THE ROLE OF THE ASYMMETRIC BOLAAMPHIPHILIC CHARACTER OF VECAR ON THE KINETIC AND STRUCTURAL ASPECTS OF ITS SELF-ASSEMBLY: A MOLECULAR DYNAMICS SIMULATION STUDY

Hye-Young Kim^{a,*}, Brian R. Novak^{b,c}, Bijay Shrestha^a, S. Emma Lee^d, and Dorel Moldovan^{b,c}

^a Department of Chemistry and Physics, Southeastern Louisiana University, Hammond, LA 70402.

^b Mechanical & Industrial Engineering Department, Louisiana State University, Baton Rouge, LA 70803.

^c Center for Computation and Technology, Louisiana State University, Baton Rouge, LA 70803.

^d George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, GA 30332.

ABSTRACT

VECAR are novel bolaamphiphilic molecules consisting of two hydrophilic molecular groups, a carnosine derivative and a chromanol group, covalently linked by a hydrophobic alkyl spacer of varying length. Despite the potential for application in various biomedical applications VECAR properties, including their bulk properties, are still largely unknown. The early stage of the self-assembly process of VECAR molecules in water is studied using molecular dynamics simulations. The study reveals that the length of the hydrophobic spacer in VECAR affects the aggregation kinetics as well as the size, shape, density, and atomistic structure of the self-assembled aggregates. A mechanism based on cooperative interactions between water, the hydrophilic hydroxyl group, and the hydrophobic benzene ring of the chromanol head is proposed to explain the ordered packings of chromanols in the self-assembled aggregate structures at the aggregate-water interface.

KEY WORDS

bolaamphiphile, carnosine, chromanol, micelle, aromatic stacking

*corresponding author.

Email addresses: hye-young.kim@selu.edu (H.-Y. Kim)

1. INTRODUCTION

Self-assembly process of amphiphilic molecules in solutions, due mostly to the hydrophobic effect [1,2], is important in a variety of areas ranging from medicine, drug discovery, and biotechnologies to industries such as electronics and nanomaterials [3,4]. The amphiphilic molecules, also known as surfactants, form aggregates of different sizes and shapes depending on their molecular structure and interactions [5-9]. Some of the bio-amphiphiles, such as phospholipids, self-assemble in nearly flat bilayer membranes that are the building blocks of vesicles (liposomes). Most of the commonly used surfactants form micelles with radii of curvature close to the length of their hydrocarbon chains [10]. The structure, thermodynamics, and dynamic properties of aggregates or films of amphiphilic molecules are of prime importance in a great variety of systems [11-16].

Bolaamphiphiles are amphiphilic molecules with two hydrophilic ends which are covalently connected by hydrophobic alkyl spacers. In this study, we report the results of molecular dynamics simulations of bulk and aqueous solutions of a novel class of bolaamphiphiles, synthesized recently, called VECAR [17]. VECAR molecules are obtained by linking the head group of α -tocopherol and carnosine derivative with a single hydrocarbon chain [17] and have the structure and properties of asymmetric bolaamphiphiles with hydrophobicity that can be tailored for specific applications by varying the length of the hydrocarbon chain.

VECAR molecules are of particular interest as they are derived from compounds occurring naturally and have potential for being used in gene- and drug-delivery vehicles in biomedical applications. Moreover, they are nontoxic, have antioxidant properties, and environmentally are safe surfactants [17]. Despite their potential in various medical applications such as prevention of oxidative stress-related diseases and atherosclerosis, or as a genetic vehicle for delivery of a wide array of therapeutic compounds through the lipoprotein transport mechanisms [17], the bulk properties of VECAR are largely unknown.

