Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Mingning Zhu April 14, 2021

Predicting Micellization Behavior of Carboxylate Surfactants from Molecular Dynamics Simulations

by

Mingning Zhu

James Kindt

Adviser

Chemistry

James Kindt

Adviser

Francesco Evangelista

Committee Member

Michelangelo Grigni

Committee Member

2021

Predicting Micellization Behavior of Carboxylate Surfactants from Molecular Dynamics Simulations

By

Mingning Zhu

James Kindt

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Chemistry

2021

Abstract

Predicting Micellization Behavior of Carboxylate Surfactants

from Molecular Dynamics Simulations

By Mingning Zhu

Surfactants are amphiphilic compounds that form aggregates such as micelles in aqueous solution at concentrations above critical micellar concentration (cmc). However, the structural disorder, polydispersity, and sensitivity to conditions of surfactant micelles present challenges to the unambiguous determination of their properties through experiments alone. The first part of the thesis aims to use molecular dynamics (MD) simulations to predict and compare the self-assembly behavior of octanoates, focusing on the effect of varying counterion identities among sodium, potassium, and tetramethylammonium cations. The second part focuses on predicting the enthalpy change and heat capacity change of micellization of anionic surfactants. Statistics from MD simulations of small systems are input in a global fitting procedure to generate equilibrium constants for surfactant micellization over micelle compositions described by the number of surfactants and the number of bound counterions. The resulting free energy statistics are used to predict the cmc, mean micelle size, and the degree of counterion binding of the investigated carboxylate surfactants. Trends in the fitted values are consistent with experimental data. Simulations at different temperatures are then conducted for each kind of surfactant to analyze enthalpy changes associated with micellization and counterion binding. Enthalpograms will be built for comparison with experimental calorimetry data by explicitly considering polydispersity, so that the model can provide insights on cluster contributions to the enthalpy change of micellization at different solution concentrations.

Predicting Micellization Behavior of Carboxylate Surfactants from Molecular Dynamics Simulations

By

Mingning Zhu

James Kindt

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Chemistry

2021

Acknowledgement

I want to sincerely thank Dr. Kindt for his invaluable mentorship and support throughout my thesis. I am thankful to Xiaokun Zhang, and all the students and faculties for their help and tireless work. I would like to express my gratitude to my committee members, Dr. Kindt, Dr. Evangelista, Dr. Grigni, for their encouragement and guidance.

Table of Contents

1. Overview

1.1 Micelle Self-Assembly

Surfactants have a wide range of applications in daily life, including detergents, wetting agents, emulsifiers and foaming agents. They are usually amphiphilic compounds with a long nonpolar (hydrophobic) hydrocarbon tail and a polar (hydrophilic) head group, which enable them to self-assemble and form aggregates above the critical micelle concentration (cmc). Self-assembly is a spontaneous organization of molecules driven by noncovalent interactions into stable aggregates [1]. One of the most commonly formed types of aggregate is the micelle, where the hydrophobic tails of the surfactants form the core of the micelle and the hydrophilic heads are in contact with the surrounding liquid. A schematic representation of a micelle structure is shown in Fig. 1.

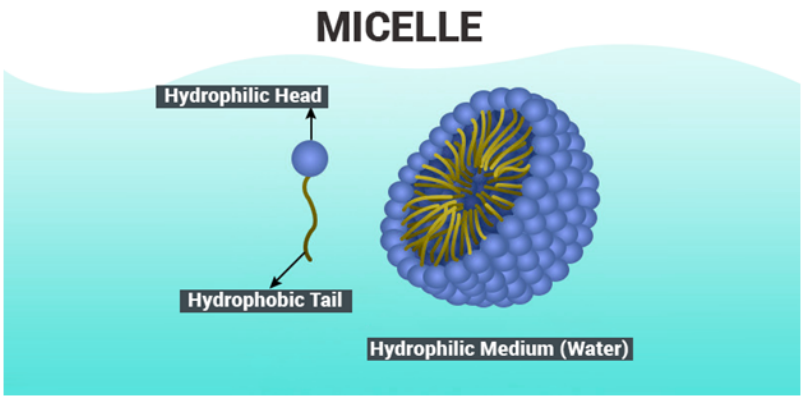


Figure 1. A schematic representation of micelle structure [2].

The driving force for self-assembly of surfactants mainly stems from the hydrophobic effect, which is considered entropy driven. The favorable entropic effect explains self-assembly as a result of an increase in water disorder when hydrophobic surfaces dispersed in water aggregate and lessens the surface area around which water molecules are more aligned [3]. This process where structured hydration water around the hydrophobe is released into the bulk is endothermic, while the transfer of the hydrocarbon chains into the micelles and

restoring the hydrogen bonding structure of the water around the micelles are assumed to be exothermic [4].

In addition, the process of rearrangement of the hydration water of the head groups upon incorporation in the micelles is also considered endothermic as a consequence of monomer association and counterions condensation [5].

The investigated surfactants in this study are alkanoates, which are fatty acid molecules with negative charges, specifically octanoates (8 carbons) and decanoates (10 carbons). The molecular structure of octanoate (OA) is shown in Fig. 2.

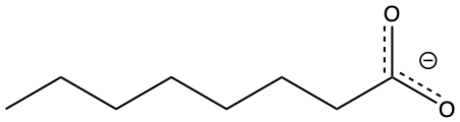


Figure 2. Chemical structure of octanoate

These anionic surfactants are characterized by low toxicity and high biodegradability, and are applied in various fields from washing, material recovery processes, environmental clean-up to encapsulation and drug delivery [6]. The self-assembly behavior of alkanoates depends on various factors such as the length of the hydrophobic tails, the concentration, the temperature, pH, and the binding to counterions. This study focuses on the high-pH limit, with negligible protonation of surfactants. Micellar properties, such as the critical micelle concentration (cmc), aggregation number, free energy change of micellization, are critical to a robust understanding of the driving forces and mechanics behind the self-assembly of ionic surfactants in different systems.

1.2 Isothermal Titration Calorimetry

Our experimental collaborator [7] investigated the micellization behavior of the long-chain carboxylates by isothermal titration calorimetry (ITC). ITC is one of the most powerful

techniques for thermodynamic characterization of aggregation processes. It is a physical technique that directly measures the heat discharged or consumed all along a chemical reaction [8].

The instrument consists of two cells: main and reference cells. Both cells contain solvent (water) and are kept at constant temperature and pressure [8]. The surfactant solution (5-15 times the value of cmc) was placed in a 250 μL syringe and titrated into the main cell [7]. The dilution of concentrated surfactant solution would result in heat change, which causes a change of temperature. Therefore, the instrument transfers the necessary power to maintain the temperature of the main cell unchanged. The heat change is then calculated by integrating the power over time and is proportional to the heat effect expressed per mole of added surfactant per injection. By integration of all peaks, the experimental heats of dilution as a function of surfactant concentration were obtained and the resulting figure is named an enthalpogram. A schematic representation of the instrument is shown in Fig. 3.

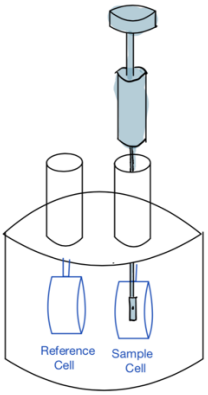


Figure 3. A schematic representation of the ITC instrument [8]

1.3 Molecular Dynamics Simulation

The structural disorder, polydispersity, and sensitivity to conditions of surfactant micelles present huge challenges to the unambiguous determination of their properties through experiments alone. Molecular dynamics (MD) simulation has long been used as a tool to obtain useful insights into the structures of micelles and their interactions with proteins and

small molecules they can solubilize [9]. It is a computer simulation method that permits the prediction of time evolution of an interacting particular system [10]. A model system consisting of N particles is selected and physically important conditions are chosen for the systems. Commonly used ensembles include canonical ensembles (NVT) and isothermal-isobaric ensembles (NPT). NVT fixes the number of molecules (N), volume of the system (V) and the temperature (T). NPT fixes the pressure of the system (P) instead of V. Modified Newton's equations of motion with the purpose of maintaining the average temperature (thermostat) are solved until the properties of the system no longer change with time, indicating that the system has been equilibrated [11]. All the actual measurements of observables will be performed after equilibration. Fig. 4 presents an initial snapshot of one of the investigated systems, with 40 sodium octanoates and 3812 water molecules.

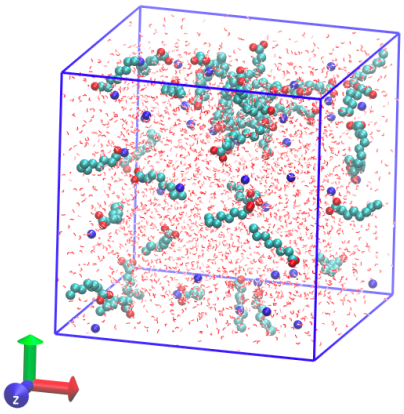


Figure 4. A snapshot from an MD simulation. Blue beads: sodium cations; cyan beads: alkyl tails; Red beads: oxygen atoms

The most time-consuming and critical part of MD simulations is the calculation of force acting on every particle in the system [11]. This is derived as a gradient of the potential energy between particles with respect to each coordinate of each particle. The force field refers to the functional form and parameters used to calculate potential energy. Usually, force field includes bonded terms and nonbonded terms. Bonded terms describe the bond, angle, and dihedral motions, while the nonbonded terms describe the long-range electrostatic and

van der Waals forces. In this study, a united-atom model is adopted to model alkyl tails of surfactants where each carbon and its bonded hydrogens are united into a single site, leading to the introduction of CH_3, CH_2, CH, and C pseudo atoms [12]. The force field parameters used in the current work are presented and discussed in Chapter 2.2.1 and Chapter 3.2.1.

2. Counterion Effect on Long-chain Carboxylate Surfactant Assembly

2.1 Introduction

Due to the importance of fatty acid surfactants, there have been a large number of experimental and theoretical studies focused on long-chain carboxylate micellar systems. In 1959 White and Benson investigated aqueous potassium octanoate (KOA) solutions at different temperatures using electrical conductivity measurements [13], and Campbell studied the micellization behavior of sodium octanoate (NaOA), sodium decanoate (NaDA), sodium dodecanoate (NaC12), and sodium tetradecanoate by applying both surface tension and electrical conductivity measurements [14, 15]. More recent experimental studies focused on the interactions between divalent or trivalent metal ions and NaOA, NaDA, and NaC12 [16]. Furthermore, NaOA and NaDA have been the subjects of intense MD simulation studies. These studies provided molecular-level details into the structures of the aggregates and the thermodynamics of aggregation [17]. Kalil, et al. not only confirmed that the hydrophobic effect is the main contribution to the self-assembly behavior of surfactants and the overall process is largely entropy-driven through simulations, but also observed the crucial favorable entropic contribution from the surfactant heads. This contribution can be attributed to the partial de-solvation of the octanoate head due to the micelle formation and the increase in counterion association during micellization.

Therefore, the nature of the counterion plays an important role in ionic micelle properties and the thermodynamic characterization of aggregation. An increase in the solution ionic strength leads to a lowering of cmc of ionic surfactants and has been used as a tool to

induce micelle growth. This effect has been attributed to the screening of the electrostatic repulsions between the charged ionic surfactant heads from counterions added to the solution [18]. Another reason lies in the common ion effect. The addition of counterions (X^+) would shift the equilibrium: $X^+ + OA^- \leftrightarrow X^+OA^-$ to the right, increasing the degree of micellization. Among the studies of the effect of counterions on surfactant self-assembly, most of them have been performed on inorganic counterions such as alkali ions, while fewer of them focused on the effect of organic counterions. The association behavior of organic counterions can be anticipated to be different from that of inorganic counterions due to their large sizes and hydrophobic characters. The charged headgroups and hydrophobic alkyl chains of the counterions bring more possibilities of interactions in the system, including strong electrostatic interactions between the anionic micelles and cationic counterions and the hydrophobicity of counterions [19]. Previous research has been performed to investigate the different effects of alkali and organic alkylmethylammonium cations on dodecylsulfate [20]. Benrraou, et al. found that at room temperature, the cmc of tetramethylammonium (TMA) dodecylsulfate is lower than sodium dodecylsulfate and the formed micelles are also larger with the presence of TMA, indicating that micellization is favored by the increased hydrophobicity of alkylmethylammonium cations. However, no one to the best of our knowledge has studied the effects of these counterions on carboxylate surfactants.

In the present work, we aim to use MD simulations to predict the self-assembly behavior of long-chain carboxylate surfactants, OA, with Na^+, K^+, and TMA cations. Fig. 5 presents the chemical structure of TMA.

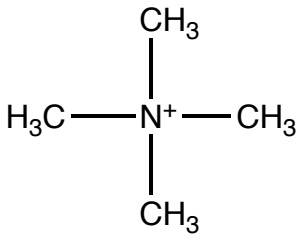


Figure 5. Chemical structure of TMA.

