time t.

Surfactant	Solvent	Simulation model	Concentration mM (wt %)	N_s	N_{solv}	L nm	t ns
MPG	TOG	C27/C36	130(5.0)	15	106	5.76	200
MPG	TOG	C27/C36	249(9.6)	30	106	5.85	200
MPG	TOG	C27/C36	321(12.3)	40	106	5.91	200
MPG	TOG	C27/C36	390(15.0)	50	106	5.97	200
MPG	TOG	C27/C36	516(19.8)	70	106	6.08	200
DPG	TOG	C27/C36	86(5.7)	10	106	5.77	200
DPG	TOG	C27/C36	163(10.8)	20	106	5.88	200
DPG	TOG	C27/C36	232(15.4)	30	106	5.99	200
DPG	TOG	C27/C36	295(19.5)	40	106	6.09	200
DPG	TOG	C27/C36	403(26.7)	60	106	6.28	200
DPG	TOG	C27/C36	450(29.8)	70	106	6.37	200
PA	TOG	C27/C36	173(5.2)	20	106	5.77	200
PA	TOG	C27/C36	328(9.9)	40	106	5.87	200
PA	TOG	C27/C36	501(15.1)	65	106	5.99	200
PA	TOG	C27/C36	655(19.7)	90	106	6.11	200
MPG	CHX	C36	18(0.9)	10	4491	9.65	200
MPG	CHX	C36	36(1.7)	20	4491	9.67	200
MPG	CHX	C36	141(6.5)	80	4491	9.80	200
MPG	CHX	C36	327(14.9)	200	4491	10.06	200
MPG	CHX	C27	133(6.5)	80	4491	9.99	200
MPG	CHX	C27	310(14.9)	200	4491	10.22	200
DPG	CHX	C36	62(5.0)	35	4491	9.76	94
DPG	CHX	C27/C36	125(9.9)	15	920	5.84	200
DPG	CHX	C27/C36	198(15.5)	25	920	5.95	200
DPG	CHX	C36	263(20.5)	35	920	6.05	200
DPG	CHX	C27/C36	322(24.8)	45	920	6.15	200
DPG	CHX	C27/C36	401(30.6)	60	920	6.29	200
PA	CHX	C36	80(2.9)	9	920	5.72	200
PA	CHX	C27/C36	131(4.7)	15	920	5.75	200
PA	CHX	C27/C36	289(10.4)	35	920	5.86	200
PA	CHX	C27/C36	432(15.4)	55	920	5.95	200

3 Equilibration and convergence

A common concern in molecular dynamics is the equilibration of the system and convergence of the values of the target measurables. As triolein is a highly viscous solvent, specific care must be taken to establish these rigorously. To determine the length of the initial relaxation period, the time evolution of the free surfactant concentration was first assessed (data not shown). Within 0-25 ns, the free surfactant concentration leveled and started to fluctuate around an average value. Drift in the free surfactant concentration and sampling time required for the system to reach a converged free surfactant concentration were additionally inspected by varying the initial relaxation time, i.e. the time period omitted in the analysis. The analysis indicated that 50-75 ns was a sufficient sampling time to obtain a converged value of free surfactant concentration. Finally, the statistical uncertainty of the size distributions were estimated by the standard method of blocks, i.e. dividing the trajectory into four blocks, and calculating the standard deviation of the block averages from the mean, see Figures S5 and S6.

In principle, the formation of large reverse micelles in the simulations could be prevented by existence of a high kinetic barrier as the simulations are self-assembly simulations. To assess whether this is plausible, an additional

simulation starting from an initial configuration with two pre-assembled micelles of 35 monopalmitins each in triolein was performed. This two micelle system corresponds to 516 mM monopalmitin in triolein. A comparison of the data set resulting from these pre-assembled micelles and those self-assembling in the simulations shows that the two different starting configurations yield similar size distributions and average free surfactant concentrations, see Figure S2. This indicates that the systems are not kinetically trapped.

