

**PCCP****Liquid Worm-like and Proto-Micelles: Water Solubilization in Amphiphile-Oil Solutions**

Journal:	<i>Physical Chemistry Chemical Physics</i>
Manuscript ID	CP-ART-01-2018-000600.R2
Article Type:	Paper
Date Submitted by the Author:	08-Apr-2018
Complete List of Authors:	Qiao, Baofu; Northwestern University, Depart. Materials Science & Engineering Littrell, Kenneth; Oak Ridge National Laboratory, Neutron Scattering Sciences Division Ellis, Ross; Oak Ridge National Laboratory

SCHOLARONE™
Manuscripts

Liquid Worm-like and Proto-Micelles: Water Solubilization in Amphiphile-Oil Solutions

Baofu Qiao,*^a Kenneth C. Littrell^b and Ross J. Ellis*^c

^a Department of Materials Science and Engineering, Northwestern University, Evanston, Illinois 60208, United States

^b Chemical and Engineering Materials Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, 37831, United States

^c Chemical Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, 37831, United States

Corresponding Authors:

*E-mail: qiaobf@gmail.com ;

*E-mail: rossellis1984@gmail.com

Present Address

R.J.E.: Borregaard, P.O. Box 162, N-1701 Sarpsborg, Norway.

Electronic Supplementary Information (ESI) available: experimental details; SAXS and SANS data collection; GIFT method; atomistic simulation details and control simulations. See DOI: 10.1039/x0xx00000x

Noncovalent interactions determine the structure-property relationship of materials. Self-assembly originated from weak noncovalent interactions represent a broad variety of solution-based transformations spanning micellization and crystallization, which, nevertheless, conform to neither colloid, nor solution sciences. Here, we investigate weak self-assembly in water-amphiphile-oil solutions to understand the connection between amphiphile molecular structure and water solubilization in oil. X-ray and neutron scattering, converged with large-scale atomistic molecular dynamics simulations, support that the amphiphiles assemble into liquid worm-like micelles and loose inverted proto-micelles. The inverted proto-micelles are energetically ready to accommodate a higher amount of water. These structures arise from a balance of intermolecular interactions controlled by the amphiphile tail-group structures. Thus, by linking aggregate morphology to molecular structure, this work sheds insights in the molecular design for the control of water solubility and dispersion in oil.

Introduction

Complementary with the covalent interactions which determine the atomic compositions inside molecules (the primary structure), the noncovalent interactions (van der Waals, H-bonds, etc.)¹ drive higher level structures inside molecules and the intermolecular structures and consequently the macroscopic properties of materials.² Therefore the precise understanding and the fine tuning of noncovalent interactions possess a great amount of challenges and opportunities for the bottom-up rational design of functional nanomaterials.³⁻⁷ Weak self-assembly is one particular case where noncovalent interactions predominantly drive the molecular scale structures. Weak self-assembly is considered the low energy limit, typically at the order of $1 k_B T$ per molecule or lower, of a wide range of solution-based transformations including micellization and crystallization, governing important chemical processes in synthesis, biomineralization, as well as separation and purification techniques.⁸ This vast category of intermediate materials sits at the intersection between molecular solution chemistry and classical colloid sciences, encompassing diverse chemical systems from weak amphiphile solutions to inorganic electrolytes. For example, liquid colloidal and polymeric structures have been recently identified in calcium carbonate solutions, illuminating the weak self-assembling processes that prelude biomineralization.⁹ Large micelle-like structures have been investigated in solutions of aliphatic alcohols in water, despite the absence of strong surfactants usually considered a prerequisite for microemulsions.¹⁰ Weak aggregation has also been linked to mesoscale ordering and phase transitions in amphiphile-oil solutions of relevance to technologically important separation processes.¹¹ In all of these solution systems, weak self-assembling processes leads to structures that conform neither to traditional concepts of supramolecular chemistry (i.e. well-defined molecule-molecule interactions forming discrete complexes in solution) nor surface-driven micellization/crystallization processes.

Even though self-assembly of dynamic liquid aggregates through weak noncovalent interactions has been known implicitly for half a century, there is lacking a theoretical understanding for these processes based on first principles that can explain their origin or how to control them.⁸ For instance, oil/water biphasic solutions with weak amphiphiles have received particular attention because of their technological relevance in separation processes.¹¹ In these systems, ions are partitioned preferentially between an aqueous electrolyte and an amphiphile-oil solution. X-ray

and neutron scattering studies have revealed the presence of nanoscale and mesoscale structures in these solutions that are sensitive to ionic strength,^{12, 13} acidity,¹⁴ molecular structure,¹⁵ and temperature.¹⁶ However, although undoubtedly present, the morphology of these aggregates is still poorly understood, and current structural concepts are rooted in classical colloidal science. For example, the majority of current studies describes the phase structure as composed of small aggregates that can be modelled as adhesive solid spheres.¹⁷ This approximation has been useful to understand colligative phenomena at the mesoscale (phase transitions), but it oversimplifies the nanoscale morphology of the supramolecular species.

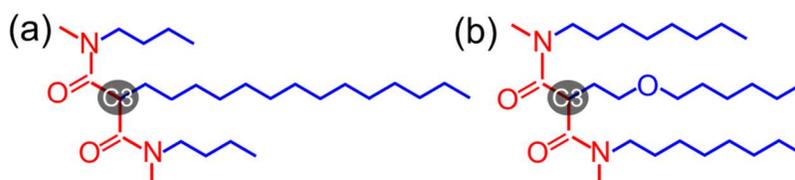
By coupling with experimental techniques, recent studies using atomistic molecular dynamics (MD) simulations have shed new detailed insights into the influence of noncovalent interactions in stabilizing protein structures and activities in oil phase,¹⁸ in self-assembling amphiphile surfactants¹⁹⁻²⁵ in oil phase, as well as in gas phase²⁶. For instance, Eu^{3+} -ligand complexes are surrounded and stabilized by H-bonds,²¹ and the Eu^{3+} -ligand centered aggregates are interacting with each other by means of dipolar interactions, which lead to the mesoscale ordering in the metalloamphiphile phase.²² In the Ce^{3+} and Yb^{3+} systems, the inter-aggregate interactions were revealed to play a crucial role in the metal ion transport from water to oil,²³ along with the enhanced metal-ligand interactions in oil.²⁷