We input statistics from small systems and generate equilibrium constants of every cluster composition observed in the simulation box through a global fitting procedure called the PEACH method [9]. The resulting dependence of free energy on micelle size and charge are rationalized by a phenomenological four-parameter model. Specifically, the free energy of micellization of a micelle composition ($\Delta G_{i,j}$) indicates the Gibbs free energy of transfer of i monomers and j counterions into the micelle per mole of surfactant. They are then used to predict cmc, micelle size, and the degree of counterion binding versus concentration of bulk solutions. The statistics are compared to ITC data measured by our collaborators [7]. As the fitted thermodynamic data are sensitive to the definitions of clusters, Chapter 2.3.1 also explores more accurate definitions of octanoate clusters with both alkali and organic cations, providing molecular-level details in the interaction dynamics between anionic surfactants and different cations.

2.2 Methods

2.2.1 General Simulation Setup

All simulations are performed with GROMACS 5.0 software package [21]. Newton's equations of motion are evaluated with the leapfrog algorithm integrator with a time step of 2 fs. Verlet list scheme that uses properly buffered lists with exact cut-off is utilized to adapt to the streaming architecture of GPUs. The TraPPE-UA potential [12] is parametrized to reproduce the phase behavior of alkanes, and HH-Alkane potential is parametrized to reproduce the enthalpy and free energy of hydration of alkanes [22]. The TIP4P-2005 water model [23] is used for all aqueous phase simulations. OPLS models [24] are selected for parameterizing the Lennard-Jones (LJ) potential energy between cations and carboxylate headgroups and the partial charges of them are adapted from the work of Hess et al. [25]

where these were developed and tested against thermodynamic properties of acetate salt solutions.

The temperature is maintained at a reference temperature of 300K (τ_T = 1.0 ps) with the velocity rescaling thermostat [26]. NpT simulations that implement the Berendsen barostat [27] with a reference pressure of 1 bar (τ_p = 5.0 ps and compressibility = 4.5 × 10^{-5} bar^{-1}) are conducted.

The Lorentz-Berthelot combining rule for LJ interactions between different atoms is used for TIP4P-2005 water, octanoates and counterions.

$$\sigma_{ij} = \frac{1}{2}\left(\sigma_{ii} + \sigma_{jj}\right), \varepsilon_{ij} = \sqrt{\varepsilon_{ii} + \varepsilon_{jj}} \quad (1)$$

Electrostatic interactions are calculated by the particle mesh Ewald method [28]. The short-range neighbor list and Coulombic cutoff radius are 1.4 nm. Van der Waals interactions implement a switch function at 1.2 nm with a cutoff radius of 1.4 nm. Table 1 presents the number of molecules, box volume, trajectory length and the pre-equilibration length of each system.

Table 1. The number of molecules, box volume, trajectory length and the pre-equilibration length of each system.

	Number of octanoates	Number of water	Trajectory Length (ns)	Pre-equilibrium Length (ns)
NaOA	40	3812	1800	200
	40	4500	1900	100
KOA	40	3812	1950	50
	40	4500	1950	50
TMAOA	30	2859	550	50
	40	2859	1950	50
	40	3812	1950	50
	50	3812	550	50

2.2.2 Cluster Definition

The ionic aggregate model [29] is used in determining the cluster compositions. r_{tail} (the cutoff distance between two OA chains) and r_{head} (the cutoff distance between one OA and

one counterion) are used to define a i-mer cluster with i surfactants and j counterions. Two OA chains are considered to be in the same cluster if the distances between either of their alkyl tail beads are within r_{tail}. The cutoff distances are adjusted according to the quality and convergence criteria of their PEACH fits (Chapter 2.2.3). Alkali cations that are within a cutoff distance r_{head} from either of the oxygen atoms of the octanoate are considered to be associated with that OA chain's cluster, as the primary interaction between counterions and surfactants is the electrostatic attraction. The radial distribution function (RDF) in a system of particles describes how density of particles varies as a function of distance from a reference particle. The first peak of the RDF function indicates the formation of contact ion pairs (CIP), where the ions of opposite charges are in direct contact with each other. The second peak of the RDF function indicates the formation of solvent shared ion pairs (SIP), where there is one solvent molecule between the cation and anion [30]. Previous research has confirmed that the cation specificity of the thermodynamics of alkali acetate solutions is determined mainly by SIP [25]. Thus, r_{head} of alkali cations is determined as the second minimum of the RDF functions of counterions over O1 on the octanoate.

However, for organic cations, the methyl groups on TMA can also hydrophobically interact with the tail of octanoates as Fig. 6 shows.

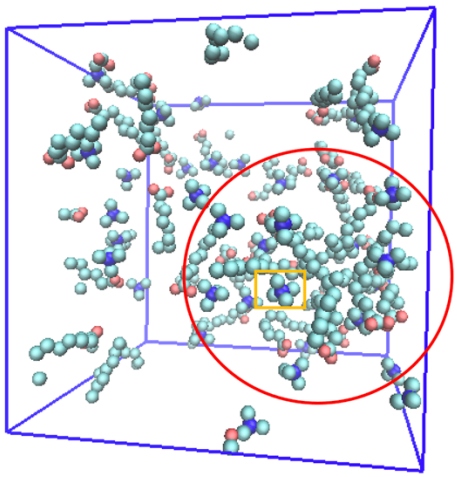


Figure 6. Snapshot of the TMA system using VMD [31], not showing water. Blue beads: N^+ on TMA cations; green beads: alkyl tails; Red beads: oxygen atoms. The circled micelle has TMA cations (orange boxed) inserted into the core.

Therefore, TMA cations with central N atoms that are within a cutoff distance r_{head} from any atom of OA are considered to be associated with that OA chain's micelle, and the cutoff distances are taken as an average from the RDF calculations of N^+ on TMA over every atom on OA. When a counterion bridges two different micelles, that counterion is treated as dividing its time evenly between the two micelles.

2.2.3 Partition-Enabled Analysis of Cluster Histogram (PEACH)

PEACH method is a global fitting strategy developed by Xiaokun Zhang, et al. [9] to find the free energy surface for aggregation of surfactant from simulations of small systems containing one or two micelles and finite number of molecules (N).

To initialize the fitting, histograms of $<n_{i,j}>$ representing the mean numbers of clusters of surfactant aggregation number i and number of associated counterions j averaged over each trajectory are generated. Each trajectory is split into 10 segments, and the standard deviation for $<n_{i,j}>$ among those segments is calculated. To account approximately for the effect of free volume on cluster statistics, we follow procedures previously used in SOS cluster simulation [9], where an effective free volume (V) that is occupied by the solvent (water) alone is used in the PEACH calculation. The histograms are then used as inputs for the PEACH method to find a globally optimized set $\{K_{i,j,fit}\}$ of equilibrium association constants for each cluster size applicable in the limit of large N. The method involves assigning an association constant to each possible cluster size i with j counterions, generating a fitted cluster size distribution for each simulation based on its total N and V, and adjusting it iteratively to find the best set $K_{i,j,fit}$ to reproduce the cluster size distributions from simulations. The initial guess of $K_{i,j,fit}$ and $K_{i,j,sim}$ are based on

$$\sigma_{i,j} = \frac{K_{i,j,sim}}{K_{i,j,fit}} = \frac{<C_{i,j}>_{sim}}{<C_{1,0}>^{i}_{sim} * <C_{0,1}>^{j}_{sim}} / \frac{<C_{i,j}>_{fit}}{<C_{1,0}>^{i}_{fit} * <C_{0,1}>^{j}_{fit}} \quad (2)$$

$$K'_{i,j,fit} = K_{i,j,fit} * \sigma_{i,j}{}^{\alpha}, \text{ where } \alpha \text{ is an adjustment factor of 0.05} \quad (3)$$

$< C_{i,j} >_{sim}, < C_{1,0} >_{sim}, < C_{0,1} >_{sim}$ are obtained from the input cluster histograms, and the initial guess of $K^{0}_{i,j,fit}$ is taken as the weighted geometric mean of $K_{i,j,sim}$ from all trajectories that feature clusters that have i surfactants and j counterions.

The standard Gibbs free energy of each cluster composition is calculated by

$$\frac{\Delta G_{i,j}}{k_B T} = -lnK_{i,j} - (i-1) * lnc_{1,0} - j * ln(c_{0,1}) \quad (4)$$

2.2.4 Predicting Micellar Properties

The bulk equilibrium distribution of cluster concentrations given the monomer OA (i) and counterion (j) concentrations is given by

$$c_{i,j} = K_{i,j} * c^{i}_{1,0} * c^{j}_{0,1} \quad (5)$$

where $K_{i,j}$ is the value obtained by the PEACH method. The method of Zhang et al. [9] is used to find pairs of OA and counterion monomers $c_{1,0}$ and $c_{0,1}$ that yield a charge-neutral system at a given total concentration, allowing concentrations $c_{i,j}$ for all cluster compositions to be predicted as a function of the total concentrations.

2.2.5 Maibaum's Model

To better understand the components of the energy contribution to the free energy of micellization and to obtain a smooth free energy surface with few noises from simulation, we fit the PEACH-derived free energy surface to a modified phenomenological model originally proposed by Maibaum, et al [32]. For a cluster composition with i OA and j counterions,

$$\frac{\Delta G_{model}(i,j)}{k_B T} = -(i-1)\Delta\mu + g(i^{\frac{2}{3}} - 1) + h(i^2 - 1) - \varepsilon j - k_B T ln(\frac{i!}{(i-j)!j!}) \quad (6)$$

The first and second terms, corresponding to a linear dependence of the free energy of transfer of a hydrophobic molecule from water into the micelle and a surface energy term scaling as the surface area, describe the free energy of nucleating oil clusters in water driven by hydrophobic interactions. The third term describes the entropic cost of packing surfactants. The fourth term describes the favorable energy contribution of counterion binding to headgroups. We assume, to a first approximation, that binding of counterions to headgroups is independent of the degree of association of the micelle or the presence of other counterions and thus the free energy of micellization is linearly proportional to the number of counterions [9]. (These assumptions about counterion binding will be tested in Chapter 3.3.2) The final term is a combinatorial factor describing the entropic contribution of distributing j counterions among i surfactant headgroups [9]. In this study, we pay particular attention to ε, which gives us a quantitative analysis of the strength of counterion binding of alkali and organic cations.

2.3 Results and Discussion

2.3.1 Cutoff-Distances to Define Clusters

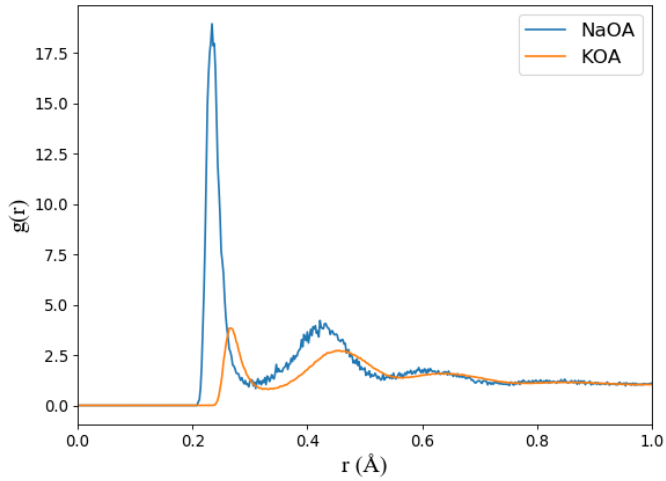


Figure 7. RDF functions of Na^+, K^+ over O1 on the octanoate

Fig. 7 shows the calculated RDF functions of Na^+, K^+ over O1 on the octanoate. N_{jc}, the coordination number of component c around component j, can be calculated by integrating over RDF.

$$N_{jc} = \rho_c 4\pi \int_0^\infty (g_{jc}(r) - 1) r^2 dr$$

$$g_{jc}(r) = \frac{dn_{jc}(r)}{4\pi r^2 dr \rho_j} \; where \; \rho_c = \frac{V}{N_j} \quad (7)$$

The observation is in qualitative agreement with experimental XAS measurements [33, 34] and the computational study by Hess, et al [25]. r_{head} of alkali cations is taken as the position of the second minimum to include both ions in direct contact and ions separated by one solvent from the headgroup. For organic cations, since we need to account for the hydrophobic interaction between the methyl groups on TMA and the tail of octanoates, r_{head} is taken as an average from the RDF calculations of atoms on TMA over every atom on OA. To access the extent of micelle insertion of TMA, we plot the RDF functions of N^+, Na^+, and K^+ over the tail carbon (C8) on OA and is shown in Fig. 8. Fig. 8 also shows the RDF function of N^+ on TMA over O1 on OA, which represents the dominant electrostatic interaction between TMA and OA. From a qualitative interpretation of the integration of RDF functions in Fig. 8, we can see that there's a comparable association between TMA and alkyl tails and TMA and acetate < 1nm (the region where ion pairs form), whereas the association between alkali ions and alkyl tails is negligible. It further confirms that hydrophobic interaction between the TMA counterions and surfactants cannot be neglected in accounting for its micelle formation.