The self-assembly process is often analyzed using simple theories which combine thermodynamics and geometrical arguments [7]. Accordingly, the ability of surfactants to form particular type of self-assembly structures is governed by the molecular packing parameter. Considering only the chemical structure of the surfactant molecules, the packing parameter depends on the size of the hydrophilic head group(s) and the length of the hydrophobic chain(s). In previous studies [18-20] with surfactants of various shapes [9], the size of the hydrocarbon chains has mostly been related to the number and length of their tails. However, the unique structure of the VECAR molecule with heterogeneous hydrophilic heads attached at both ends of a hydrophobic chain is expected to lead to a different mechanism of self-assembly. In general, based on experimental investigations [21,22] alone, it is difficult to formulate an understanding of how the structure of self-assemblies is related to the chemical composition and geometry of individual surfactant molecules, and in turn, how structural features of the aggregates can be tuned by controlling specific parameters, and how these would in the end modulate the bulk properties of the self-assembled systems. The atomistic level simulations presented in this study will contribute to the development of a better understanding of the formation of the self-assembled structures formed by these novel molecules, and thereby provide new insight and enhance the interest in further experimental and theoretical investigations of VECAR or other asymmetric bolaamphiphilic molecules.

In the present study, the atomistic details of the self-assembled structures formed by VECAR molecules in water have been studied with large-scale molecular dynamics (MD) simulations using GROMACS [23,24]. The goal is to investigate the kinetic and structural characteristics of the early stages of the self-assembly of VECAR molecules. Our data analysis includes cluster size distributions, dynamics of the fusion of two spheroidal micelles, atomistic and molecular morphology of aggregates of varying sizes and shapes, and stacking of aromatic rings. The main focus will be on understanding the behavior and the role of the two heterogeneous hydrophilic heads and the hydrophobic alkyl spacer on the size, structure, and

kinetic process of self-assembly. Bulk VECAR was also simulated in order to compare the arrangement of the atoms in aqueous VECAR aggregates with the corresponding structure of the bulk liquid.

2. SIMULATION METHODOLOGY

The length of the alkyl spacer (n) of the synthesized VECAR molecules varied from 4 to 22 [17]. In the present MD simulation study we specifically chose the shortest ($n=4$) and the longest ($n=22$) chain lengths that have been synthesized and an additional medium ($n=13$) chain length to investigate the dependence on the chain length. Abbreviated notations of VC-C4, VC-C13, and VC-C22 will henceforth be used for VECAR molecules with chain lengths $n=4$, 13, and 22, respectively. The α -tocopherol end of the VECAR molecules is 2,5,7,8-tetramethyl-6-chromanol, and will be referred to as simply chromanol; the other head group, the carnosine derivative, is methyl (2S)-2-(3-formamidopropanoylamino)-3-(1H-imidazol-4-yl) propanoate and will be referred to as carnosine henceforth. Figure 1 presents the chemical structure of VECAR molecules in which the atoms are labeled according to their location. The atomic site identification number (ID) used to label the atoms is an arbitrary number assigned to each atom or united atom in the VECAR molecule starting with 1 at the far end of the chromanol side and ending with 62 at the far end of the carnosine side as depicted in Fig. 1 for VC-C22. The united atoms in the alkyl spacer have site IDs ranging from 17 through 20 for VC-C4, 17 through 29 for VC-C13, and 17 through 38 for VC-C22. These site IDs will be referred to in the subsequent analysis. Some of the sites represent individual atoms (like in hydrophilic groups OH and NH) and some represent united atoms. Henceforth, a united atom will be denoted as either an atom or site for simplicity.

Figure 1. Chemical structure of VECAR. The numbered labels are the atom site IDs used in the discussion. Site IDs for alkyl chains of varying lengths are: 17-20 for $n=4$, 17-29 for $n=13$, and 17-38 for $n=22$.

Since VECAR is a recently synthesized molecule which has not been previously investigated with MD simulations, a united atom parameterization for VECAR molecules with each of the three chain lengths is developed using the Automated Topology Builder (ATB) [25]. To build the topologies and parameters at the B3LYP/6-13G* level, which limits the total number of atoms to a maximum of 40, we first obtained topologies from ATB for four overlapping smaller pieces of a

VECAR molecule and then constructed the final topology from those. According to the ATB Validation Statistics Tool, the vacuum RMSD values between the minimized structures from DFT and using the ATB parameters of the four sections we used range from 0.00761 to 0.01325 nm, which suggests that the ATB parameters are reasonable. GROMOS 53a6 united atom force field [26] is used, and the topology files for these VECAR molecules are available in supplementary information. The simple point charge (SPC) model [27] is adopted for water molecules.