Though the scattering measurements are frequently employed in studying the morphology of aggregates and in the subsequent interpretation of the interactions between aggregates, the precise and unique interpretation of experimental SAXS is still challenging. For instance, the calculation of real space probability distribution $p(r)$ is one of the most frequently employed approaches in this regard. Nevertheless, approximations exist. For instance, the influence of polydispersity is not considered. Therefore, in the last years, with the fast development of atomistic simulation approaches, researchers are resorting to computer simulations to calculate SAXS data. This way, we can compare the calculated SAXS with the experimental SAXS. When reasonable agreement is reached, it can be reasonably argued that the structures in the atomistic simulations are representative of those in the real experimental solutions. This method has been employed by the groups of Prof. Guilbaud, Zemb and co-workers,^{10, 19, 20, 26, 28} and us.^{22, 23, 29}

In the present work, we are aiming to move forward to investigate the impacts of weak noncovalent interactions in the absence of metal ions on how amphiphile molecular structure influences water dispersion in amphiphile-oil solutions. We focus our studies on two diamide molecules (DMDBTDMA: $\text{N,N}'$ -dimethyl $\text{N,N}'$ -dibutyl tetradecylmalonamide and

DMDOHEMA: N,N'-dimethyl-N,N'-dioctylhexylethoxymalonamide, **Scheme 1**),³⁰ which differ only in their tail-group structures. Though it has been reported that trace amount of water is capable of dramatically affecting chemistry and materials,^{31, 32} quantitative studies on water dispersion in oil is very limited. The presence of liquid worm-like and loose 'proto-inverted-micelle' aggregates are demonstrated, with the latter more readily accommodating waters. The amphiphile tail-group structure is pivotal in defining a balance between 'worm-like' and 'proto-micelle-forming' aggregates, controlling the dispersion of water in the hydrocarbon matrix. Our results show a great level of structural complexity in weakly self-assembling water-amphiphile-oil systems, and elucidates new strategies for controlling water solubility and dispersion via molecular design.

Results and Discussion



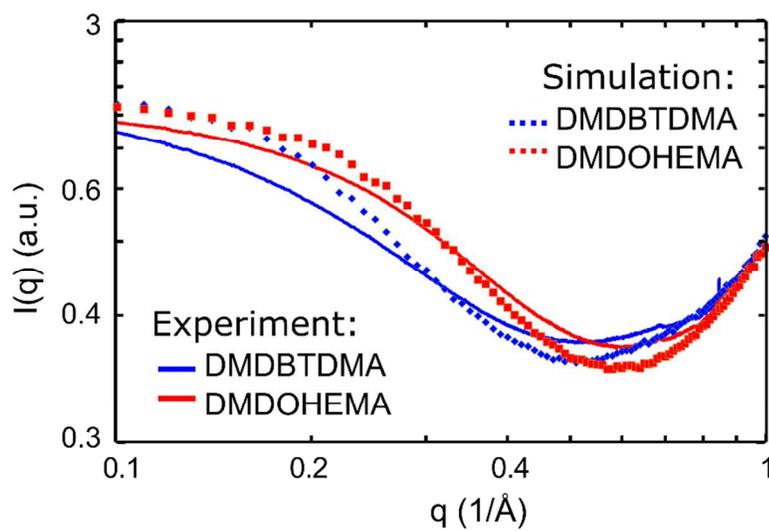
Scheme 1. Diamides investigated: (a) DMDBTDMA and (b) DMDOHEMA. Headgroups and aliphatic tail-groups are colored in red and blue, respectively. The highlighted C3 atom represents the central carbon atom on the headgroup.

In the present work, two diamide amphiphile were employed, DMDBTDMA and DMDOHEMA (**Scheme 1**). They have the same headgroup. But the tail-groups are different in the facts that (a) though the overall numbers of hydrocarbon atoms on the three tail chains are similar, their lengths are quite different, expected to drive different structures; (b) the presence of one extra oxygen atom in the DMDOHEMA provide weak hydrophilicity in its tail regime. 0.5 M diamide amphiphile in *n*-heptane solutions²¹⁻²³ was saturated with water by mixing the amphiphile-oil solution with deionized water, centrifuging the phases and removing the organic layer. The different tail-group structures of the amphiphiles resulted in different water concentrations in the amphiphile-oil solutions, with DMDOHEMA solubilizing 35% more water than DMDBTDMA (0.103 M versus 0.076 M; see also Table S1). This shows that the tail-group structure has a

notable impact on the amount of water that the organic phase can accommodate. Our obtained water concentration in oil is in consistent with an earlier work, where 0.093 M water was extracted to tetrahydrogenated propylene with 0.0492M DMDBTDMA,³³ and approximately 0.1 M water was extracted to *n*-heptane with 0.5 M DMDOHEMA.³⁴ In what follows, we will explain the origin of the difference in the water solubilization on the basis of the structures and energies of the aggregates that emerge from weak noncovalent interactions in the amphiphile-oil phases.

Aggregates evidenced by SAXS, SANS and atomistic MD simulations

To explore structures in the solutions, we collected small-angle X-ray scattering (SAXS) and small-angle neutron scattering (SANS) data, and conducted atomistic MD simulations under the experimental conditions (Table S1). The usage of SAXS and SANS have previously employed in studying the aggregation structure (aggregation number, shape, and size) by Jeng et al.^{35,36} X-rays are scattered by regions of electron density inhomogeneity in the solutions, which is caused by the aggregation of molecular groups that have more electrons (diamide headgroups, water) than the surrounding aliphatic media (diamide tail-groups, heptane). Neutrons are scattered by atomic nuclei depending on the scattering length density of the isotope, and isotopic substitution can then be used to highlight different parts of the molecules that make up the aggregates. In particular, hydrogen and deuterium have very different neutron scattering length densities, so that by using deuterated solvent (d-heptane) the hydrogen-rich aliphatic chains of the diamide