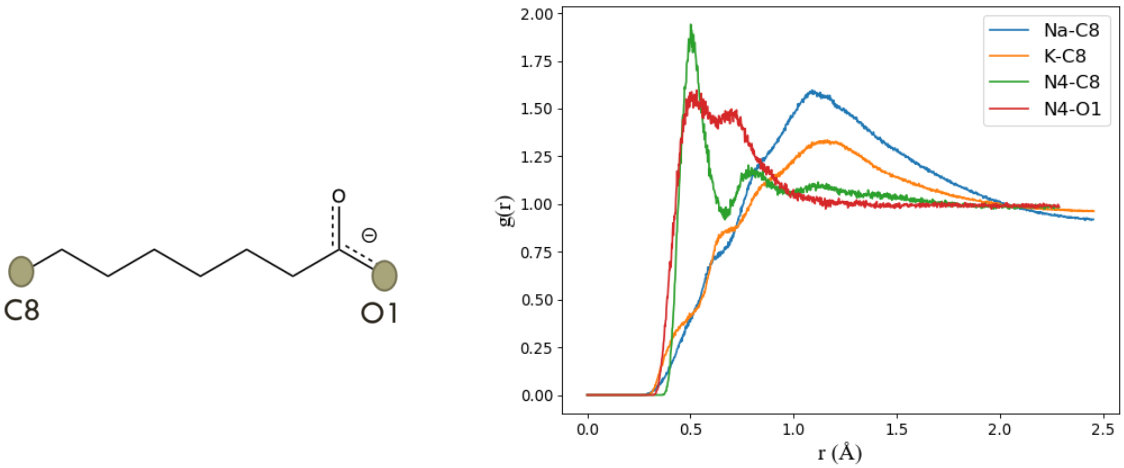


Figure 8. RDF functions of Na^+, K^+, N^+ over C8 on OA and the RDF function of N^+ over O1 on OA.

The value of r_{tail} is directly adapted from the value that Xiaokun, et al. uses in the simulation of octyl phosphocholine [35], and it is adjusted in the TMAOA system according to the convergence criteria of the PEACH fit. It is worth noting that changing r_{head} and r_{tail} makes small differences to the qualitative trends observed in these three systems, although they yield small quantitative changes to cmc, mean micelle size and the micelle charge. Table 2 presents the optimal cutoff distances used to define clusters of NaOA, KOA and TMAOA.

Table 2. The optimal cutoff distances are used to define clusters of NaOA, KOA and TMAOA.

	r_{tail} (nm)	r_{head} (nm)
NaOA	0.45	0.52
KOA	0.45	0.56
TMAOA	0.40	0.60

2.3.2 Micellar Properties derived by PEACH

Fig. 9 (Left) shows the cluster distribution of micelles irrespective of counterion numbers (summing $<n_{i,j}>$ over all j) at the same concentration. The fit is successful in reproducing the distribution of surfactant compositions of micelles of the simulated systems. The poorest agreement is found in the TMAOA system, where the fewest and smallest clusters are

observed. A higher concentration is required for TMAOA to form micelles as shown in Fig. 9 (Right).

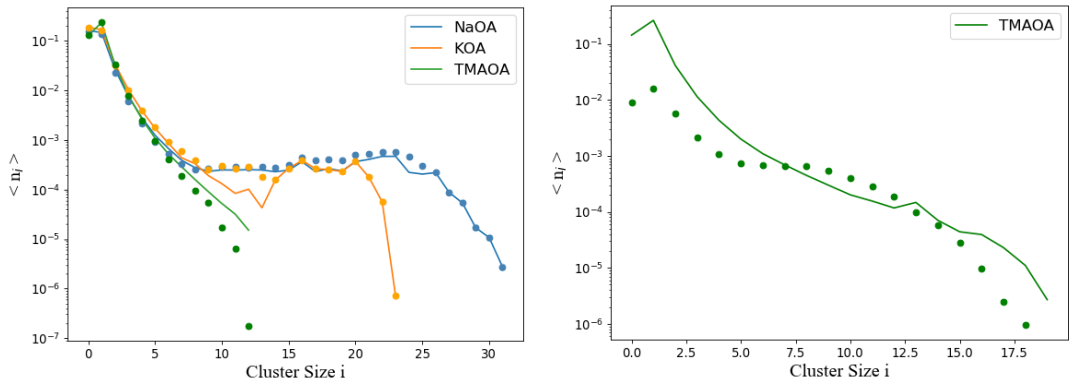


Figure 9. Left: Cluster distribution of NaOA, KOA, TMAOA irrespective of counter ion numbers at a concentration of 40 surfactants and 3812 water. Right: Cluster distribution of TMAOA irrespective of counter ion numbers at a concentration of 50 TMAOA and 3812 water. Dots: Raw data from PEACH input; lines: the PEACH fit

Since the concentration in the experimental data is expressed in molarity (mol/(L of solution)), all the concentrations in Fig. 10, and Table 3 are converted from molality (mol/(kg of solvent)) to molarity. Fig. 10 plots the fraction of OA in micelles over the total concentration of surfactants where micelles are defined as clusters having more than 7 OAs. cmc is approximated by estimating the point where there's a huge increase in the fraction of OA in micelles. We observe that there's not a large difference between the cmc of alkali cations, although Na^+ has a slightly smaller value. That's because the parameters of partial charges and LJ parameters are very similar to the oxygen atoms in both water and acetates. Thus, the alkali cations will not have a strong preference for either acetate or water oxygen, so the difference in alkali cations won't affect the interaction between them and surfactants much. On the other hand, the insertion of TMA cations into the hydrophobic tails significantly affects the onset of micellization of carboxylates. We are very surprised to see that the counterion effects follow the opposite trend of the dodecyl sulfate, where Na^+ induces

the most micelle growth with the lowest cmc value. These inserted TMA ions disrupt the hydrophobic interactions between surfactant tails so much that lead to a delay in the micellization, while in dodecyl sulfate systems, the hydrophobic attraction to OA and the dehydration of TMA cations due to their hydrophobic characters are considered to induce micelle growth.

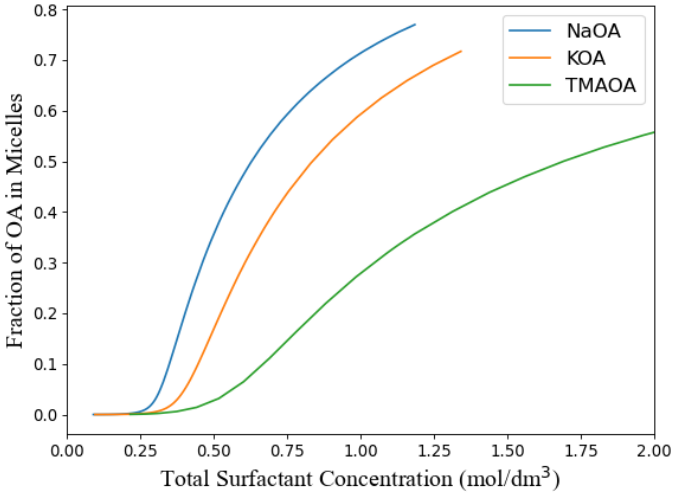


Figure 10. The fraction of OA in micelles over the total concentration of surfactants.

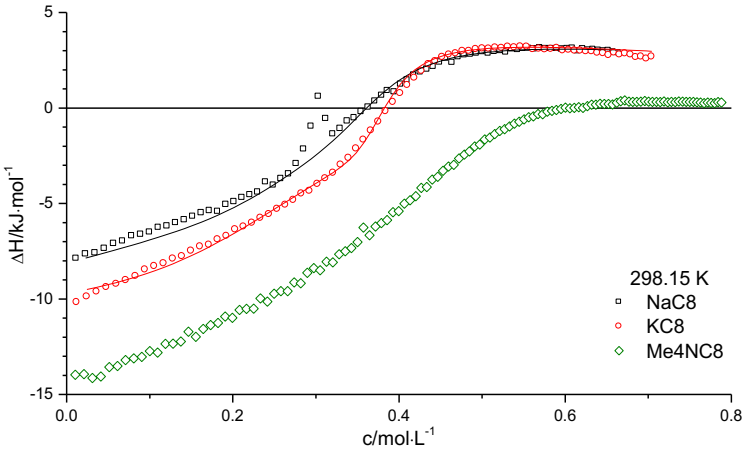


Figure 11. The enthalpogram of NaOA, KOA, TMAOA measured by ITC (Note: this is from personal communication with Žiga Medoš)

ITC measurements are performed later to verify our results and are shown in Fig. 11. Since the formation of large micelles is an endothermic process, the concentration at which micellization happens fastest coincides with the concentration where there's most rapid

change with the enthalpy. Thus, the inflexion point of the enthalpogram gives an approximation of the cmc values, and the PEACH-derived trend is confirmed by the ITC measurement. More detailed understanding and prediction of enthalpogram features from simulation data will be explored in Chapter 3.

Finally, we compare the degree of counterion (β) binding of alkali ions predicted by PEACH and the two-step model proposed by our collaborators. β is calculated by the ratio of the number of OA in the micelle to the number of counterions in the micelle. At the cmc value, $\beta_{PEACH} = 0.58$ for NaOA, and $\beta_{PEACH} = 0.51$ for KOA, both are within 20% errors of the experimental fit: β_{exp}, where $\beta_{exp} = 0.8 \pm 0.1$ for NaOA, and $\beta_{exp} = 0.6 \pm 0.1$ for KOA. The observation is consistent with the law of matching water affinities where the kosmotropic Na^+ interacts more strongly with the kosmotropic acetate anion [36].

2.3.3 Modelling the PEACH-Derived Free Energy Surface

Free energy surfaces are calculated by Eq. 4 using the PEACH-derived equilibrium constants of different cluster compositions. Fig. 12 shows the 2-D and 3-D representation of the free energy surface of NaOA. The plots of the other two systems are shown in Fig. S1 and Fig. S2. For better visualization of the well in the surface, where the minimum free energy indicates the most stable cluster composition, the concentration of surfactant and counterion monomers are chosen around cmc.

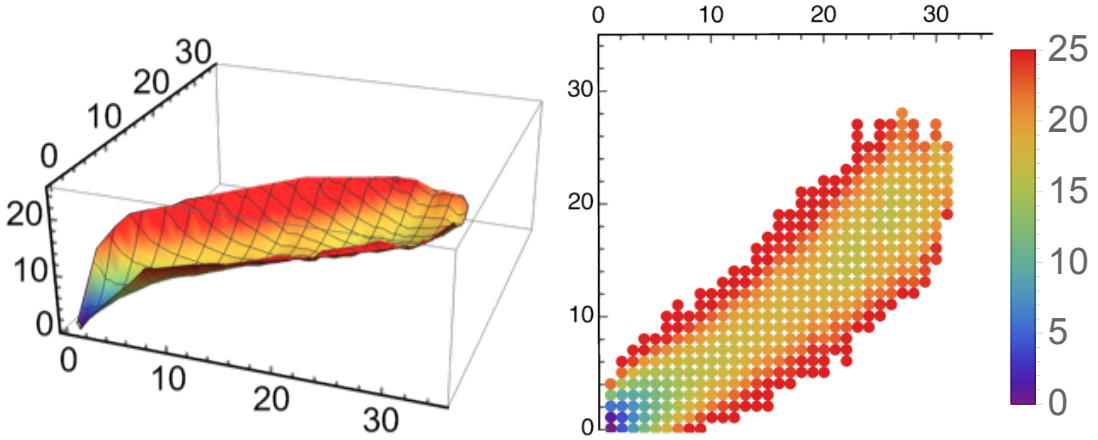


Figure 12. The 3-D (left) and 2-D (right) representation of the free energy surface of NaOA at concentration i = 0.149 mol/dm^3 and j = 0.233 mol/dm^3: x-axis: No. of surfactant; y-axis: No. of counterions; z-axis (color bar): Free energy of formation (k_BT)

The PEACH-derived free energy surfaces are then fitted to the phenomenological Maibaum's Model. Note that due to the addition of a combinatorial factor, overcharged micelles (j > i) are not allowed, but previous work [9] has shown that their impact on the overall cluster distributions is of negligible importance except at very high surfactant concentrations, which are not explored here. To obtain a fitting at cluster composition where the sampling is relatively good, we first fit parameters g, and h by performing a 1-D fit along the cut j = i/2. Then we subtract the nonlinear terms and the combinatorial factor and perform linear regression on the other two parameters $\Delta\mu$ and ε. The first (1-D) fitting surface of NaOA is shown in Fig. 13, and fitting of the free energy surface is shown in Fig. 14. Those of KOA and TMAOA are shown in Fig. S3 – Fig. S6. Note that by comparison with the PEACH-derived free energies, we find that substitution of i^4 for i^2 in Eq. 6 in TMAOA system results in a much better fit to the free energy of micellization.