Five distinct systems with VECAR molecules of varying spacer lengths or varying concentrations are simulated in water at 300 K and 1 atm. The details of the systems studied are listed in Table 1, and each aqueous system contains 160,000 water molecules. To reduce the computational cost and yet to be able to observe the dynamics of the early stage of self-assembly until all molecules merge to form a final single assembly of a reasonable size within a reasonable time, high molecular concentrations are studied. We have also carried out additional simulations of bulk systems with 125 VECAR molecules at 300K and 1 atm. The initial random configurations of all simulations are generated in a cubic box using packmol [28].

Table 1. Details of VECAR aqueous systems. Each aqueous system contains 160,000 water molecules.

System	Number of VECAR	Total run time (ns)	wt. % VECAR
VC-C4/500	500	166.1	8.4
VC-C13/500	500	191.0	10.2
VC-C13/403	403	208.1	8.4
VC-C22/500	500	148.6	11.9
VC-C22/338	338	308.1	8.4

MD simulations are carried out using the GROMACS 4.6 simulation package [23,24]. Periodic boundary conditions are used in all directions. First, the steepest descent algorithm is used for energy minimization (EM) until the maximum force on any atom is less than 100 kJ/mol/nm. For the aqueous systems, constant pressure and temperature (NPT) simulations are run for 1 ns with the positions of the non-hydrogen sites in the VECAR molecules restrained to allow the water to equilibrate around the VECAR molecules before they are allowed to aggregate. The restrained simulations are followed by unrestrained NPT simulations. For bulk systems, unrestrained NPT simulations of 50 ns are carried out following EM.

The simulation settings for non-bonded interactions are as follows. Two cutoffs are used for the van der Waals (vdw) interactions. The vdw interactions are calculated every 1 and every 10 time steps, respectively, within a 1.0 nm radius and between 1.0 nm and 1.4 nm. The short-range electrostatic interactions are calculated up to 1.0 nm. The potentials are not shifted at the cutoff. For long-range electrostatic interactions, the fast particle-mesh Ewald method is used with 0.12 nm grid spacing, order 4, and tolerance of 10^{-5} .

The simulation settings related to the integration of the equations of motions and the coupling to the external thermostat and barostat are as follows. All bonds are constrained and the equations of motion are integrated with a time step of 2 fs using the leap-frog algorithm [29]. Temperature and pressure are controlled via the velocity rescaling thermostat [30] and isotropic Berendsen pressure coupling [31] with time constants of 0.1 and 2.5 ps, respectively. The coordinates of atom sites are stored every 8 ps, and subsequently used for analyses. The

analyses of the MD trajectories are carried on by using Visual Molecular Dynamics (VMD) software package [32].

3. RESULTS AND DISCUSSION

3.1. KINETICS OF VECAR SELF-ASSEMBLY

3.1.1. CLUSTER SIZE

Due to their bolaamphiphilic nature, VECAR molecules aggregate in all aqueous systems investigated (see Table 1). The growth of the self-assembled structures in the simulated aqueous solutions can be described by the variation in time of the cluster size distributions. Figure 2 depicts the size of the largest cluster as well as the average cluster size as functions of time. The cluster size, or the number of molecules per cluster, is evaluated by using the *g_clustersize* function in GROMACS simulation package. According to the algorithm implemented in the *g_clustersize* function two molecules are considered to belong to the same cluster if the shortest separation distance between any pair of atoms, each belonging to a different molecule, is less than the cutoff distance of 0.35 nm. The average cluster size being smaller than the size of the largest cluster indicates that in addition to the largest cluster there are one or more smaller clusters in the system.

Figure 2. Time evolution of both the average cluster size (number of molecules per cluster) and the largest cluster size. The dotted and solid lines indicate the largest and the average cluster size, respectively.