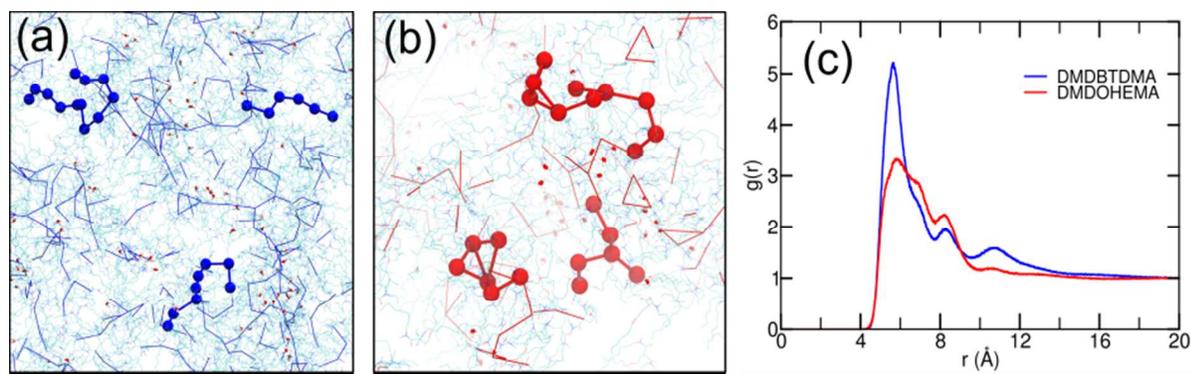


amphiphiles dominate neutron scattering.

Figure 1. SAXS. The normalized experimental SAXS (solid lines) and the MD-SAXS generated from the simulation trajectory (dotted lines). The size of the simulation box defines the low- q limit of the SAXS region ²² ($> 0.1 \text{ \AA}^{-1}$) to 1 \AA^{-1} . See Figure S1 for the corresponding SWAXS in full q range.

Experimental small- and wide-angle X-ray scattering (SWAXS) measurements were collected from both diamide solutions, and the theoretical SWAXS data were generated from the MD simulation trajectories for comparison. The experimental data were normalized using a standard (water), and MD-SWAXS results were normalized against the experimental data using the WAXS solvent (heptane) peak.^{22, 23} The SAXS region ($< 1 \text{ \AA}^{-1}$) holds information pertaining to the structures of supramolecular aggregates. As seen from the simulated and experimental SAXS data in **Figure 1**, the scattering intensity for both diamide solutions increases at low- q values, indicating electron-rich structures on the supramolecular scale. In the present work, the low- q regime refers to the SAXS and SANS with

This is consistent with particle scattering (i.e. discontinuous regions of electron density contrast) as expected from a dispersion of nanoscale aggregates of amphiphile and water in oil (heptane).³⁷ The shapes of the scattering curves for the two diamide solutions are slightly different, suggesting that the nanoscale polar aggregate structures are sensitive to the tail-group structures. The MD-SAXS results are reflective of the experimental data in terms of the positions of the scattering minima, the intensities of the scattering maxima, as well as the subtle differences in the shape of the scattering functions between the two diamide solutions (**Figure 1**). These agreements help corroborate the simulations with experimental data, justifying the use of the atomistic simulations to investigate in detail the structural correlations that drive self-assembly. It is noteworthy that even though SAXS scattering data have been previously calculated from atomistic simulation by Guilbaud, Zemb and co-workers,^{10, 19, 20, 26, 28} and us,^{22, 23, 29} the exact reproduction of experimental SAXS in the very low- q range (long distance in the real space) is still highly challenging to date due to the finite size effect and the accuracy of atomistic



simulation potentials.

Figure 2. Aggregation of diamide headgroups. Structures in the final simulation configurations of (a) DMDBTDMA and (b) DMDOHEMA systems, where the neighbors of diamide headgroups are represented by blue and red lines, respectively. Three aggregates are highlighted using beaded markers in each system. (c) headgroup-headgroup RDFs collected over the simulation trajectory. The C3 atoms (**Scheme 1**) were employed to represent the headgroups, which are defined as neighbors when closer than a cutoff distance of 7.7 Å (the first minimum on the RDF).

Inspection of the final simulation snapshots gives us a first glimpse of how the water-amphiphile-oil solutions are structured. **Figure 2a** shows the final snapshot for the DMDBTDMA solution, highlighting structural correlations between neighbor amphiphile headgroups using blue lines. It is demonstrated that linear aggregates dominate, which are made up from amphiphile headgroups. Three of them are highlighted by beaded markers. In contrast, the final snapshot for the DMDOHEMA solution (**Figure 2b**) shows some closed loop structures, along with many branched structures made up from amphiphile headgroups.

The different structures in the two diamide solutions is also indicated from the headgroup-headgroup radial distribution functions (RDFs) (**Figure 2c**), which shows stronger correlations for DMDBTDMA. In particular, the stronger correlations in the long-distance region (from 10 Å to 14 Å) indicate the formation of worm-like chains in the DMDBTDMA system. To quantitatively calibrate the aggregation behavior of these two types of diamide molecules, we calculated the average number of diamides per aggregate, using the C3-C3 cutoff distance of 7.7 Å. A higher value of 4.5 ± 0.2 was obtained for DMDBTDMA than 3.6 ± 0.1 for DMDOHEMA, indicating the presence of larger aggregates in the DMDBTDMA solution. The relatively larger error bar in the former is expected to result from the broader distribution of the aggregates (Figure S3). In Figure S3, the histogram distribution of the cluster size are plotted for both systems. Similarly the calculated coordination number of C3 around C3 carbon atoms is slightly larger for DMDBTDMA (1.3) than for DMDOHEMA (1.2). The calculated aggregation numbers agree with the previous reported data of 4 in tetrahydrogenated propylene or 6 in benzene

solution for DMDBTDMA,³³ and around 4 DMDOHEMA molecules per aggregate at DMDOHEMA concentration less than 1 M in *n*-heptane solution.³⁴