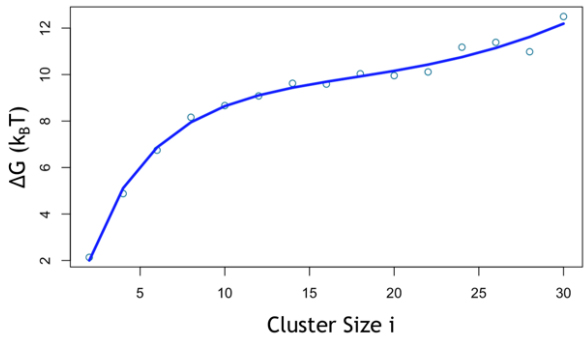


Figure 13. The 1-D fitting of NaOA. Dots: PEACH-derived values; lines: fitted values

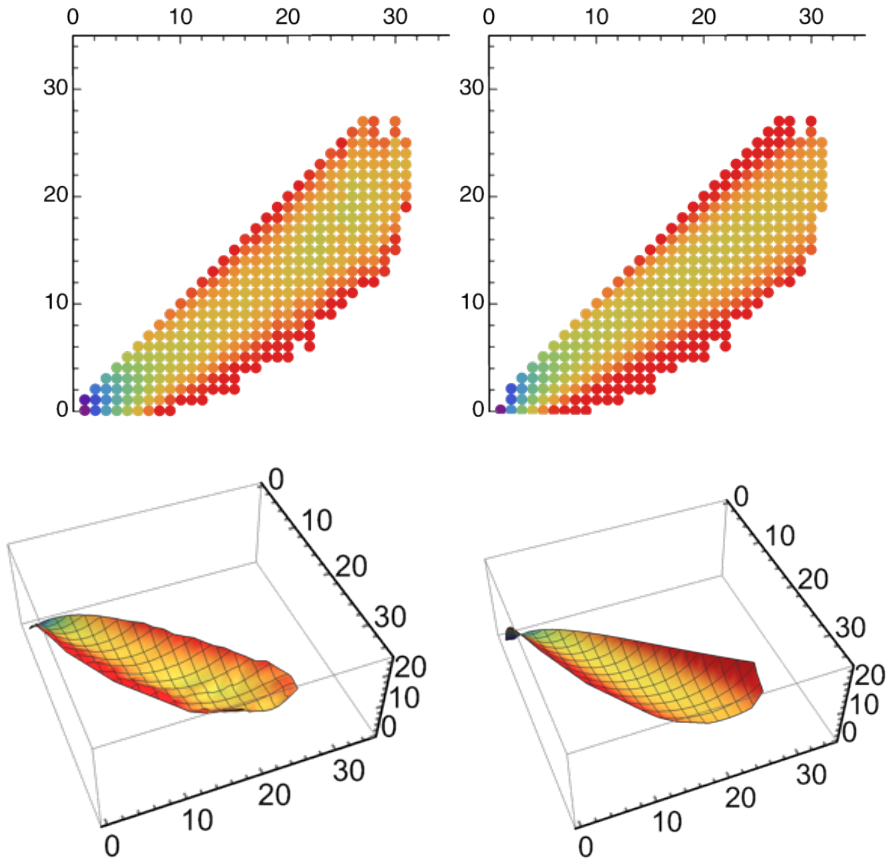


Figure 14. The fitted free energy surfaces of NaOA. Left: 2-D (above) and 3-D (below) representations of PEACH-derived surfaces. Right: 2-D (above) and 3-D (below) representations of fitted surfaces

Table 3 presents the numerical values (with standard errors) of the fitted parameters and the R^2 values to assess the quality of the fit.

Table 3. The numerical values (with std. errors) of the fitted parameters and the R^2 values

	$\Delta\mu\ (k_BT)$	g (k_BT)	h (k_BT)	$\varepsilon(k_BT)$	$R^2(1D)$	$R^2(2D)$
NaOA	2.444±0.0150	10.422±1.156	0.0255±0.004	0.770±0.017	0.986	0.997
KOA	2.615±0.0237	10.777±3.973	0.0193±0.0213	0.446±0.025	0.919	0.992
TMAOA	0.331±0.0139	4.770±1.026	-5.295e-06 ±5.523e-06	0.173±0.018	0.978	0.906

All the R^2 values are greater than 0.90, indicating that the PEACH-derived free energy surfaces do a good job in reproducing the physical dependence of free energy of micellization on cluster compositions. The fitted surface successfully smooths out the statistical noises in simulations and may be used as a good starting point to predict the enthalpy change of

micellization, which will be explored in Chapter 3. ε, the energy of binding per counterion to available headgroup at cmc, indicates the difficulty of counterion binding. A more negative value indicates a stronger attraction between the cations and the surfactants. ε has the same qualitative trend observed in simulations where $\varepsilon_{TMA} < \varepsilon_{K^+} < \varepsilon_{Na^+}$, proving that Maibaum's Model is a good representation of the free energy of micellization.

However, taking a closer look at the fit of every composition, we find that the fitting is poor in the premicellar range (i < 6) and the poorly sampled largest cluster region. The derivation of the exponent of the second term in the Maibaum's model $i^{\frac{2}{3}}$ involves an assumption that the forming micelle is spherical. In the pre-micelle stage, since the cluster is composed of just a few surfactants, the shape of the micelle is far from being spherical, leading to large deviation from the PEACH-derived surface. Concentrations of cluster compositions that are poorly sampled will also be inaccurately represented in the fitting to the Maibaum's model. As illustrated in Fig. S2, the PEACH-derived free energy surface of KOA is rough, indicating that there are larger statistical noises in the sampling of KOA, and the model fitting is not as good as that of NaOA as shown in Fig. S3 and Fig. S4, especially for the largest clusters of each cluster size (boundary of the surface). Another reason behind the poorer fitting of KOA is that the simulations have not reached a concentration where the largest clusters of KOA form, as the free energy surface has not 'curved up' in Fig. S4. Therefore, an accurate functional representation of free energy of micellization requires adequate and accurate sampling of cluster distributions in simulations.

2.4 Conclusion

In this work we present the comparison of the micellization behavior of octanoates in three different systems with counterions Na^+, K^+, and TMA. Equilibrium association constants are obtained from PEACH and are used to calculate free energy of association for each cluster size. They are then used to predict the micelle properties and cmc values are

estimated from the plot of fraction of OA in micelles over the concentration of surfactant solutions. They have a similar trend to experimental data, showing that the cmc of TMAOA is about 25% percent higher than those of KOA and NaOA. These findings are, intriguingly, contrary to trends reported for sodium vs. TMA dodecyl sulfate surfactant [20]. The difficulties of octanoate aggregation in TMA system are partly due to the competing hydrophobic attraction between methyl groups on TMA and the octanoate tails. The core-inserted TMA ions disrupt the hydrophobic interactions between surfactant tails so much that they finally lead to a delay in the micellization.

The physical dependence of PEACH-derived free energy of micellization on cluster compositions can be successfully described by a phenomenological model. The model smooths out the statistical noise and gives a quantitative estimation of energy contributions of surfactant heads, tails and counterions. In the next chapter, we aim to build an enthalpogram model based on the PEACH-derived free energy surfaces and reproduce experimental trends in the enthalpy changes of micellization of long-chain carboxylates, so that we can offer more physical insights into the thermodynamics of surfactant self-assembly.

3. Predicting the Enthalpy Change of Micellization

3.1 Introduction

In Chapter 2, we successfully used a global fitting strategy and optimized cluster definitions to investigate the free energy of micellization in three carboxylate systems with different counterions. However, it remained an inherent difficulty associated with the estimation of the change of thermodynamic quantities across different temperatures. Studies have shown that zwitterionic surfactants typically exhibit a minimum in the cmc as a function of temperature [37]. The minimum reflects the point of compensation between two opposing effects: the decrease in the hydration of the hydrophilic headgroup upon increasing the temperature, which favors micellization or a lower cmc (as the hydrophobicity of the whole

molecule increases), and the disruption of water structure around hydrophobic portions of the surfactant as the temperature increases, which disfavors micellization and promotes an increase in the cmc [38]. Del Rio and Jones [39] showed that during the micellization of surfactants, the standard enthalpy change (ΔH°) sometimes changes its sign and contributes only slightly in this region to the standard free energy change (ΔG°) at room temperatures. Huang and Chandler [40] showed that the weak van der Waals interactions between nonpolar solutes or planar hydrophobic surfaces and water, a crucial component of the unfavorable enthalpic contribution to surfactant self-assembly, affects the excess chemical potential of the solution but not its temperature dependence. More recently, Kalil [17] used the method of potential of mean force (pmf) and umbrella sampling to study the micellization behavior of NaOA, separating the Gibbs energy of association and its enthalpy ($\Delta H_{micellization}$) and entropy ($\Delta S_{micellization}$) terms into contributions arising from the polar headgroup and the hydrophobic tail. The analysis relies mostly on the average pmf and gives fewer attention to the variety of processes that take place in a single association event, such as the fluctuations of counterion concentration at the micelle interface. The computational results were partially confirmed by electric conductivity [41] and ITC measurements [7] of the micellization process of carboxylates.

It was found by our experimentalists that the heat capacity of micellization (ΔC_p) is negative due to the removal of water molecules from contact with nonpolar surface area upon micelle formation for nonionic surfactants [42], which is also found to be the case for NaDA even though the value is less negative [7]. It means that $\Delta H_{micellization}$ decreases with increasing temperature. The micellization process is observed to be endothermic up to ~320K, and then it becomes exothermic for NaOA and most other long-chain carboxylates [7]. Similar trends were observed for alkali decyl and alkali dodecyl sulfates [43]. When temperature increases, the structure of water molecules in the aqueous phase is disturbed, and

consequently, the role of hydrophobic effect and dehydration of counterions and polar surfactant head becomes weaker since less energy is required to break up the original water structures [7]. Therefore, the process becomes more exothermic. On the other hand, the entropic contribution to the free energy is always positive, reflecting the entropic gain upon release of water molecules around the hydrophobic tails, whereas the increase in entropy ($\Delta S_{micellization}$) decreases as temperature increases due to the disruption of original water molecules [38]. In conclusion, as temperature increases, both $\Delta H_{micellization}$ and $\Delta S_{micellization}$ decrease.

However, the interpretation of ITC experiments depends on complex modeling since $\Delta H_{micellization}$ is related to both the cluster size distribution and the enthalpy change of formation of every cluster composition present in the solution. Our collaborators proposed a two-step model where they approximated micelle polydispersity to two different cluster sizes to obtain $\Delta H_{micellization}$ [7]. In this study, we hope to find a more realistic model with fewer parameters to capture the effect of polydispersity and enthalpic contributions from different cluster compositions that is in qualitative agreement with the ITC measurements, with particular emphasis on the physical events that happen at concentrations smaller than the cmc (pre-cmc stage). The experimentalists observed that the $\Delta H_{micellization}$ at pre-cmc stage is temperature-independent and hypothesized that it is dominant by ion-pair formation. We hope to test out their hypothesis by explicitly considering cluster size distribution and investigating the amount of contribution from individual clusters.

Based on the Van't Hoff equation, $\Delta H_{micellization}$ can be determined by a change in the $\Delta G_{micellization}$ to the change in temperature. Under the circumstances, direct assessment of $\Delta H_{micellization}$ by monitoring changes in potential energy is not practical since we need to also consider the enthalpy of solvent-solvent interaction and to track all the association and dissociation events of different clusters. Thus, we propose and assess two methods to obtain a

functional form of the PEACH-derived $\Delta G_{micellization}$ to calculate enthalpy change of formation of every cluster composition ($\Delta H_{i,j}$), and then convert the cluster statistics to construct enthalpograms. By independently tuning the contribution of different cluster sizes and comparing the resulting enthalpograms to the ITC data, we are able to provide structural details in the dominant enthalpic contribution during the pre-cmc stage.

3.2 Methods

3.2.1 General Simulation Setup

The simulation setup is the same as present in Chapter 2.2.1 except for the temperature and the kind of surfactants investigated. In this study, we choose sodium decanoate (NaDA) as our model surfactants as the sampling of their aggregation behavior is found to be better than the other ones. Table S1 presents the number of molecules, temperature, box volume, trajectory length and the pre-equilibration length of each system investigated.

3.2.2 Generating Enthalpograms

3.2.2.1 Converting Cluster Statistics to Enthalpograms

Since the ITC curve expresses the experimental heats of dilution as a function of surfactant concentrations, we simulate the experimental setup where a concentrated surfactant (stock) solution is added to distilled water. We first calculate the total enthalpy per unit volume $H(c_{tot})$ at a particular surfactant solution concentration c_{tot} as

$$H(c_{tot}) = \sum_{i,j} \Delta H_{i,j} * c_{i,j} \quad (8)$$

where $\Delta H_{i,j}$ is the molar enthalpy change of *i* surfactants and j ions aggregating into a cluster, and $c_{i,j}$ is the equilibrium concentration of that cluster composition calculated by the method presented in Chapter 2.2.4 and Eq. 5. To account for the effect of adding stock solution first, the final enthalpy change is calculated as

$$\Delta H = \frac{q}{n_{stock}} = \frac{H(c)*(V_0+\Delta V)-H_0(c_0)*V_0-H_{c_{stock}}*\Delta V}{n_{stock\ in\ \Delta V}} \quad (9)$$

where $H_0(c_0)$ and $H(c)$ are the enthalpy for the system before and after adding one injection of the stock solution. n_{stock} (mol) is the total amount of surfactants in the injected volume.