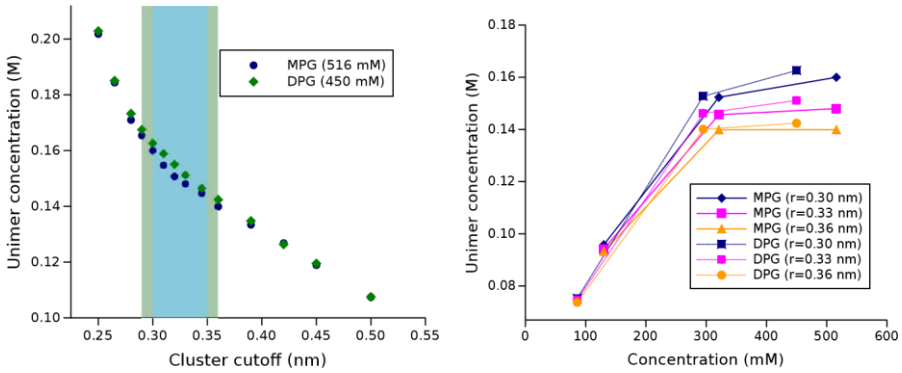

Figure S1: Effect of cluster cutoff on free unimer concentration (left) and its concentration dependence (right). The green and blue regions show the region where the cutoff based, calculated distribution matched a manually, by visual inspection counted size distribution. To reduce computational cost, the time-averaged distributions were calculated from simulation configurations saved every 1000 ps over the last 125 ns of the 200 ns simulation trajectories.

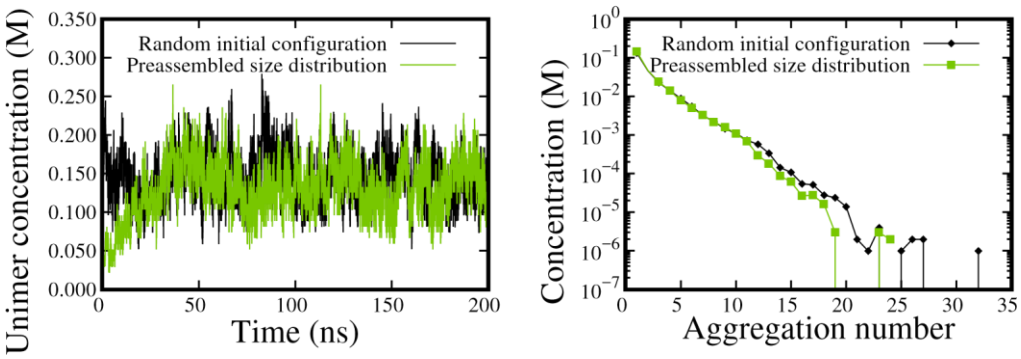

Figure S2: Evolution of free surfactant concentration vs. simulation time (**left**), and size distributions after 50 ns (**right**) calculated from monoglyceride in triglyceride solvent simulations started from randomly dispersed surfactants and two preformed micelles of 35 monoglycerides. Both systems correspond to 516 mM monoglyceride concentration.

4 Influence of cluster cutoff distance on aggregation energy

In dilute solutions, following Equation 3 of the main manuscript, the association free energy of aggregates of size g (aggregation number), $\Delta\mu_g^o$, can be extracted from aggregate size distributions using equation

$$\frac{\Delta\mu_g^\circ}{RT} = -\frac{1}{g}\ln\frac{c_g}{c_1^g} \tag{1}$$

where c_g the concentration of aggregates of size g. In this, the activity coefficients have been neglected as the solution is dilute. Figure S3 plots the $\Delta\mu_g^o$ values calculated using different cluster cutoff distances. The data shows that for all examined cutoff distances, $\Delta\mu_g^o$ shows considerable concentration dependence. Notably, in an ideal solution in which there are no interaggregate interactions, the data would not show concentration dependence, unless the cutoff distance

is poorly chosen. However, the concentration dependence in dimer and trimer aggregation energies persist until the cluster cutoff distance is as small as 0.07 nm which is clearly unrealistically small. In total, this means the concentration dependence has, at least partially, another origin than the cutoff distance choice, e.g. it rises from interaggregate interactions.