In the studies of the aqueous systems, the MD simulations are terminated when all VECAR molecules in the system belong to a single self-assembled structure. The formation of one assembly indicates that the structures formed are almost certainly smaller than the average sizes of structures that would actually form at these concentrations in much larger simulation systems. That is, it is very likely that if the simulations contained a larger number of VECAR molecules then larger self-assembled structures would form. As depicted in the following section, the final configurations of the self-assembled structures do not bear the shape of any commonly known self-assembled structures. This may be due to the fact that the structures have not reached their full size and even at their current sizes have likely not equilibrated completely. Given the already large size of the systems simulated, possible resolution of these issues will require coarse grained simulations. Thereby the focus here will be only on the kinetics and structure formation in the early stages of self-assembly process of VECAR molecules.

3.1.2. DESCRIPTION OF THE EARLY STAGE OF SELF-ASSEMBLY PROCESS

The self-assembly process that leads to the final structure is a complex phenomenon occurring in three-dimensions and it is difficult to present with a series of two-dimensional snapshots or even with video animations. In addition, given that the self-assembly process has an inherent stochastic component the details of the assembly process will vary when different initial conditions are chosen. Despite these limitations, in this section, in order to give a sense of the typical events and time scales involved, we describe the self-assembly processes observed with VMD renderings of trajectory snapshots. We note here that as a testing case, we carried out two independent simulations of the smallest size system (VC-C4/500) with two different random initial configurations as discussed later in this section.

In the VC-C22/500 system, starting from an initial configuration with random distribution of VECAR molecules in water, the early stage is characterized by the formation of multiple small aggregates, which subsequently interact with each other and coalesce leading to the formation of larger clusters. After about 50 ns, the VMD rendering shows a system composed of two structures, one large branched aggregate and one small spheroidal structure. The bent branches of the aggregate collapse inward over time into an overall more spherically shaped structure. The spheroid composed of 12 VECAR molecules is quite stable and moves about in the vicinity of the large aggregate for about 63 ns. This spheroid eventually merges with the large aggregate, forming the one large, final aggregate. The total time for the evolution of this system from the initial random configuration to the final configuration containing one large self-assembled structure is about 113 ns. Figure 3 contains a few snapshots of the simulated system depicting the first 50 ns of the simulation that leads to the formation of the branched aggregate. The final configurations of all the aqueous systems simulated are depicted in Figure 4.

In the VC-C22/338 system, the initial random configuration of molecules evolves quickly, leading to many small spheroidal aggregates which subsequently interact with each other and merge to form larger ones. After about 55 ns, the system is composed of one large aggregate and two small spheroidal aggregates. The two spheroids containing 35 and 31 molecules each are stable for about 27 ns before merging to form a prolate-shaped aggregate, which again remains stable for about 179 ns. This merging process is explored in great detail in a later section. Finally, the prolate-shaped aggregate merges into the large aggregate forming the final self-assembly comprising of all 338 VC-C22 molecules. Due to the lower VECAR concentration, the overall time required to form one final aggregate is larger than in the VC-C22/500 system; 261 ns compared to 113 ns.

Figure 3. Snapshots of the VC-C22/500 system at six different times. The early formation of a branched aggregate is visible in the snapshots taken at 4 ns to 9 ns. Carbon, oxygen, nitrogen, and hydrogen atoms are colored (online) in cyan, red, blue, and white, respectively. Water molecules are not shown for clarity.

The final configurations of both VC-C13/500 and VC-C13/403 systems are elongated self-assemblies which percolate the periodic simulation box; the ends in Figure 4 are connected across the periodic boundaries. Since these elongated structures occur in both simulations, it is likely that VC-C13 actually forms elongated structures for concentrations around those used here, but more simulations or experiments would be required to confirm this.

In the VC-C13/500 system, the molecules of the initial random configuration aggregate quickly into a long branched assembly across a periodic boundary and three small spheroids that are stable for about 40 ns. Two of these spheroids then merge into the elongated assembly one after another over a period of 10 ns. The configuration of one elongated assembly and one small spheroid composed of 22 molecules lasts for about 75 ns until the small spheroid merges into the large assembly, forming the final configuration. This gives a total time of about 160 ns from the initial random configuration to the final aggregation.