3.2.2.2 Calculating $\Delta H_{i,j}$

The new challenge is then to obtain values for $\Delta H_{i,j}$ from simulation results. Gibbs Helmholtz Equations [44] at constant pressure (1 bar) is used to derive the enthalpy change of micellization from the free energy change of micellization at different temperatures. The general formulation of the equation is shown in Eq. 10.

$$\left(\frac{\partial(\Delta G/T)}{\partial T}\right)_p = \frac{-\Delta H}{T^2} \quad (10)$$

To get an accurate estimation of the ΔH of every cluster composition, we need a good model to describe $\Delta G_{i,j}$ across different temperatures that doesn't overfit the simulation data, which are subject to noise but captures the important physical trend. Here we try two methods to accomplish that.

Method 1 uses the functional form of the Maibaum's Model to describe the isobaric heat capacity of micellization. As the heat capacity should be temperature-independent and thus A'', B'', C'', and D'' are constants.

$$\Delta c_{\mathrm{p,model}}(\mathrm{i},j) = A''(\mathrm{i}-1) + B''\left(i^{\frac{2}{3}}-1\right) + C''(i^2-1) - D''\mathrm{j} \quad (11)$$

Using the definition of isobaric heat capacity, we have

$$\int_{T_r}^{T} \partial \Delta H_{\mathrm{model}}(i,j,T) = \int_{T_r}^{T} \Delta c_{\mathrm{p,model}}(i,j)\partial T \quad (12)$$

We choose 300K as our reference temperature (T_r), and we get $\Delta H_{\mathrm{model}}(i,j,T)$

$$\Delta H_{\mathrm{model}}(i,j,T) = \Delta H_{\mathrm{model}}(i,j,T_r) + A''(T-T_r)(i-1) + B''(T-T_r)\left(i^{\frac{2}{3}}-1\right) +$$

$$C''(T-T_r)(i^2-1) - D''(T-T_r)j \quad (13)$$

Running the Gibbs-Helmholtz Equation on Eq. 10, we get $\Delta G_{\mathrm{model}}(i,j,T)$

$$\int_{T_r}^{T} \partial(\Delta G_{\mathrm{model}}(i,j,T)/T) = \int_{T_r}^{T} -\frac{\Delta H_{\mathrm{model}}(i,j,T)}{T^2}\partial T$$

$$\Delta G_{model}(i,j,T) = \frac{T}{T_r}\Delta G_{model}(i,j,T_r) + \Delta H_{model}(i,j,T_r)\left(1-\frac{T}{T_r}\right) - A''\left(T\,ln\frac{T}{T_r} + T_r - T\right)(i-1) - B''\left(T\,ln\frac{T}{T_r} + T_r - T\right)\left(i^{\frac{2}{3}} - 1\right) - C''\left(T\,ln\frac{T}{T_r} + T_r - T\right)(i^2-1) + D''\left(T\,ln\frac{T}{T_r} + T_r - T\right)j \quad (14)$$

The functional form of Maibaum's Model is also used to describe the $\Delta G_{\text{model}}(i,j,T_r)$ and $\Delta H_{\text{model}}(i,j,T_r)$

$$\Delta G_{\text{model}}(i,j,T_r) = A(T_r)(i-1) + B(T_r)\left(i^{\frac{2}{3}} - 1\right) + C(T_r)(i^2-1) - D(T_r)j - k_B T \ln\left(\frac{i!}{(i-j)!\,j!}\right)$$

$$\Delta H_{\text{model}}(i,j,T_r) = A'(T_r)(i-1) + B'(T_r)\left(i^{\frac{2}{3}} - 1\right) + C'(T_r)(i^2-1) - D'(T_r)j \quad (15)$$

The second method is a brute-force strategy, where we directly apply Eq. 10 to every cluster composition, hoping that the noises would average out during regression. We assume that the enthalpy change of micellization is independent of temperature and integrate Eq. 10 with respect to T for each cluster size. To incorporate heat capacity effects independently for each cluster size and composition would require fitting a parabola through a small number of points, whose scatter would cause overfitting.

$$\frac{\Delta G_{i,j}(T_2)}{T_2} - \frac{\Delta G_{i,j}(T_1)}{T_1} = \Delta H_{i,j}\left(\frac{1}{T_2} - \frac{1}{T_1}\right) \quad (16)$$

Since $\Delta G_{i,j} = -lnK_{i,j}$, we have

$$-lnK_{i,j,T_2} + lnK_{i,j,T_1} = \frac{\Delta H_{i,j}}{R}\left(\frac{1}{T_2} - \frac{1}{T_1}\right) \quad (17)$$

According to Eq. 14, we perform linear regression on $-lnK_{i,j}$ over 1/T, and the slope indicates the value of $\frac{\Delta H_{i,j}}{R}$.

3.3 Results and Discussions

3.3.1 Failure of the Maibaum's Model

The assumption to represent the heat capacity, the free energy and the enthalpy changes of micellization using the Maibaum's model leads to large deviations from PEACH-derived values, especially at the formation of small clusters as illustrated by Fig. 15.

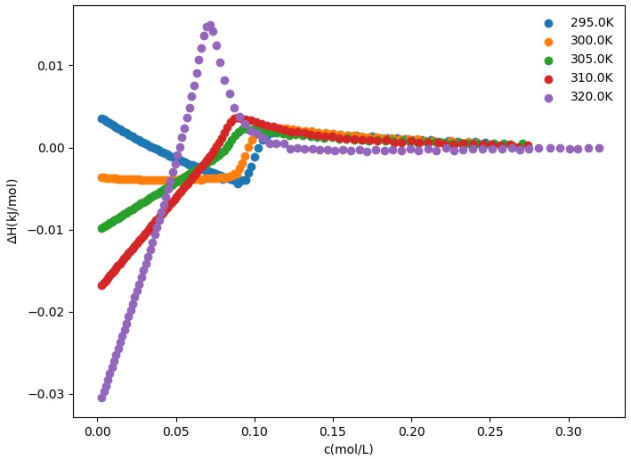


Figure 15. The enthalpogram generated by the Maibaum's Model

This is not surprising given that the Maibaum's model is based on a continuous treatment of the cluster geometries and interactions as mentioned in Chapter 2.3.3. This model yields strongly exothermic cluster formation enthalpies for i < 6 (especially at 295K). To better sample the small clusters, separate simulations and PEACH fitting are performed at the concentration of 10 NaDA and 3000 water molecules for all temperatures investigated. Hereafter, we will call them the dilute systems, as opposed to the concentrated systems at other concentrations. Fig. 16 plots the $\Delta G_{micellization}$/T over 1/T of dimer formation in dilute systems, where the slope indicates the enthalpy of micellization according to Eq. 10. The temperature dependence from the diluted simulations indicate that the formation of small clusters is weakly exothermic to endothermic, whereas the Maibaum's model predicts that dimer formation has negative enthalpy at even 295K. Since dimer formation is dominant in the pre-cmc stage (further validated in Chapter 3.3.3), the inaccurate representation of the

Maibaum's model at small clusters will significantly affect the derivation of the enthalpogram.

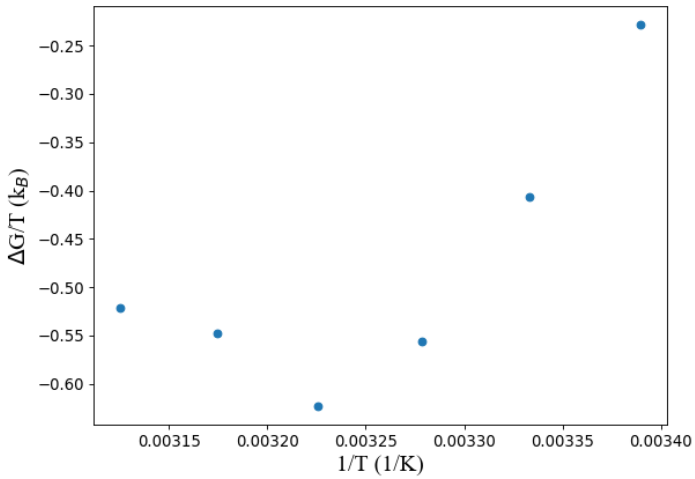


Figure 16. ΔG/T over 1/T (T=295K, 300K, 305K, 310K, 315K, 320K) of dimer formation. The slope between two points indicates the enthalpy of association.

3.3.2 Exploration of Counterion Binding Thermodynamics

Next, we investigate the assumption in the Maibaum's model that the binding energy per counterions to available headgroups (ε) is independent of the micelle size or the presence of other counterions. Fig. 17 plots the cluster free energy without the combinatorial factor ($\Delta G_{micellization}$-constant) over the number of counterions (j) at five clusters sizes, where the slope indicates the value of epsilon in Eq. 6. Fig. 17 shows that the slope of every line stays constant at a micelle size, proving that ε is independent of the presence of other counterions.

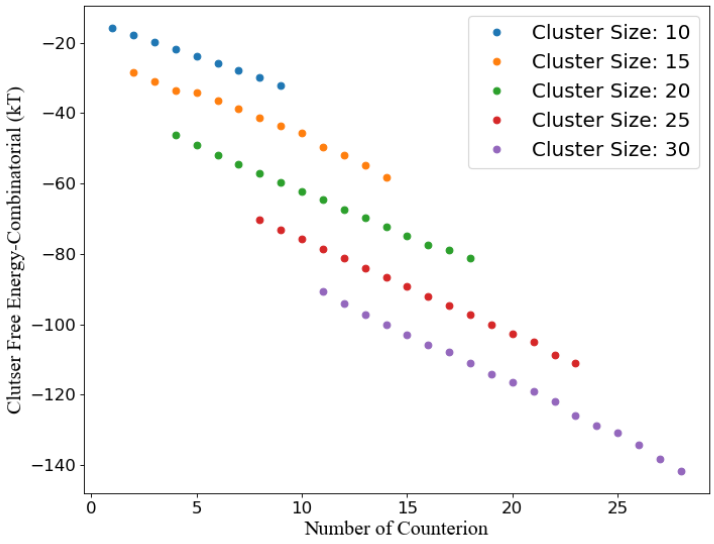


Figure 17. Cluster free energy without the combinatorial factor (ΔG-constant) at 295K over the number of counterions (j) at i = 10, 15, 20, 25, 30

The independence of ε on other counterions confirms the existence of short-range repulsion between counterions, where no two counterions can occupy the same available site on the micelle. Fig. 18 plots the value of ε over the cluster sizes across temperatures.

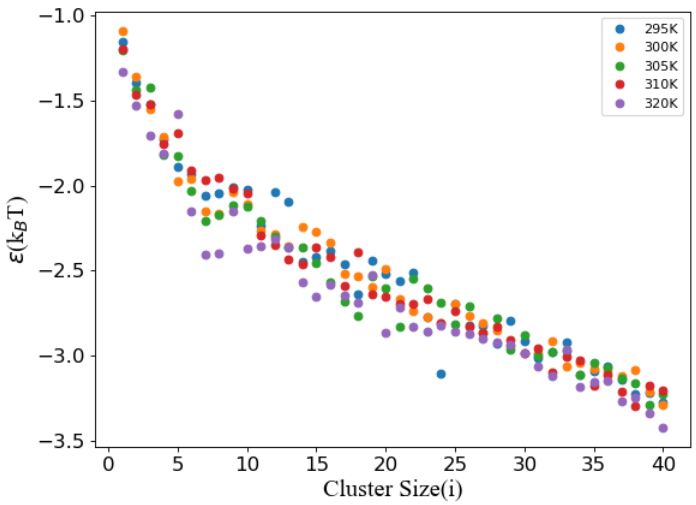


Figure 18. ε over cluster sizes at T = 295, 300, 305, 310, 320K

Fig. 18 shows that the favorable ε gets stronger as cluster size increases at all investigated temperatures (with no obvious dependence on temperature). The effect is particularly significant during the pre-cmc stage. There are several factors ignored in treating ε as independent of micelle size [45]. First is the steric interaction between the surfactant heads

and the adsorbed counterions on the micelle surface. The presence of bounded counterions at the micelle surface leads to an increase in the steric repulsions, so an increase in micelle size will reduce this unfavorable hindrance. Secondly, the model ignores the main driving force promoting counterion binding: the reduction in the electrostatic repulsions between acetate heads. The electrostatic contribution of counterions depends on the micelle size and structure. Thirdly, due to the cooperative effects of counterions at micelle surface that are absent in the bulk solution, the octanoate heads and associated counterions are further dehydrated. Since the hydration entropy around a charged species is negative due to the strong water organization around a charged species, the further dehydration of the charged micelles leads to an increase in the water entropy [17]. In the next chapter, we investigate the applicability of Method 2 to model the enthalpogram, particularly focusing on the physical event that happened in the pre-cmc stage.

3.3.3 Pre-cmc Cluster Enthalpy

Experiments have shown that $\Delta H_{micellization}$ of large clusters has negative heat capacity, meaning that $\Delta H_{micellization}$ decreases as temperature increases. Qualitatively, our simulations fail to reproduce the trend of heat capacity as simulated $\Delta H_{micellization}$ of many large clusters is observed to be constant by comparing their $\Delta G_{micellization}$ across temperatures. Thus, we focus on the enthalpogram at our reference temperature (T_r=300K).

Fig. 19 shows the modeled enthalpogram (left) using Method 2 in the concentrated systems, and the ITC data (right).