Morphologies of aggregates

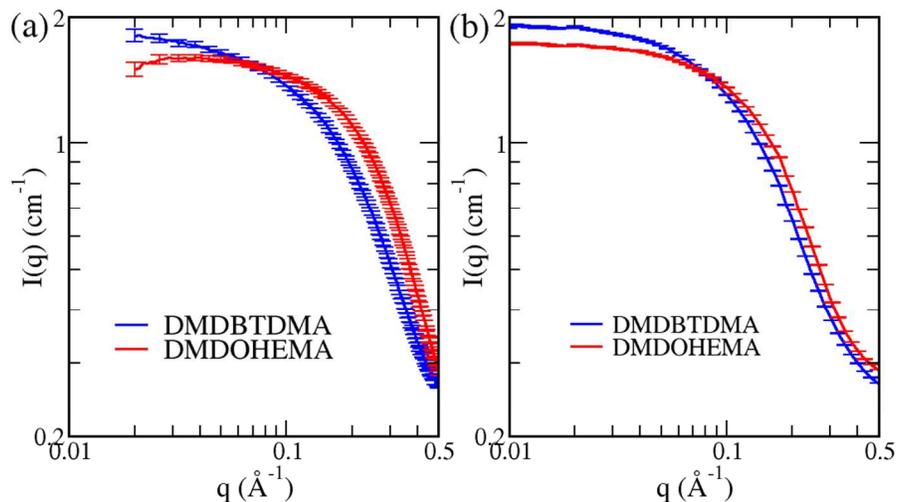


Figure 3. Background-subtracted and normalized SAXS and SANS. Experimental (a) SAXS and (b) SANS data after removal of solvent background and normalization with water standard. The upper limit of the low- q regime is set at 0.5 \AA^{-1} due to reliability of background removal in the higher- q domain. The error bars are also included.

To derive quantitative information about aggregate morphology from the experimental SAXS and SANS data, the first step is to subtract the scattering background arising from solvent contributions.^{38, 39} The general shapes of the background-subtracted SANS functions are similar to the SAXS functions, as shown in **Figure 3**. Some qualitative information could be first inferred that a slightly more sloped profile is exhibited in the low- q region for DMDBTDMA than for DMDOHEMA. A sloped profile in the low- q is associated with more elongated structures, whereas flat profiles in the low- q are more consistent with spherical globules.³⁷

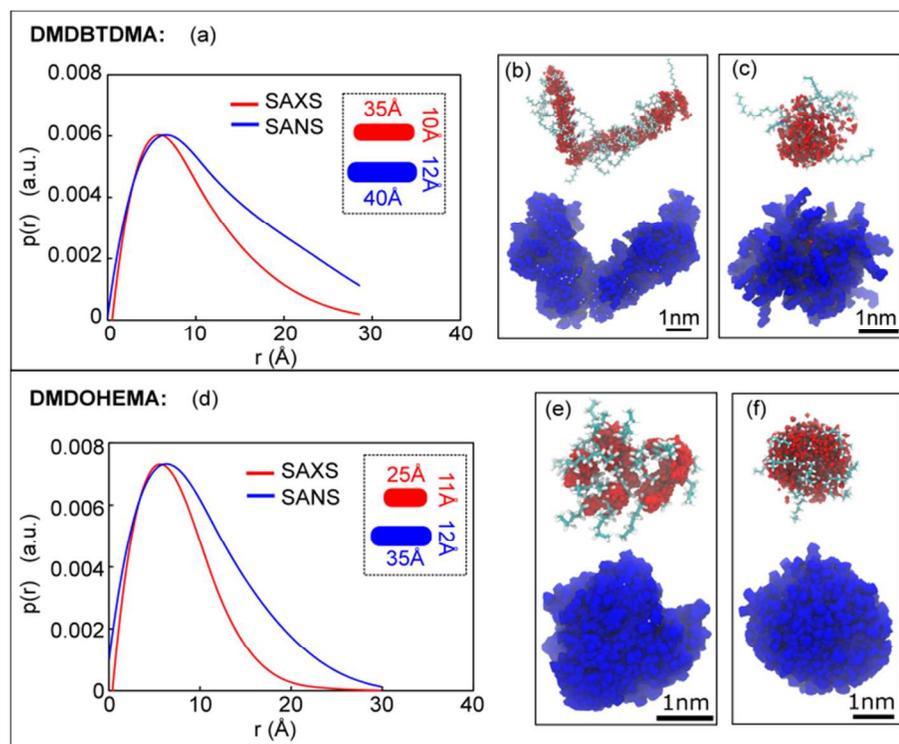


Figure 4. Real space functions from SAXS and SANS. Real space $p(r)$ functions corresponding to experimental aggregate structures in the (a) DMDBTDMA and (d) DMDOHEMA solutions. The calculated dimensions are provided in the insets. Panels (b, c, e, f) are visualizations of aggregates over MD simulations with SAXS visible polar core atoms (N and O) colored in red, and SANS visible atoms in blue. (b) is a DMDBTDMA 16-mer chain over 0.25 ns simulation time; (c) water centered (6 waters) DMDBTDMA aggregate over 0.28 ns. (e) is a DMDOHEMA 5-mer over 1 ns; (f) water centered (5 waters) DMDOHEMA aggregate over 1 ns. Diamides and waters are based on the last frame configurations, with C/O/N/H colored in cyan/red/blue/white, respectively. The other atoms are omitted for the display.

Subsequently the generalized indirect Fourier transform (GIFT) method (SI)^{40, 41} was applied to convert from reciprocal space into real space functions, as we previously reported.^{22, 38, 42, 43} Presented in **Figure 4a,d** are the real space probability distribution $p(r)$ at given scattering distances (r) between particles in the assemblies as viewed by SAXS and SANS. The SAXS $p(r)$ functions for both diamide solutions are asymmetrical bell-shaped curves with positive displacement into high r . Such functions are indicative of globular aggregates with ellipsoidal

geometry.⁴⁴ The SANS $p(r)$ functions are slightly larger than those obtained from SAXS, which is expected because of the larger neutron scattering cross-sections imparted by the hydrogenated amphiphile tail-groups. For DMDOHEMA, the SANS $p(r)$ function is similar in profile to the SAXS $p(r)$, with a bell-shaped curve with asymmetrical displacement towards high r . In contrast, the SANS $p(r)$ function for DMDBTDMA is notably different from the corresponding SAXS $p(r)$, with a more pronounced linear slope that extends towards high r . This $p(r)$ characteristic is associated with worm-like or cylindrical aggregates.⁴⁴ In both cases, the dimensions of the aggregates can be described with a long and a short axis. The average short axis diameter is estimated from the second inflection point of the $p(r)$ function (calculated using the second differential), and the maximum linear extent of the ellipsoid/worm-like aggregate is taken from the point where the function decays to 0.⁴⁴ These measurements are shown in the insets of **Figure 4a,d**, supporting that the aggregates for DMDBTDMA are more elongated than for DMDOHEMA.