In the VC-C13/403 system, the molecules in the initial random configuration assemble into three large and two small aggregates. The two small aggregates quickly merge into one of the large ones. Two of the three large aggregates merge to form a larger elongated one which is connected across the simulation box through a periodic boundary. These two merging aggregates are relatively flat and round, and the planes perpendicular to their shortest principal axes are twisted relative to each other during merging. The third of the original large aggregates is prolate-shaped and is composed of 124 molecules. This configuration with the large elongated assembly and the smaller prolate-shaped aggregate remains stable for about 113 ns until the prolate-shape aggregate merges into the larger assembly to form the final elongated self-assembly.

In the VC-C4/500 systems, the VC-C4 molecules quickly form many small elongated aggregates, which then assemble into the final structure within the initial 75 ns. In Fig. 4, VC-C4/500* is an alternative final configuration to VC-C4/500 obtained with a different initial configuration. Although the final self-assemblies of the VC-C4/500 and VC-C4/500* systems differ in the overall shape, all of our analyses demonstrate that they share the same atomistic structure within statistical error. Henceforth, we present only the results of the VC-C4/500 system.

In summary, our MD studies of the early stages of the self-assembly process of VECAR in aqueous solutions lead to three main conclusions. (1) The time it takes for the self-assembly process to reach the end point characterized by the formation of a single aggregate is larger for VECAR molecules with larger hydrocarbon chains at a given VECAR concentration, which indicates that the small aggregates of longer chain molecules are more stable. (2) The size of the small spheroidal and prolate-shaped aggregates tends to be larger for longer chain lengths for a given shape and VECAR concentration. A similar trend of increasing size of the initial aggregates as the chain length increases was also observed in the initial 30 ns of the MD studies of the aggregate formation of ionic aqueous systems [33]. (3) At higher VECAR concentrations, the small aggregates tend to merge into the larger ones more readily as demonstrated in VC-C22 systems in Fig. 2.

Figure 4. Depiction of the final configuration of each simulated system. The coloring scheme is the same as for Fig. 3. Water molecules are not shown for clarity. VC-C4/500* is an alternative final configuration to VC-C4/500 obtained with a different initial configuration.

3.1.3. ANALYSIS OF THE MERGING OF TWO SPHEROIDAL ASSEMBLIES

As mentioned previously, we observe two spheroids merging to form a prolate-shaped aggregate in the VC-C22/338 system. This event provides a relatively easy means to analyze the merging process since the geometries of the merging structures are simple. The spheroidal assemblies are composed of 35 and 31 VC-C22 molecules. When these two spheroids merge, a visible nitrogen band appears at the interface. We note here that all four nitrogen atoms in each VECAR molecule are from the strongly hydrophilic carnosine side.

Figure 5. The distance distributions of the core-site (CS) and the nitrogens (N) as the two spheroids merge to form a prolate-shape assembly in VC-C22/338 system. The snap shots were taken at: (a) 91.6 ns, (b) 113.8 ns, and (c) 224.2 ns. The nitrogen atoms are represented in larger, dark blue (online) spheres for clarity. Solid and broken curves are for the atoms belong to the spheroids on the right and left, respectively, of the center of geometry of the two-spheroids system. The symbols represent the distribution profile of sum of those two curves.

Figure 5 presents the atomistic distribution at varying distances from the geometric center of the two spheroids system along its longest principal axis. The distribution profile of each time, t_0 , is an average over 21 time frames over the time range between $t_0 - 0.08$ ns and $t_0 + 0.08$ ns. The distributions are shown for the nitrogen atoms and the hydrocarbon atom with site ID 23 at three different times during the merging process. The distribution of each site is calculated as the number of sites in each bin of width 0.1 nm. The site ID 23 atoms are consistently found at the innermost core of each spheroid, which henceforth will be denoted as the core-site. The distribution of nitrogen represents the mean of the distributions of the four nitrogen atoms (site

ID's of 41, 47, 56, and 59). Also, shown in Fig. 5 are the snapshots of the corresponding assemblies. For clarity, the nitrogen atoms are shown in larger diameter than the other atoms. In both the snapshots and the graphs in Fig. 5, the spheroids composed of 35 and 31 VC-C22 molecules are presented on the left and right side of the geometric center of the merging assembly, respectively. Right after they merge at 91.6 ns, the nitrogens from left and right spheroids overlap in a narrow region as shown in Fig. 5(a) and (b), which appeared as a visible nitrogen band in the VMD rendering. As the merging process advances, the nitrogen atoms disperse across the boundary resulting in a wider overlap region as shown in Fig. 5(c). On the other hand, the distribution of core-sites does not penetrate into the other side much throughout the entire merging process observed here.