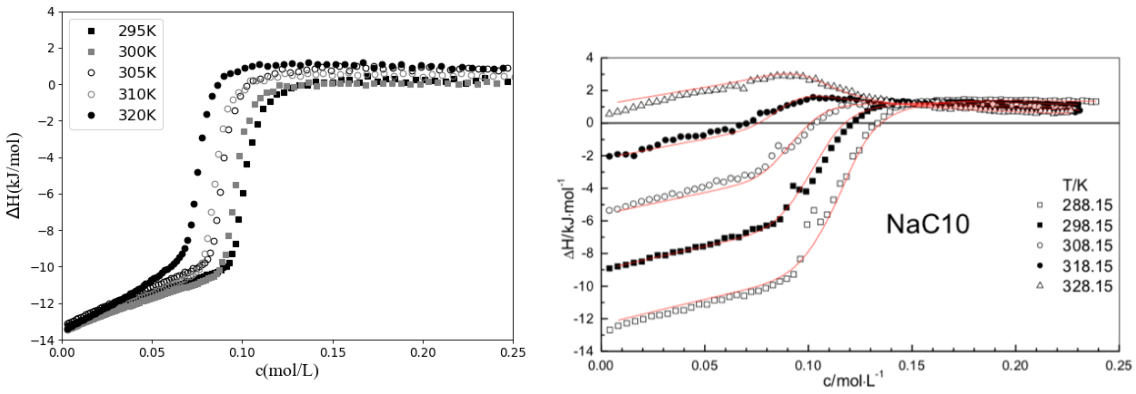


Figure 19. The modeled enthalpogram (left) and the experimental ITC data (right).

Qualitatively, the simulation results show a larger jump in the heat of dilution at the onset of micellization where the slopes are steeper and the cmc predicted at 300K is lower. The differences in initial ΔH in the ITC data are due to the differences in ΔH of stock solutions across temperatures, so they follow from the heat capacity effect that we are neglecting. Quantitatively, experimental fitting to the ITC curve predicts that $\Delta H_{per\ monomer\ ITC}$ = 10.9±0.3 kJ/mol at 300K [7], where $\Delta H_{per\ monomer}$ indicates the enthalpy change of associating a monomer with the micelle.

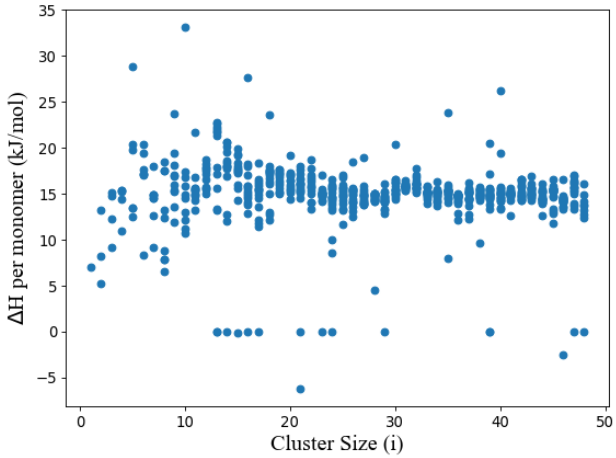
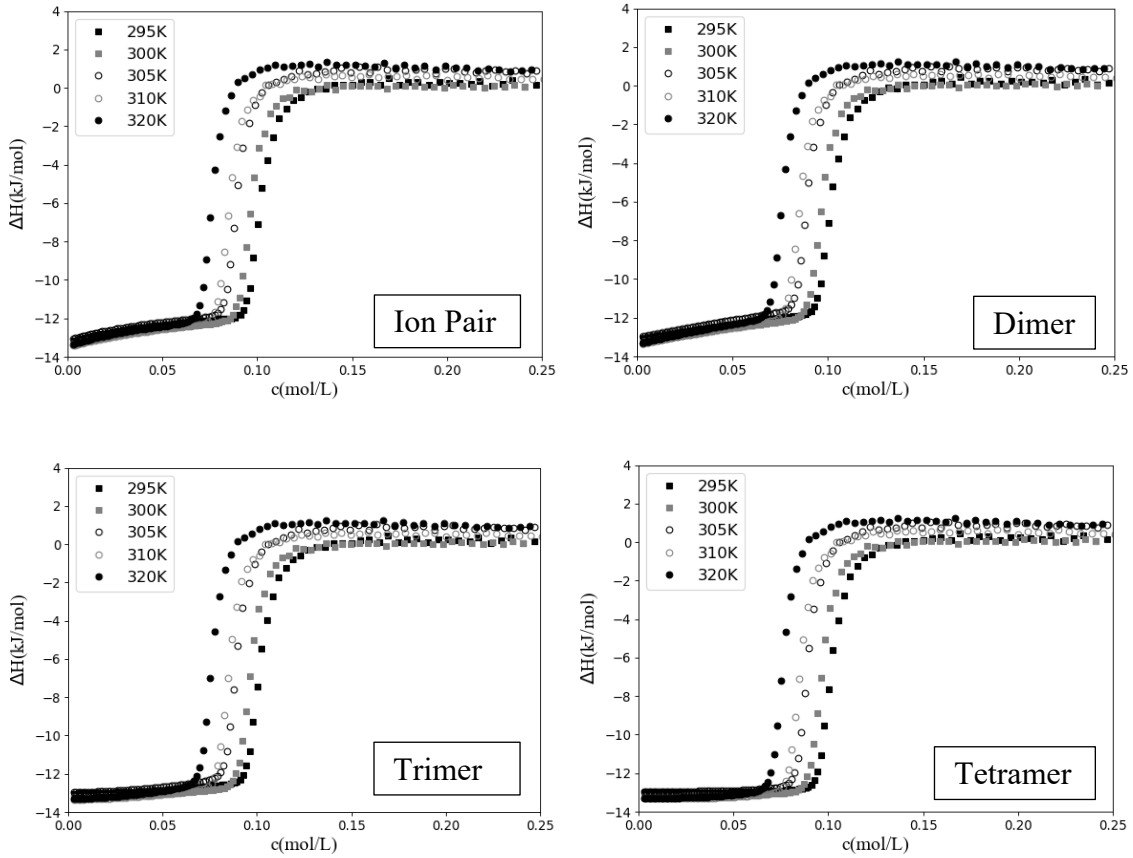


Figure 20. ΔH per monomer in the cluster derived by Method 2 over the cluster sizes

Fig. 20 plots the simulated ΔH per monomer in the cluster derived by Method 2 over the cluster sizes, where $\Delta H_{per\ monomer\ sim}$ = 15.0 kJ/mol, in fair agreement with the ITC fit. The overestimation is due to our assumption of constant $\Delta H_{micellization}$ for both small and large

clusters across temperatures, while experimental $\Delta H_{micellization}$ decreases as the temperature increases for large clusters.

The ITC enthalpogram shows that ΔH is temperature-independent ($\Delta C_p \approx 0$) during the pre-cmc stage as the initial slopes of the enthalpogram are constant across temperatures. Since ΔC_p decreases as the loss of water accessible to the non-polar headgroup increases [43], the experimentalists argued that in the pre-cmc stage, the nonpolar alkyl chains in the cluster core are in strong contact with water, leading to a near zero ΔC_p [7]. They hypothesized that the formation of ion pairs is dominant in the pre-cmc stage. To verify their hypothesis, we set $H_{i,j}$ = 0 for all i<6 except for the investigated cluster size in concentrated systems and compare the initial slope to that in the ITC data and the results are shown in Fig. 21.



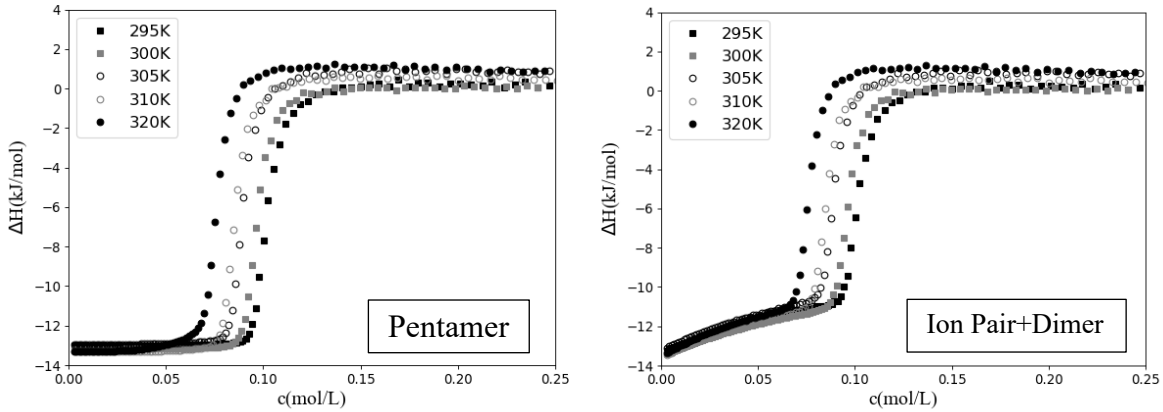


Figure 21. Enthalpograms of different cluster contributions in the pre-cmc stage in concentrated systems.

By considering the enthalpy contribution from both ion pairs and dimers, the initial slopes of the enthalpograms correspond well with the experimental ITC data. On the other hand, the contributions from trimers, tetramers and pentamers are negligible. The amount of enthalpic contributions from ion pairs and dimers separately are also verified in the dilute systems (Fig. 22) where the slopes of their curves are similar to those calculated from the concentrated systems.

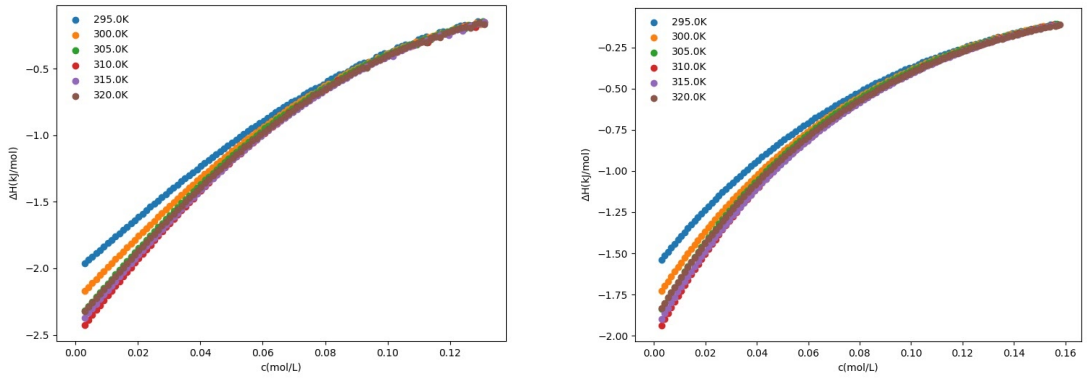


Figure 22. Enthalpograms of different cluster contributions in the pre-cmc stage in diluted systems. Left: contribution of ion pairs; Right: contribution of dimers. The initial slopes are similar to those in the concentrated systems and ITC curves.

According to Eq. 8, the calculation of H_{total} depends on $c_{i,j}$ and $H_{i,j}$. Fig. 23 shows the PEACH-derived standard ΔG of ion pairs and dimers across temperatures in dilute systems.

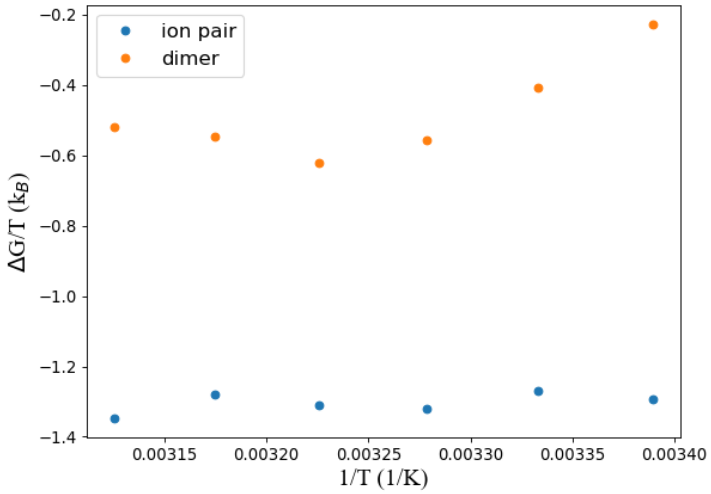


Figure 23. The PEACH-derived standard ΔG of ion pairs and dimers across temperatures in dilute systems.

Fig. 23 shows that $\Delta G_{1,1}$ varies little in the temperature range investigated, indicating a very small enthalpy change of ion pair formation, while there's a clear temperature dependence of $\Delta G_{2,0}$. We cannot tell whether ΔH of ion pair formation is positive or negative. But from our analysis, a combination of cluster size distribution and their corresponding $\Delta H_{i,j}$ is enough to give a pre-cmc enthalpy change that's comparable to experiments, suggesting that dimer formation cannot be ignored. However, the simulation still fails to reproduce the heat capacity trend of small cluster formation. The ITC curve shows zero heat capacity at pre-cmc stage, while the simulated data shows a negative heat capacity.