The sizes of the aggregates inferred from the experimental $p(r)$ functions (**Figure 4a,d**) are in qualitative accordance with the average number of diamides per aggregate in the atomistic simulations (4.5 ± 0.2 DMDBTDMA per aggregate, 3.6 ± 0.1 for DMDOHEMA). Defining average size in weak aggregates is challenging because the aggregates are polydisperse (Figure S2) and very dynamic: merging, dissipating and changing continuously over time. The histogram distribution of diamide number per aggregate shows no critical aggregate size in the systems investigated (Figure S2). Nevertheless, the reasonably good agreements between the SAXS data in **Figure 1** support consistent structures in experiments and simulations. The atomistic MD simulations are also a useful visualization tool to connect morphological information from the $p(r)$ functions to the molecular structure of aggregates. In the simulations, the atom coordinates are saved over the entire simulation trajectory. By inspection of simulation trajectories we can illustrate how isolated aggregates might produce experimental SAXS and SANS scattering profiles. To obtain the averaged visualization of the X-ray scattering polar cores of aggregates, we plot the accumulated distributions of O and N atoms in the aggregate polar cores over time.^{20,}
²² Similarly we plot the accumulated distributions of the neutron scattering atoms that are dominated by the amphiphile tail-groups to obtain the neutron scattering components of the same aggregate. The translation and rotation of the diamide headgroups have been first removed to

exclude the influence of diffusion behaviors, and then the distributions of the X-ray or neutron scattering atoms are plotted over the same amount of simulation time in a single panel. The accumulated scattering profiles of some selected aggregate examples are given in **Figure 4**(b, c, e, f), showing the accumulated positions of the X-ray sensitive polar atoms in red and the neutron sensitive atoms in blue.

As discussed above and from inspection of the simulation snapshots, there appears to be two different types of aggregate that coexist: (i) aggregates consisting of self-associated diamide amphiphiles that interact through noncovalent interactions between headgroups; and (ii) water-centered structures where a cluster of water molecules forms a core around which amphiphile molecules aggregate. **Figure 4b** shows a self-associated worm-like DMDBTDMA aggregate that results in a rather thin and elongated polar core consisting of a chain of O and N atoms that has a time-averaged diameter of 3 - 10 Å. This rather “ropey” electron cloud with a small X-ray scattering cross-section is expected to contribute less to the SAXS signal than the globular water centered aggregate type in **Figure 4c**, which has a much larger diameter of around 20 Å. In contrast, the neutron scattering profile of the worm-like aggregate in **Figure 4b** is much thicker than its X-ray scattering profile, thanks to the long aliphatic amphiphile tail-groups that yield a time-averaged cylindrical-shaped profile with a diameter of more than 20 Å. The elongated profile of the self-associated worm-like DMDBTDMA aggregates may account for the elongated nature of the $p(r)$ functions for DMDBTDMA (**Figure 4a**), especially for the neutron scattering $p(r)$ function that is suggestive of cylindrical-shaped structures. Unlike DMDBTDMA, elongated time averaged aggregate structures are much rarer for DMDOHEMA, which instead ends to form more branched or looped self-associated structures (**Figure 4e**, **Figure 2b**). The time averaged X-ray and neutron scattering profiles of these self-associated structures are globular/elliptical in morphology, and similar to the water-centered DMDOHEMA aggregates in **Figure 4f**. This likely explains why the experimental $p(r)$ functions for DMDOHEMA are less elongated than for DMDBTDMA. Although the structures shown in **Figure 4**(b, c, e, f) are isolated snapshots, they give examples of time averaged morphologies that roughly reflect those suggested by the experimental $p(r)$ functions (**Figure 4a**, d). The quantitative definition and calculation of these different aggregate morphologies are unavailable to date. Moreover, the calculation of the

histogram distributions of the diamide headgroups support that the DMDBTDMA favor larger aggregates, whereas smaller aggregates for DMDOHEMA molecules (Figure S2).

Energetic origin of the difference in aggregate morphologies

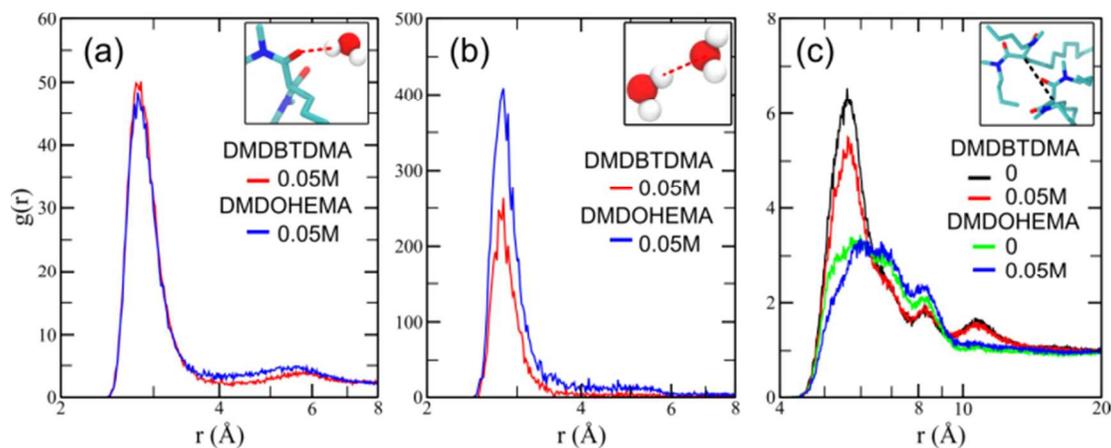


Figure 5. Correlations between different molecules. (a) Diamide carbonyl-water, (b) water-water and (c) headgroup-headgroup RDFs in the control simulations at different water concentrations (0, 0.05 M). Oxygen is employed for carbonyl and water, and C3 (**Scheme 1**) for headgroup. The insets show the corresponding correlation atom pairs, where the red dashed lines stand for the H-bonds and the black dashed line for the C3-C3 neighbor.