Figure 6. The average speed of each atom in spheroid 1 and spheroid 2 along the longest principal axis of the two spheroids system. Spheroid 1 (triangles) and spheroid 2 (circles) are composed of 35 and 31 VC-C22 molecules, respectively.

To investigate the dynamics of the merging process, we calculate the mean speed of each spheroid along the longest principal axis which goes through the two spheroids. Observation of the trajectory shows that the two isolated spheroids move toward each other starting at about 82.1 ns until they merge to form one assembly at 91.6 ns as shown in Fig. 5(a). We calculate the mean linear speed of an atom by measuring the positions of the atom relative to the center of mass of two spheroids along the longest principal axis at both 82.1 ns and 91.6 ns, and then by dividing the absolute displacement along the principal axis by the travel time (9.5 ns). Figure 6 presents the average speed over the period as functions of the site IDs. On average, spheroid 2 composed of 31 molecules has a larger speed than spheroid 1 composed of 35 molecules. Using the simulation data depicted in Fig. 6 we found the ratio of the average speeds of the two spheroids to be $v_2/v_1 = 1.128$. Interestingly, this value is very close to the ratio of the masses of the two spheroids $m_1/m_2 = 35/31 = 1.129$; result that seems to be consistent with the conservation, in average, of the linear momentum of the two spheroids system.

3.2. STRUCTURE OF VECAR SELF-ASSEMBLIES

As described in the previous sections, due to the relatively slow kinetics for the typical MD time scales, it is not possible for us to determine the fully equilibrated configurations in the present simulation studies. However, the self-assemblies observed in the present studies are representative for self-assemblies of VECAR in the early stages of self-assembly and/or at very low concentrations. Additional insight into the self-assembly process can be obtained by

performing detailed analysis of the structural characteristics of some of these self-assembled structures.

3.2.1. SPHEROIDAL ASSEMBLIES: RADIAL DISTRIBUTION AND SHAPE

To address the issue of whether the small spheroidal assemblies of VECAR molecules of varying alkyl spacer length have a micelle structure, we calculate the normalized radial occupancy and density profiles. For these calculations, the assemblies are assumed to be spherically symmetric. Both normalized radial occupancy and density profiles are averaged over 50 time frames. First, the normalized radial occupancy is calculated as the number of a given type of atom in a bin of width $\Delta r = 0.025$ nm at a radial distance r from the spheroid's geometric center divided by the total number of molecules in the spheroid. Thus, the sum of the normalized radial occupancy is equal to one for each type of atom. Figure 7 shows the normalized radial occupancy profiles for three atom types in a spheroid composed of 35 VC-C22 molecules from the VC-C22/338 system. The normalized radial occupancy profiles indicate that the spheroid has a well-developed micelle structure with the hydrophobic alkyl chain (ID 23) located at its core, the strongly hydrophilic carnosine end (ID 60) located at the outermost shell, and weakly hydrophilic chromanol end (ID 1) located at the inner shell. See Figure 1 for the definition of the atom IDs. There is a significant overlap between the wide occupancy profiles of the three sites in Figure 7, especially between the two hydrophilic shells. We also find indication of well-developed micelle structures based on the normalized radial occupancy profiles of spheroidal assemblies found in the VC-C22 and VC-C13 aqueous systems. On the other hand, the normalized radial occupancy profile for a short-lived spheroidal assembly in the aqueous VC-C4 system shows no indication of a micelle structure. The three sites of VC-C4 are distributed almost evenly across the spheroidal self-assembly. Thereby, we conclude that the self-assemblies in the aqueous VC-C4/500 system do not form micelle. It is also important to note that some water molecules are found within the interior of these assemblies; this in part provides an explanation for the existence of some hydrophilic ends throughout the assemblies.