3.4 Conclusion

In this work we present a model to predict the enthalpy change of micellization of long-chain carboxylates in the presence of counterions. We evaluate the applicability and assumptions of the two methods to model the free energy change of micellization. The continuous micelle geometry assumed in the Maibaum's model leads to large deviations from PEACH-derived ΔG for small clusters. The binding energy per counterion to available headgroups is independent of other counterions but gets stronger as micelle size increases. By

assuming that the enthalpy change of micellization is independent of temperature, we successfully generate the enthalpogram at the reference temperature that is consistent with the experimental ITC data. $\Delta H_{per\ monomer}$ that we directly calculated from simulation results, 15.0 kJ/mol, is in fair agreement with the indirect two-step fit that experimentalists performed on the ITC data, 10.9±0.3 kJ/mol at 300K. Through comparing the initial slope of predicted and experimental enthalpograms, we conclude that at pre-cmc stage, an important contribution to the total enthalpy change of micellization comes from dimer formation, while the contribution from ion pair is hard to estimate. However, the calculated enthalpogram predicts a negative ΔC_p for formation of dimers and other small clusters, which is not supported by experiments.

4. Future Directions

This project presents a strategy to perform detailed analysis of the thermodynamics of anionic surfactants self-assembly using small-N atomistic simulations with a converged cluster size distribution. The PEACH method and the enthalpogram model can be further applied to other 2-component surfactant systems and provide molecular-level details of surfactant aggregation with different kinds of counterions. Future work will be focused on developing a temperature-dependent enthalpogram model that's not subject to noise in simulations and be able to capture the experimental trend of heat capacity of micellization. Other phenomenological models to fit the free energy change or models that don't rely on known functional forms, e.g., symbolic regression [46] can be explored to fulfill the goal. By optimizing the simulation setup, fitting procedures, cut-off distances and cluster definitions, we hope that we can perform a more holistic thermodynamic analysis of bicomponent surfactant systems in a computationally efficient way that is able to complement the experimental study.

5. References

1. Whitesides, G.; Mathias, J.; Seto, C. Molecular Self-Assembly and Nanochemistry: A Chemical Strategy for the Synthesis of Nanostructures. *Science* **1991**, *254*, 1312–1319.
2. Micelle - Reverse Micelles: Supercells: Applications of Micelle https://byjus.com/chemistry/micelle/ (accessed Feb 24, 2021).
3. De Jesus, A. J.; Yin, H. Supramolecular Membrane Chemistry. *Comprehensive Supramolecular Chemistry II* **2017**, 311–328.
4. Luczak, J.; Jungnickel, C.; Joskowska, M.; Thöming, J.; Hupka, J. Thermodynamics of micellization of imidazolium ionic liquids in aqueous solutions. *Journal of Colloid Interface Science* **2009**, *336*, 111−116.
5. Chatterjee, A.; Moulik, S. P.; Sanyal, S. K.; Mishra, B. K.; Puri, P. M. Critical Evaluation of Micellization Behavior of Nonionic Surfactant MEGA 10 in Comparison with Ionic Surfactant Tetradecyltriphenylphosphonium Bromide Studied by Microcalorimetric Method in Aqueous Medium. *The Journal of Physical Chemistry B* **2006**, *110*, 9815−9821.
6. Fameau, A.-L.; Arnould, A.; Saint-Jalmes, A. Responsive Self-Assemblies Based on Fatty Acids. *Current Opinion in Colloid & Interface Science* **2014**, *19*, 471–479.
7. Medoš, Ž.; Bešter-Rogač, M. Two-Step Micellization Model: The Case of Long-Chain Carboxylates in Water. *Langmuir* **2017**, *33*, 7722–7731.
8. Srivastava, V. K.; Yadav, R. Isothermal Titration Calorimetry. *Data Processing Handbook for Complex Biological Data Sources* **2019**, 125–137.
9. Zhang, X.; Patel, L. A.; Beckwith, O.; Schneider, R.; Weeden, C. J.; Kindt, J. T. Extracting Aggregation Free Energies of Mixed Clusters from Simulations of

Small Systems: Application to Ionic Surfactant Micelles. *Journal of Chemical Theory and Computation* **2017**, *13*, 5195–5206.

10. Roy, K.; Kar, S.; Das, R. N. Computational Chemistry. *Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment* **2015**, 151–189.
11. Frenkel, D.; Smit, B. *Understanding Molecular Simulation: from Algorithms to Applications*; Academic Press: San Diego, CA, 2002.
12. Martin, M. G.; Siepmann, J. I. Transferable Potentials for Phase Equilibria. 1. United-Atom Description of n-Alkanes. *The Journal of Physical Chemistry B* **1998**, *102*, 2569-2577.
13. White, P.; Benson, G. C. The Temperature Variation of the Heat of Micellization of Potassium Octanoate in Aqueous Solutions. *Transactions of the Faraday Society* **1959**, *55*, 1025.
14. Campbell, A. N.; Kartzmark, E. M.; Lakshminarayanan, G. R. Conductances of Aqueous Solutions of Sodium Octanoate at 25° and 35° and the Limiting Conductance of the Octanoate Ion. *Canadian Journal of Chemistry* **1962**, *40*, 839–844.
15. Campbell, A. N.; Lakshminarayanan, G. R. Conductances and Surface Tensions of Aqueous Solutions of Sodium Decanoate, Sodium Laurate, and Sodium Myristate, at 25° and 35°. *Canadian Journal of Chemistry* **1965**, *43*, 1729–1737.
16. Pereira, R. F.; Valente, A. J.; Fernandes, M.; Burrows, H. D. What Drives the Precipitation of Long-Chain Calcium Carboxylates (Soaps) in Aqueous Solution? *Physical Chemistry Chemical Physics* **2012**, *14*, 7517.

17. Bernardino, K.; de Moura, A. F. Aggregation Thermodynamics of Sodium Octanoate Micelles Studied by Means of Molecular Dynamics Simulations. *The Journal of Physical Chemistry B* **2013**, *117*, 7324–7334.

18. Srinivasan, V.; Blankschtein, D. Prediction of Conformational Characteristics and Micellar Solution Properties of Fluorocarbon Surfactants. *Langmuir* **2005**, *21*, 1647–1660.

19. Li, P.; Jansson, M.; Bahadur, P.; Stilbs, P. NMR Study of Organic Counterion Binding and Micellization of Decylammonium Dicarboxylate Surfactants. *The Journal of Physical Chemistry* **1989**, *93*, 6458–6463.

20. Benrraou, M.; Bales, B.L.; Raoul Z. Effect of the Nature of the Counterion on the Properties of Anionic Surfactants. 1. Cmc, Ionization Degree at the cmc and Aggregation Number of Micelles of Sodium, Cesium, Tetramethylammonium, Tetraethylammonium, Tetrapropylammonium, and Tetrabutylammonium Dodecyl Sulfates. *Journal of Physical Chemistry B* **2003**, *107,* 13432-13440.

21. Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; *et al.* GROMACS 4.5: A High-Throughput and Highly Parallel Open-Source Molecular Simulation Toolkit. *Bioinformatics* **2013**, *29*, 845–854.

22. Ashbaugh, H. S.; Liu, L.; Surampudi, L. N. Optimization of Linear and Branched Alkane Interactions with Water to Simulate Hydrophobic Hydration. *The Journal of Chemical Physics* **2011**, *135*, 054510.

23. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *The Journal of Chemical Physics* **1983**, *79* (2), 926−935.

24. Murzyn, K.; Bratek, M.; Pasenkiewicz-Gierula, M. Refined OPLS All-Atom Force Field Parameters Forn-Pentadecane, Methyl Acetate, and Dimethyl Phosphate. *The Journal of Physical Chemistry B* **2013**, *117*, 16388–16396.

25. Hess, B.; van der Vegt, N. F. Cation Specific Binding with Protein Surface Charges. *Proceedings of the National Academy of Sciences* **2009**, *106*, 13296–13300.

26. Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling Through Velocity Rescaling. *The Journal of Chemical Physics* **2007**, *126*(1), 014101.

27. Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular-Dynamics with Coupling to an External Bath. *The Journal of Chemical Physics* **1984**, *81*(8), 3684−3690.

28. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *The Journal of Chemical Physics* **1995**, *103*, 8577−8593.

29. Burov, S. V.; Shchekin, A. K. Aggregation Work at Polydisperse Micellization: Ideal Solution and ″Dressed Micelle″ Models Comparing to Molecular Dynamics Simulation. *The Journal of Chemical Physics* **2010**, *133*, 244109.

30. Humphrey, W.; Dalke, A.; Schulten, K. VMD - Visual Molecular Dynamics. *The Journal of Molecular Graphics* **1996**, *14*, 33−38.

31. Winstein, S.; Clippinger, E.; Fainberg, A. H.; Heck, R.; Robinson, G. C. Salt Effects and Ion Pairs in Solvolysis and Related Reactions. III. Common Ion Rate Depression and Exchange of Anions during Acetolysis. *Journal of the American Chemical Society* **1956**, *78*, 328–335.

32. Maibaum, L.; Dinner, A. R.; Chandler, D. Micelle Formation and the Hydrophobic Effect. *The Journal of Physical Chemistry B* **2004**, *108*, 6778–6781.

33. Uejio, J. S.; Schwartz, C. P.; Duffin, A. M.; Drisdell, W. S.; Cohen, R. C.; Saykally, R. J. Characterization of Selective Binding of Alkali Cations with Carboxylate by x-Ray Absorption Spectroscopy of Liquid Microjets. *Proceedings of the National Academy of Sciences* **2008**, *105*, 6809–6812.
34. Aziz, E. F.; Ottosson, N.; Eisebitt, S.; Eberhardt, W.; Jagoda-Cwiklik, B.; Vácha Robert; Jungwirth, P.; Winter, B. Cation-Specific Interactions with Carboxylate in Amino Acid and Acetate Aqueous Solutions: X-Ray Absorption and Ab Initio Calculations. *The Journal of Physical Chemistry B* **2008**, *112*, 12567–12570.
35. Zhang, X.; Arce Nunez, J. G.; Kindt, J. T. Derivation of Micelle Size-Dependent Free Energies of Aggregation for Octyl Phosphocholine from Molecular Dynamics Simulation. *Fluid Phase Equilibria* **2019**, *485*, 83–93.
36. Collins, K. D. Charge Density-Dependent Strength of Hydration and Biological Structure. *Biophysical Journal* **1997**, *72*, 65–76.
37. Wang X, Yu L, Jiao J, Zhang H, Wang R, Chen H. Aggregation Behavior of COOH Functionalized Imidazolium-based Surface-active Ionic Liquids in Aqueous Solution. *Journal of Molecular Liquids* **2012**, *173*, 103–107.
38. Adriana P. G.; Paulo F.A. Costa; Faruk N.; Frank Q. Micellization and Adsorption of Zwitterionic Surfactants at the Air/water Interface. *Current Opinion in Colloid & Interface Science* **2017**, *32*, 48-56.
39. del Rio, J. M.; Jones, M. N. Thermodynamics of the Hydrophobic Effect. J *The Journal of Physical Chemistry B* **2001**, *105*, 1200−1211.
40. Huang, D. M.; Chandler, D. The Hydrophobic Effect and the Influence of Solute−Solvent Attractions. *The Journal of Physical Chemistry B* **2002**, *106*, 2047−2053.

41. Medoš, Ž.; Bešter-Rogač, M. Thermodynamics of the Micellization Process of Carboxylates: A Conductivity Study. *The Journal of Chemical Thermodynamics* **2015**, *83*, 117–122.

42. Kresheck, G. C. The Temperature Dependence of the Heat Capacity Change for Micellization of Nonionic Surfactants. *Journal of Colloid Interface Science* **2006**, *298*, 432−440.

43. Ropers, M. H.; Czichocki, G.; Brezesinski, G. Counterion Effect on the Thermodynamics of Micellization of Alkyl Sulfates. *The Journal of Physical Chemistry B* **2003**, *107*, 5281–5288.

44. Atkins, P. W.; Paula, J. D. *Physical Chemistry*; Oxford University Press: Oxford, 2006.

45. Srinivasan, V.; Blankschtein, D. Effect of Counterion Binding on Micellar Solution Behavior: 2. Prediction of Micellar Solution Properties of Ionic Surfactant−Electrolyte Systems. *Langmuir* **2003**, *19*, 9946–9961.

46. Udrescu, S. M.; Tegmark, M. AI Feynman: A Physics-inspired Method for Symbolic Regression. *Science Advances* **2020**, *6* (16).