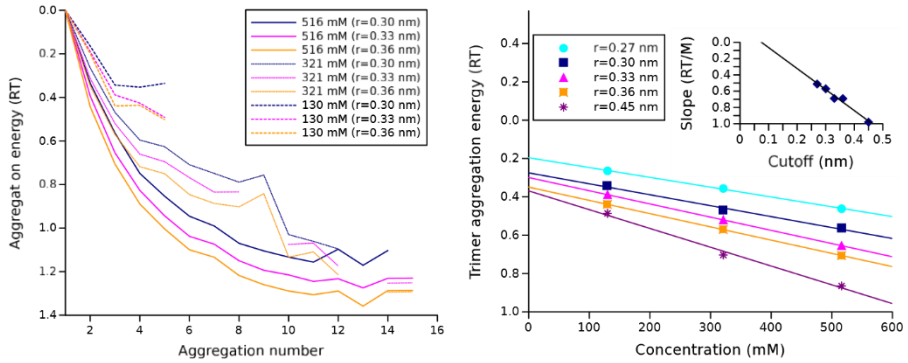

Figure S3: Effect of aggregate cutoff and concentration on the aggregation energy of monopalmitin in triolein (**left**) and monopalmitin trimer in triolein (**right**) in the simulations.

5 Photographs of experimental TCNQ dye solubilization samples

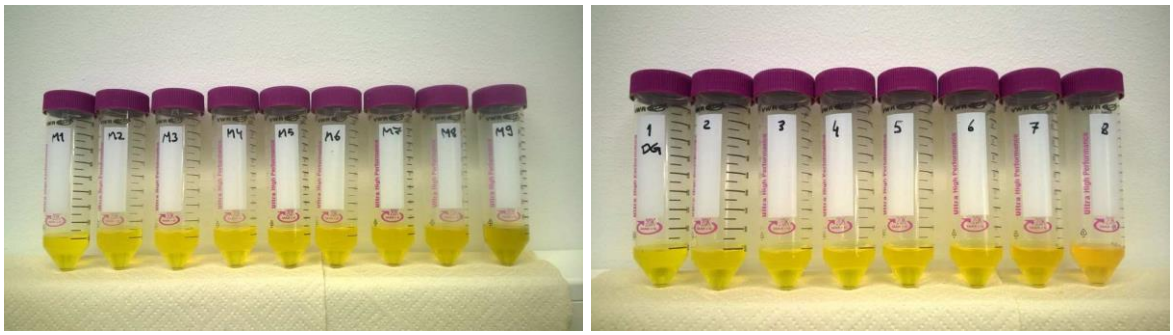

Figure S4: Photographs of the TCNQ dye solubilization samples after TCNQ addition at different concentrations of monoglyceride (**left**) and diglyceride (**right**) in vegetable oil. The samples are ordered such that the monoglyceride and diglyceride concentration increases from left to right.