Figure 7. Normalized radial occupancy for a spheroidal assembly composed of 35 VECAR molecules in the VC-C22/338 system.

The radial density distribution of all atoms is calculated across spheroidal assemblies with a bin width of 0.05 nm. Presented in Fig. 8 are the density distributions for two spheroidal micelles composed of 22 VC-C13 and 35 VC-C22 molecules found in the VC-13/500 and VC-C22/338 systems, respectively. Even though the spheroidal assembly composed of 25 VC-C4 molecules

in the VC-C4/500 system lacks the micelle structure as discussed above, its density distribution profile is very close to that of the spheroidal micelle composed of 22 VC-C13, and thus is not shown here for clarity. From the mass density profiles shown in Fig. 8, we make a rough estimate of the radius of a spheroid as follows. For a spheroidal assembly of total mass M , assuming the structure as a perfect sphere of uniform density ρ_t , the geometric radius R of the sphere is defined as:

$$\left(\frac{4\pi}{3}R^3\right)\rho_t = M \quad (1)$$

Here, we estimate ρ_t by averaging the mass density up to the radial distance at half its maximum density. The average density (and the radius) thus estimated are 0.89 g/cm³ (2.30 nm), 0.91 g/cm³ (1.85 nm), and 0.95 g/cm³ (1.77 nm) for the spheroidal micelles composed of 35 VC-C22 molecules, 22 VC-C13 molecules, and 25 VC-C4 molecules, respectively. This estimation reveals that the density of VECAR assemblies in aqueous systems decreases as the alkyl chain length increases.

Figure 8. The mass density profiles of spheroidal self-assemblies.

Additionally, we investigate the shape of each spheroidal self-assembly by calculating its radii of gyration about the three principal axes using the *g_gyrate* function in GROMACS. From those radii, the semi-axis lengths of an ellipsoid with the approximate shape of the self-assembly were computed [34]. The relative shape anisotropy [35] which ranges from 0 for a sphere to 1 for a line was also computed from the radii. Table 2 lists the time averages (over 50 time frames) of the radii of gyration with their standard deviations, which represent how much the shape fluctuates dynamically. Note that since the spheroid in VC-C4/500 system exists for only a very short period of time, its radii of gyration are not averaged. The semi-axis lengths and relative shape anisotropies are also included in the table. The semi-axis lengths and relative shape anisotropies indicate that the shape of each assembly is slightly aspherical. The assemblies become more spherical as the alkyl chain length increases which is likely due to increased conformational flexibility, although the VC-C22/338 assembly is also significantly larger than the other assemblies.

Table 2. Radii of gyration (R_{gx} , R_{gy} , R_{gz}), semi-axis lengths (a , b , c), and relative shape anisotropies (κ^2) of spheroidal assemblies.

	VC-C4/500	VC-C13/500	VC-C22/338
Number of VECAR	25	22	35
R_{gx} (nm)	1.003	1.093 ± 0.027	1.373 ± 0.002
R_{gy} (nm)	1.165	1.169 ± 0.024	1.466 ± 0.003
R_{gz} (nm)	1.211	1.234 ± 0.036	1.519 ± 0.003
a (nm)	2.132	2.058	2.535
b (nm)	1.670	1.838	2.260
c (nm)	1.497	1.611	2.078
κ^2	0.012	0.005	0.003

3.2.2. MOLECULAR MORPHOLOGY: CROSS-SECTION OF SELF-ASSEMBLIES

Most of the self-assembled structures have shapes that are much more complicated than simple spheroids as can be seen in Figs. 3 and 4. For structures of such complex shapes, the radius of gyration does not provide enough information about the structural features of these structures. To gain additional insights, about the structure of these complex shapes, we also examined the size and shape of the cross-sections of the final self-assemblies. In VMD renderings of trajectories, we observe that the cross-sections of VC-C22 assemblies are either circular or elliptical while the cross-sections of VC-C13 and VC-C4 assemblies are consistently elliptical. As we explore the size and shape of each cross-section by fitting an encompassing ellipse over all atoms while in VMD renderings, we find that the following common numbers appear for a given molecular system as one of the semi-axis lengths: 3.0 ± 0.2 nm for VC-C22, 2.4 ± 0.2 nm for VC-C13, and 1.80 ± 0.01 nm for VC-C4. We note here that these rough estimates were made independently for various self-assembled structures.