To understand how the aggregate morphologies affect water solubilization, we use atomistic simulations to investigate more deeply the noncovalent interactions between the polar solutes. To provide comparable conditions, control atomistic simulations were conducted at 0 and 0.05 M water concentrations for both diamides. The diamide carbonyl oxygen-water oxygen RDFs for both diamides are shown in **Figure 5a**. The primary peaks at 2.9 Å indicate the presence of H-bonds between them.²¹ The intensity of this peak is strikingly similar for both diamides, suggesting that the different tail-group structures have an insignificant effect on the strength of the H-bonds between the diamide headgroups and water. In contrast, the primary peak between water molecules in the oxygen-oxygen RDFs in **Figure 5b** is higher for DMDOHEMA than for DMDBTDMA. This shows that the water-water interactions – a prerequisite for inverted micelles or inverted proto-micelles – are influenced by the tail-group structure. In specific, the enhanced water-water correlations up to around 6 Å suggest the presence of stronger and larger

water aggregates in the DMDOHEMA system (see also Figure S4). Tail-group structure also has a significant influence on diamide headgroup-headgroup correlation (**Figure 5c**). With the increase of water concentration, the intensity of the diamide headgroup-headgroup RDF maxima is significantly decreased for DMDBTDMA, but remains approximately constant for DMDOHEMA. Additional calculations of the RDFs between diamide tail carbon atoms and n-heptane carbon atoms suggested very weak correlation between these hydrophobic groups (Figure S4).

Table 1. Noncovalent Energies (E in $k_B T$) at Different Water Concentrations ^a

[H ₂ O] /M	H ₂ O-headgroup	H ₂ O-H ₂ O	headgroup-headgroup
0 (DMDBTDMA)	-	-	-7.8±0.3
0.05 (DMDBTDMA)	1.22±0.04	-0.9±0.2	-6.8±0.2
0 (DMDOHEMA)	-	-	-5.5±0.4
0.05 (DMDOHEMA)	1.16±0.04	-1.5±0.3	-5.2±0.1

a) The total energies (Coulomb + van der Waals) have been rescaled by the numbers of waters in the simulation box for $E(\text{H}_2\text{O}-\text{H}_2\text{O})$, and by the numbers of diamides for $E(\text{H}_2\text{O}-\text{headgroup})$ and $E(\text{headgroup}-\text{headgroup})$. The total interaction energies are provided in Table S2. The standard deviations are provided as the uncertainties.

We further calculated the average interaction energies between waters and diamide headgroups by using an in-house program.¹⁸ In these calculations, waters and diamide headgroups are considered as neighbors when their distances between water oxygen (O) and diamide C3 carbon atom are less than the corresponding cutoff distances, which are obtained from the first minima of the O-C3, O-O and C3-C3 RDFs (6.0 Å, 3.5 Å and 7.7 Å, respectively). Then for the labelled neighbors the Coulomb and van der Waals interactions between all the atoms on group A (e.g., water) and all the atoms on group B (e.g., diamide headgroup, see **Scheme 1**) are calculated. This is repeated over all the neighbors found in the whole simulation trajectory, with the total energies per frame reported in **Table 1**. Note that due to the difference in the sizes of the simulation systems (ESI), the interaction energies have been rescaled for the convenience of comparisons. It is shown that with the introduction of waters the DMDBTDMA headgroup-

headgroup interactions drop heavily by around 13% ($-7.8 k_B T$ vs. $-6.8 k_B T$), whereas it is negligibly affected for DMDOHEMA with the uncertainties considered ($-5.5 \pm 0.4 k_B T$ vs. $-5.2 \pm 0.1 k_B T$), in agreement with the findings in the RDFs (**Figure 5c**). At 0.05 M water, the water-water interactions are stronger in the DMDOHEMA solution ($-1.5 k_B T$) than in the DMDBTDMA ($-0.9 k_B T$), in line with the structural difference in **Figure 5b**; whereas the water-diamide headgroup interactions show approximately negligible dependence on the types of diamides investigated here ($-1.22 \pm 0.04 k_B T$ vs. $-1.16 \pm 0.04 k_B T$), in consistence with **Figure 5a**.

In addition to the noncovalent interactions, we also calculated the cohesive energy ($-k_B T \ln(g(r))$) based on the $g(r)$ in **Figure 5**.⁴⁵⁻⁴⁷ The cohesive energy represents the effective interactions between different molecules in certain solution. The obtained results are listed in Table S3, which are similar to the calculated noncovalent interactions in **Table 1**. Consequently, the calculations of both the structures (**Figure 5**) and energies (**Table 1** and Table S3) demonstrate that the presence of water has a disruptive effect on the DMDBTDMA phase but not for DMDOHEMA.

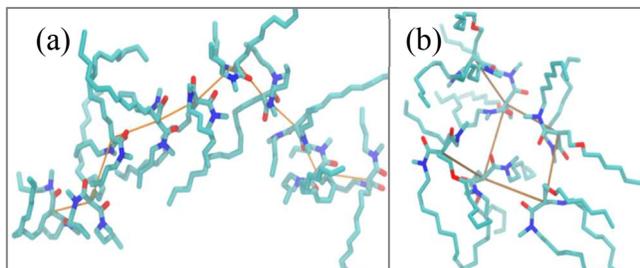


Figure 6. Snapshots of aggregates in the (a) DMDBTDMA and (b) DMDOHEMA control simulations without water. Orange lines represent the formation of C3 carbon neighbors, which are just to guide the eyes. The diamide hydrogen atoms and heptane molecules are omitted for the display.