6. Supplementary Materials

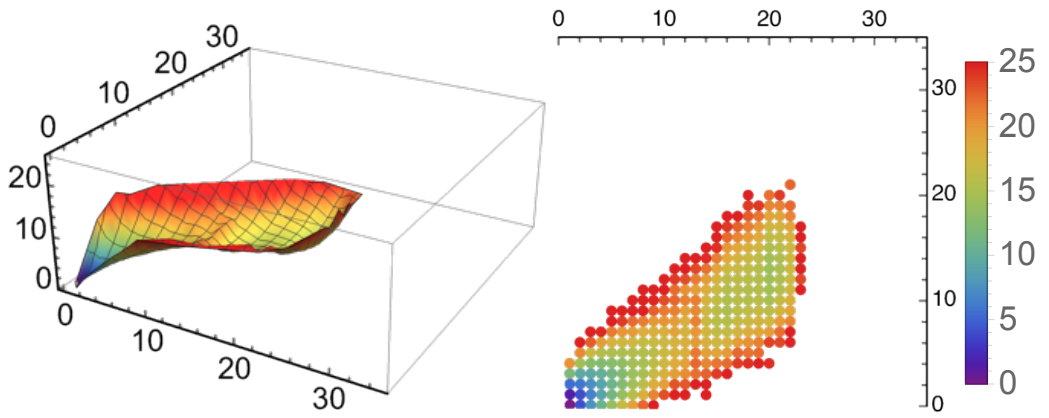


Figure S1. The 3-D (left) and 2-D (right) representation of the free energy surface of KOA at concentration i = 0.170 mol/dm^3 and j = 0.286 mol/dm^3: x-axis: No. of surfactant; y-axis: No. of counterions; z-axis (color bar): Free energy of formation (k_BT)

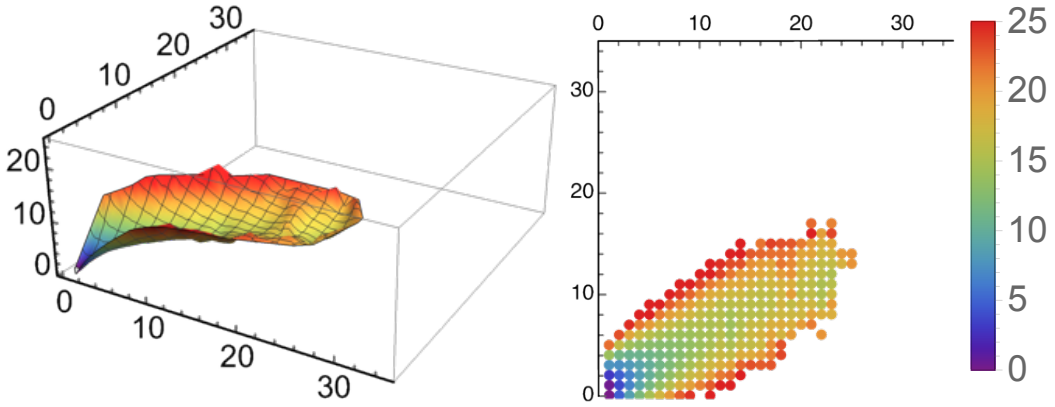


Figure S2. The 3-D (left) and 2-D (right) representation of the free energy surface of TMAOA at concentration i = 0.215 mol/dm^3 and j = 0.317 mol/dm^3: x-axis: No. of surfactant; y-axis: No. of counterions; z-axis (color bar): Free energy of formation (k_BT)

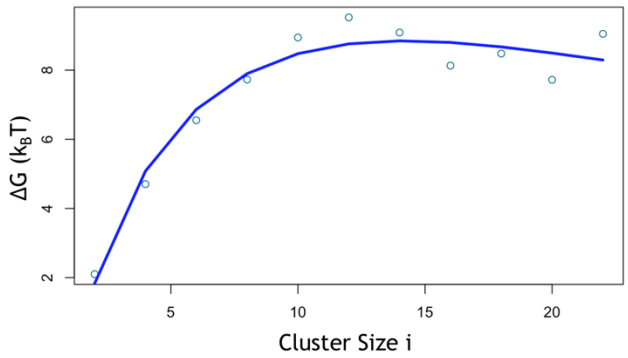


Figure S3. The 1-D fitting of KOA. Dots: PEACH-derived values; lines: fitted values

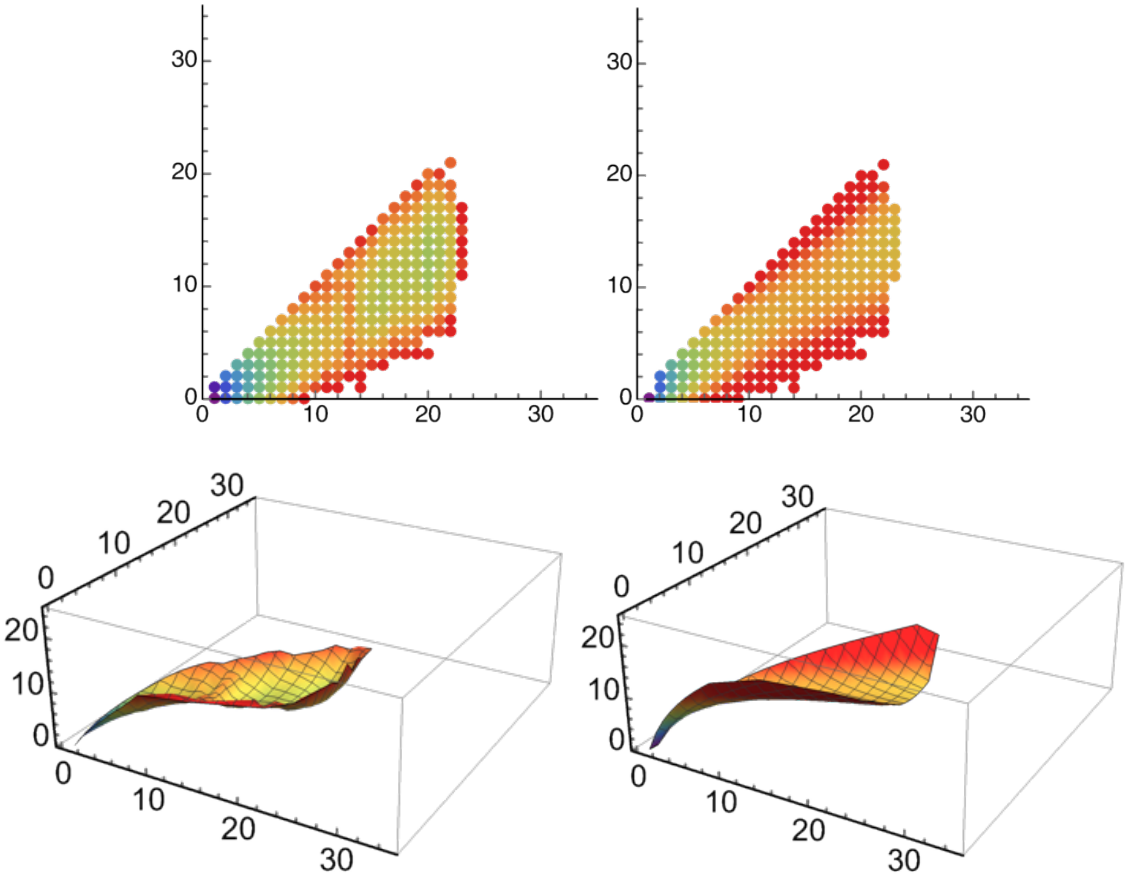


Figure S4. The fitted free energy surfaces of KOA. Left: 2-D (above) and 3-D (below) representations of PEACH-derived surfaces. Right: 2-D (above) and 3-D (below) representations of fitted surfaces

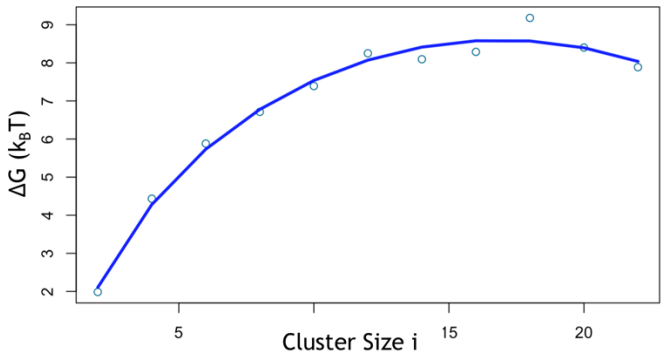


Figure S5. The 1-D fitting of TMAOA. Dots: PEACH-derived values; lines: fitted values

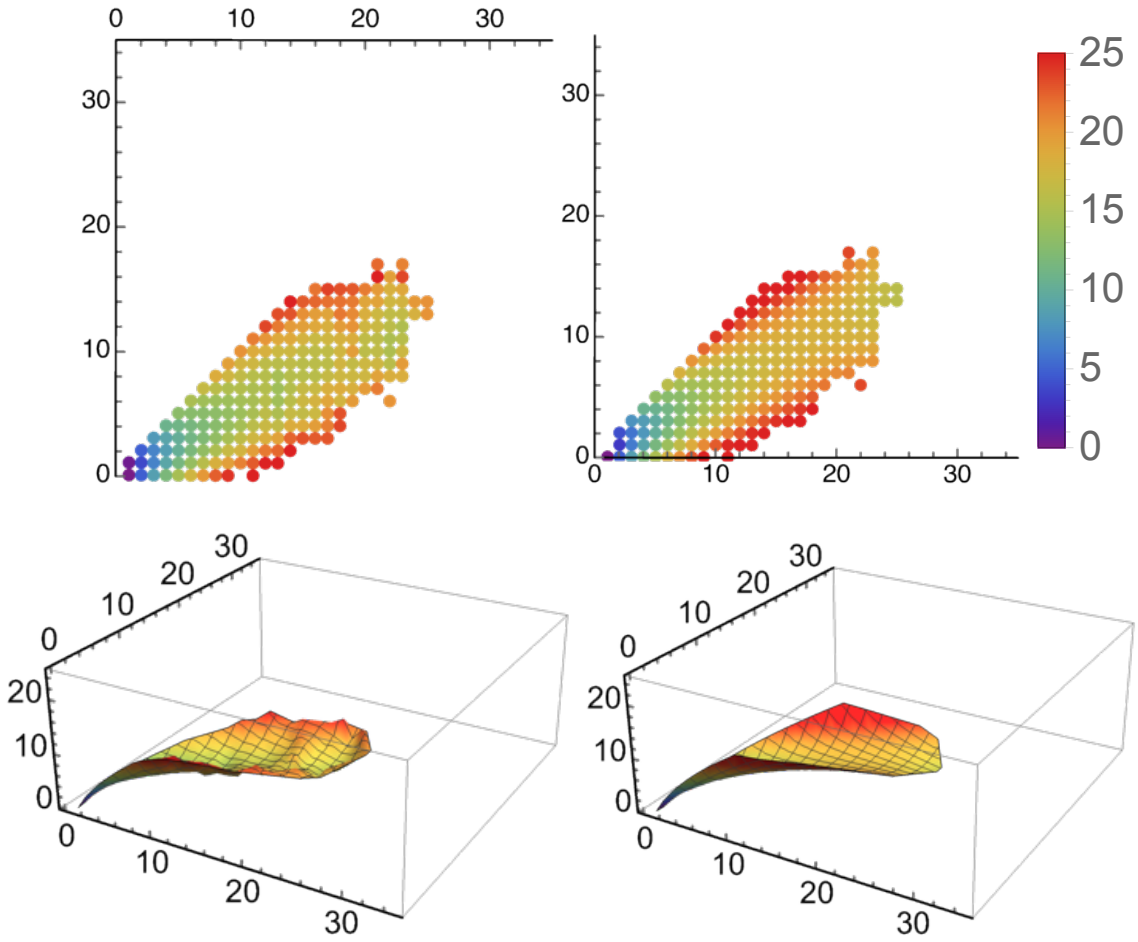


Figure S6. The fitted free energy surfaces of TMAOA. Left: 2-D (above) and 3-D (below) representations of PEACH-derived surfaces. Right: 2-D (above) and 3-D (below) representations of fitted surfaces

Table S1. The temperature, number of molecules, trajectory length and the pre-equilibration length of each system.

Temperature	# of NaDA	# of water	Traj. Length	Pre-equilibration length
295	10	3000	300ns	50ns
295	20	4512	2000ns	50ns
295	25	4512	703ns	50ns
295	30	4512	1081ns	50ns
295	35	4512	1200ns	150ns
295	40	4512	1980ns	150ns
295	45	8875	900ns	100ns
295	50	8875	1140ns	100ns

Temperature	# of NaDA	# of water	Traj. Length	Pre-equilibration length
305	10	3000	300ns	50ns
305	20	4512	2000ns	50ns
305	25	4512	789ns	50ns
305	30	4512	803ns	50ns
305	35	4512	801ns	50ns
305	40	4512	2000ns	150ns
305	45	8875	900ns	500ns
305	50	8875	900ns	200ns

Temperature	# of NaDA	# of water	Traj. Length	Pre-equilibration length
310	10	3000	300ns	50ns
310	20	4512	900ns	100ns
310	25	4512	900ns	100ns
310	30	4512	900ns	100ns
310	35	4512	900ns	150ns
310	40	4512	900ns	150ns
310	45	8875	1000ns	50ns
310	50	8875	900ns	100ns

Temperature	# of NaDA	# of water	Traj. Length	Pre-equilibration length
320	10	3000	219	50
320	20	4512	884	50
320	25	4512	884	50
320	30	4512	880	50
320	35	4512	876	50
320	40	4512	870	50
320	45	8875	612	50
320	50	8875	600	50