6 Aggregate size distributions

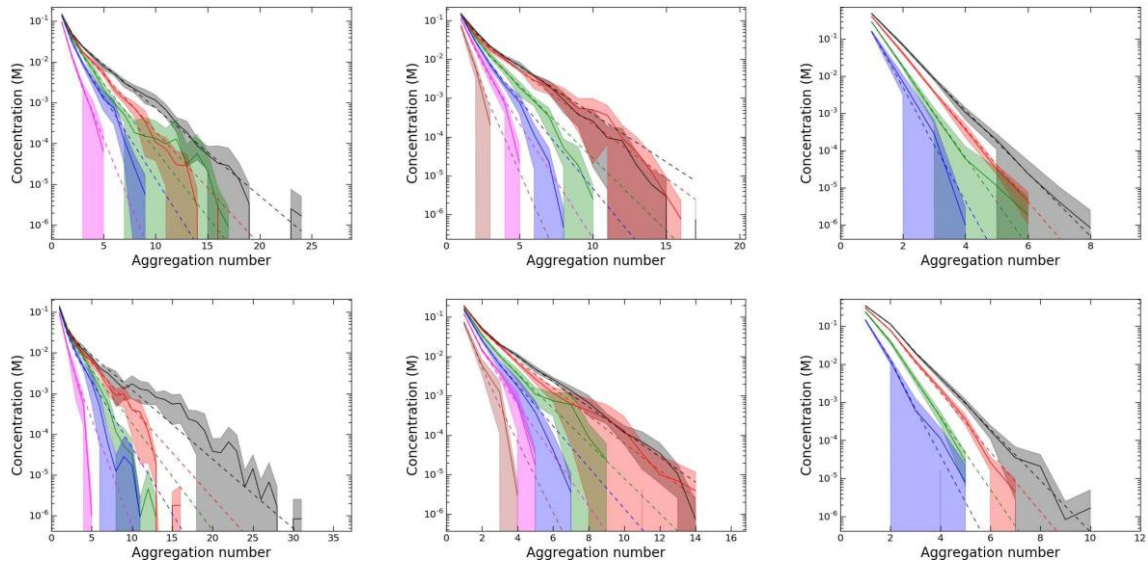

Figure S5: Aggregate size distributions of monopalmitin (left), dipalmitin (middle) and palmitic acid (right) in triolein. Top row corresponds to simulations with C36 model and bottom row to C27 model. Straight lines correspond to simulated size distribution, dashed lines show the fit to Equation 3, while shaded areas indicate the standard deviation calculated based on four blocks. The concentrations of the surfactant decrease in the order black > red > green > blue > pink > brown. Table S1 presents the actual concentrations.

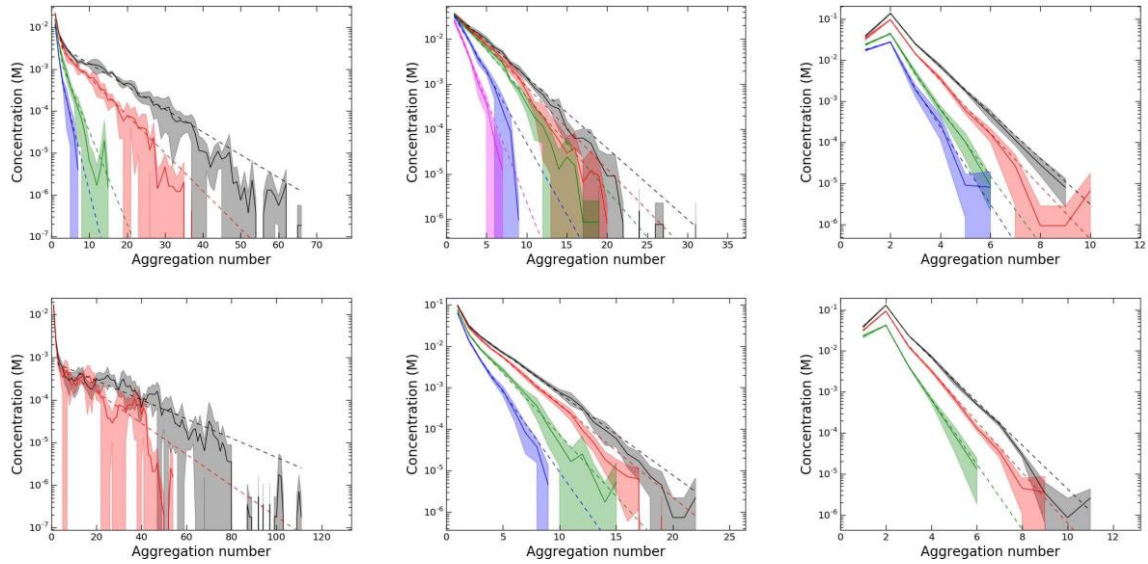

Figure S6: Aggregate size distributions of monopalmitin (left), dipalmitin (middle) and palmitic acid (right) in cyclohexane. Top row corresponds to simulations with the C36 model and bottom row to the C27 model. See caption of Figure S5 for full description of the presented data.