Table 1 also shows that the diamide headgroup-headgroup interactions are stronger in the DMDBTDMA solution than in the DMDOHEMA (e.g., $-7.8 k_B T$ vs. $-5.5 k_B T$ in the absence of water). The inspection of the aggregates in the simulation trajectory indicates that DMDBTDMA molecules favor the orientation that their headgroups are approximately aligned to each other

(**Figure 6a**). This orientation evidences the significant role of dipolar interactions in stabilizing the worm-like aggregates. In contrast, the DMDOHEMA headgroups favor curved aggregates with loops and branches (**Figure 6b**), where the headgroups are interacting with each other by means of relatively weaker dipolar interactions as indicated by the deformed alignment of the headgroups. These curved aggregates in the DMDOHEMA solution thus provide the local environment to readily accommodate water molecules in forming (water-centered) inverted proto-micelles. The calculated average numbers of diamides per aggregate are 5.2 ± 0.6 (0 M H₂O) and 4.4 ± 0.2 (0.05 M H₂O) for DMDBTDMA, and 3.5 ± 0.3 (0 M H₂O) and 3.5 ± 0.1 (0.05 M H₂O) for DMDOHEMA, consistent with the changes in the headgroup-headgroup noncovalent energies. It could thus be concluded that in the DMDBTDMA solution the energy barrier is relatively higher upon the water dispersion in breaking the long worm-like aggregates into small inverted (water centered) proto-micelles. In comparison, DMDOHEMA are pre-organized into inverted proto-micelles structures that readily allow the incorporation of water. These findings combined explain why it is 35% more effective at solubilizing water (0.103 M) in the DMDOHEMA solution than in the DMDBTDMA solution (0.076 M water).

Conclusions

By combining X-ray and neutron scattering with atomistic MD simulations, we studied how diamide amphiphiles self-assemble in heptane oil solutions via weak noncovalent interactions. Worm-like aggregates and loose inverted proto-micelles are formed depending on the diamides employed, the balance of which is defined by the amphiphile tail-group structures. The inverted proto-micelles are found to be more effective at solubilizing water in oil. This is ascribed to the pre-organized structures of the inverted proto-micelles, which are more energetically ready in accommodating water. Our results further our understanding of self-assembly driven by weak noncovalent interactions, elucidating rich structure at the intersection between solution and colloidal science.

Materials and Methods

The experimental details are provided in the ESI. They are summarized here. The extraction of water from the aqueous phase to the organic *n*-heptane phase was experimentally performed by contacting 0.5 M diamide of *n*-heptane solution with equal volume of water, followed with centrifugation for phase separation. The oil phase was subsequently separated from the water

phase. And the water concentration in oil was determined using the Karl Fisher titration method. The SAXS and SANS measurements were subsequently performed for the water-amphiphile-oil solutions. The SAXS measurements were conducted at the Advanced Photon Source (APS) at Argonne National Laboratory. And the SANS were performed at the High Flux Isotope Reactor (HFIR) facility at Oak Ridge National Laboratory.

The atomistic MD simulations were performed using the modified OPLS-AA force field for the long hydrocarbon molecules.⁴⁸ The amounts of molecules were determined to reflect the experimental concentrations. All the molecules were initially randomly distributed, followed by annealing simulations²¹⁻²³ to equilibrate the systems. The production simulations were performed using the NTP ensemble, with a simulation time of 50 ns each. The SWAXS calculations were then performed based on the production simulation trajectories.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

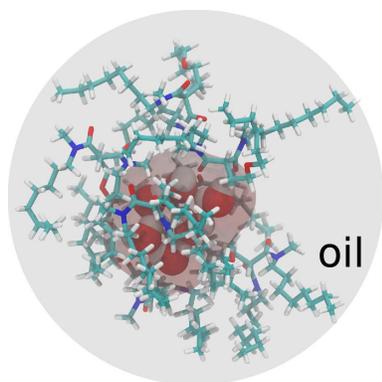
The authors thank Prof. Monica Olvera de la Cruz for valuable suggestions. This work is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Division of Chemical Sciences, Biosciences and Geosciences, under contract numbers DE-AC02-05CH11231.

References

1. P. Hobza and J. Řezáč, *Chem. Rev.*, 2016, 116, 4911-4912.
2. K. T. Mahmudov, M. N. Kopylovich, M. F. C. Guedes da Silva and A. J. L. Pombeiro, *Coord. Chem. Rev.*, 2017, 345, 54-72.
3. T. Aida, E. W. Meijer and S. I. Stupp, *Science*, 2012, 335, 813-817.
4. W. E. Bentley and G. F. Payne, *Science*, 2013, 341, 136-137.
5. T. Gunnlaugsson, *Nat. Chem.*, 2015, 8, 6.
6. H.-L. Au-Yeung, A. Y.-Y. Tam, S. Y.-L. Leung and V. W.-W. Yam, *Chem. Sci.*, 2017, 8, 2267-2276.
7. Z. Yu, A. Erbas, F. Tantakitti, L. C. Palmer, J. A. Jackman, M. Olvera de la Cruz, N.-J. Cho and S. I. Stupp, *J. Am. Chem. Soc.*, 2017, 139, 7823-7830.
8. T. Zemb and W. Kunz, *Curr. Opin. Colloid Interface Sci.*, 2016, 22, 113-119.
9. R. Demichelis, P. Raiteri, J. D. Gale, D. Quigley and D. Gebauer, *Nat. Commun.*, 2011, 2, 590.
10. T. N. Zemb, M. Klossek, T. Lopian, J. Marcus, S. Schöettl, D. Horinek, S. F. Prevost, D. Touraud, O. Diat, S. Marčelja and W. Kunz, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, 113, 4260-4265.

11. T. Zemb, C. Bauer, P. Bauduin, L. Belloni, C. Dejumat, O. Diat, V. Dubois, J.-F. Dufreche, S. Dourdain, M. Duvail, C. Larpent, F. Testard and S. Pellet-Rostaing, *Colloid Polym. Sci.*, 2015, 293, 1-22.
12. R. J. Ellis, Y. Meridiano, J. Muller, L. Berthon, P. Guilbaud, N. Zorz, M. R. Antonio, T. Demars and T. Zemb, *Chem. Eur. J.*, 2014, 20, 12796-12807.
13. M. P. Jensen, T. Yaita and R. Chiarizia, *Langmuir*, 2007, 23, 4765-4774.
14. C. Dejumat, S. Dourdain, V. Dubois, L. Berthon, S. Pellet-Rostaing, J.-F. Dufreche and T. Zemb, *Phys. Chem. Chem. Phys.*, 2014, 16, 7339-7349.
15. P. Bauduin, F. Testard, L. Berthon and T. Zemb, *Phys. Chem. Chem. Phys.*, 2007, 9, 3776-3785.
16. R. J. Ellis, *J. Phys. Chem. B*, 2014, 118, 315-322.
17. C. Erlinger, L. Belloni, T. Zemb and C. Madic, *Langmuir*, 1999, 15, 2290-2300.
18. B. Panganiban, B. Qiao, T. Jiang, C. DelRe, M. M. Obadia, T. D. Nguyen, A. A. A. Smith, A. Hall, I. Sit, M. G. Crosby, P. B. Dennis, E. Drockenmuller, M. Olvera de la Cruz and T. Xu, *Science*, 2018, 359, 1239-1243.
19. G. Ferru, D. Gomes Rodrigues, L. Berthon, O. Diat, P. Bauduin and P. Guilbaud, *Angew. Chem., Int. Ed.*, 2014, 53, 5346-5350.
20. P. Guilbaud and T. Zemb, *Curr. Opin. Colloid Interface Sci.*, 2015, 20, 71-77.
21. B. Qiao, T. Demars, M. Olvera de la Cruz and R. J. Ellis, *J. Phys. Chem. Lett.*, 2014, 5, 1440-1444.
22. B. Qiao, G. Ferru, M. Olvera de la Cruz and R. J. Ellis, *ACS Cent. Sci.*, 2015, 1, 493-503.
23. G. Ferru, B. Reinhart, M. K. Bera, M. Olvera de la Cruz, B. Qiao and R. J. Ellis, *Chem. Eur. J.*, 2016, 22, 6899-6904.
24. B. Qiao, J. V. Muntean, M. Olvera de la Cruz and R. J. Ellis, *Langmuir*, 2017, 33, 6135-6142.
25. D. M. Brigham, A. S. Ivanov, B. A. Moyer, L. H. Delmau, V. S. Bryantsev and R. J. Ellis, *J. Am. Chem. Soc.*, 2017, 139, 17350-17358.
26. P. Guilbaud and T. Zemb, *ChemPhysChem*, 2012, 13, 687-691.
27. B. Qiao, G. Ferru and R. J. Ellis, *Chem. Eur. J.*, 2017, 23, 427-436.
28. S. Schoettl, J. Marcus, O. Diat, D. Touraud, W. Kunz, T. Zemb and D. Horinek, *Chem. Sci.*, 2014, 5, 2949-2954.
29. A. W. Knight, B. Qiao, R. Chiarizia, G. Ferru, T. Forbes, R. J. Ellis and L. Soderholm, *Langmuir*, 2017, 33, 3776-3786.
30. L. Berthon, J. M. Morel, N. Zorz, C. Nicol, H. Virelizier and C. Madic, *Sep. Sci. Technol.*, 2001, 36, 709-728.
31. L. Xie, D. K. Harris, M. G. Bawendi and K. F. Jensen, *Chem. Mater.*, 2015, 27, 5058-5063.
32. T. Okuhara, *Chem. Rev.*, 2002, 102, 3641-3666.
33. L. Nigond, C. Musikas and C. Cuillerdier, *Solvent Extr. Ion Exch.*, 1994, 12, 261-296.
34. Y. Meridiano, L. Berthon, X. Crozes, C. Sorel, P. Dannus, M. R. Antonio, R. Chiarizia and T. Zemb, *Solvent Extr. Ion Exch.*, 2009, 27, 607-637.
35. U. S. Jeng, T.-L. Lin, C.-S. Tsao, C.-H. Lee, T. Canteenwala, L. Y. Wang, L. Y. Chiang and C. C. Han, *J. Phys. Chem. B*, 1999, 103, 1059-1063.
36. U. Jeng, T. L. Lin, Y. Hu, T. S. Chang, T. Canteenwala, L. Y. Chiang and H. Frielinghaus, *J. Phys. Chem. A*, 2002, 106, 12209-12213.
37. O. Glatter, in *Neutrons, X-Rays and Light*, eds. P. Lindner and T. Zemb, Elsevier Science B.V., Amsterdam, 2002, pp. 73-102.
38. R. J. Ellis and M. R. Antonio, *Langmuir*, 2012, 28, 5987-5998.
39. R. Chiarizia, D. Stepinski and M. R. Antonio, *Sep. Sci. Technol.*, 2010, 45, 1668-1678.
40. G. Fritz and O. Glatter, *J. Phys.: Condens. Matter*, 2006, 18, S2403-S2419.
41. G. Fritz, A. Bergmann and O. Glatter, *J. Chem. Phys.*, 2000, 113, 9733-9740.

42. R. J. Ellis, Y. Meridiano, R. Chiarizia, L. Berthon, J. Muller, L. Couston and M. R. Antonio, *Chem. Eur. J.*, 2013, 19, 2663-2675.
43. R. J. Ellis, M. Audras and M. R. Antonio, *Langmuir*, 2012, 28, 15498-15504.
44. O. Glatter, *J. Appl. Crystallogr.*, 1979, 12, 166-175.
45. D. A. McQuarrie, *Statistical Mechanics*, Harper Collins, New York, 1976.
46. D. Chandler, *Introduction to Modern Statistical Mechanics*, Oxford University Press, New York, 1987.
47. C.-Y. Leung, L. C. Palmer, B. Qiao, S. Kewalramani, R. Sknepnek, C. J. Newcomb, M. A. Greenfield, G. Vernizzi, S. I. Stupp, M. J. Bedzyk and M. Olvera de la Cruz, *ACS Nano*, 2012, 6, 10901-10909.
48. S. W. I. Siu, K. Pluhackova and R. A. Bockmann, *J. Chem. Theory Comput.*, 2012, 8, 1459-1470.



Weak noncovalent interactions control water dispersion and solubility in oil.