To better understand these common characteristic length-scales measured from the cross-sections of many small and large assemblies in a given molecular system, we calculated the distance between carbon atoms with site IDs 39 and 15, which represents the end-to-end length of the hydrophobic chain (L_{CC}) for a given molecule. We also calculated the distance between hydrogens with site IDs 60 and 1, which represents the end-to-end length of a VECAR molecule (L_{HH}). The results are shown in Fig. 9. The probability density of the end-to-end lengths occurrence is defined as the number of molecules with a length L in a bin of width 0.025 nm divided by the total number of molecules in the system. Each probability density is calculated as a time average over 50 time frames of the last configurations in which all molecules form a single self-assembled structure. The VC-C13/500 and VC-C22/500 systems are not shown since they exhibit similar probability density profiles as their corresponding lower density systems. In Figs. 9(a)-(c), the common dimension found from the cross-sections of assemblies in each aqueous system is indicated with an arrow for visual aid; the common dimensions are comparable to the maximum occurrence for L_{HH} in each aqueous system. From this, we conclude that one of the semi-axes of the elliptical cross-section of all VECAR assemblies is limited by the most frequently occurring end-to-end length of the VECAR molecules, which increases as the hydrocarbon chain length increases. We note here that having a dimension of

the assemblies limited by the length of the hydrocarbon chain agrees with the findings of Fattal *et. al.* [36], in which it was found that the radius of cylindrical micelles is comparable to the chain length.

Figure 9. Probability density of the end-to-end lengths occurrence of the hydrocarbon chains (L_{CC}) and of the entire VECAR molecule (L_{HH}) are presented with dotted and solid lines, respectively, for (a) VC-C4/500, (b) VC-C13/403, (c) VC-C22/338, (d) bulk VC-C4, (e) bulk VC-C13, and (f) bulk VC-C22.

The profiles of probability density of the end-to-end lengths occurrence in system consisting of pure bulk VECAR are also shown in Figs. 9(d)-(f). By comparison, we find the following: First, the occurrence profile of L_{CC} does not change significantly between aqueous and bulk systems. This may be due to the fact that the hydrophobic chains reside within assemblies and thus their average length profile is not directly influenced by the presence of water even in aqueous systems. Second, the profile of L_{HH} occurrence, in all aqueous systems, clearly spreads farther into the smaller length region by about 0.5 nm compared to that of bulk systems. Third, only for the VC-C22 molecular systems, the maximum occurrence of L_{HH} shifts from 2.7 nm in the bulk system to 3.5 nm in the aqueous system. From the second and third observations, we infer that the hydrophilic heads residing at the surface of assemblies in aqueous systems might bend or orient themselves in such ways to generate the wider range of molecular length L_{HH} profile due to the hydrophilic interaction with surrounding water molecules at the interface in addition to the attractive and repulsive electrostatic interactions between their partial charges. In the following section, we continue to examine this conjecture.

3.2.3. STACKING OF AROMATIC RINGS

The two ends of a bolaamphiphilic VECAR molecule are composed of a strongly hydrophilic carnosine derivative and a weakly hydrophilic chromanol group. Each hydrophilic end has an aromatic ring in their structure, which henceforth will be denoted simply as the imidazole (IDs 55-56-57-59-61) and the benzene ring (IDs 3-4-11-10-6-7). Based on the chemical structure, the π - π stacking among these aromatic rings is expected to play an important role in the assembly process of VECAR molecules. This seems to be true for benzene rings in the VECAR self-assemblies of the present study.

Figure 10 presents the pair distribution profiles for the imidazole-imidazole and the benzene-benzene rings in the VC-C22/338 self-assemblies. Each pair distribution is calculated with the geometrical centers of each ring, and is averaged over approximately 150 time frames of the configurations in which all VECAR molecules form a single self-assembly. For the imidazole pair distribution profile, in addition to the sharp first peak at 0.5 nm, there is a broad second peak at 0.84 nm. The pair distribution profiles of the hydrophilic imidazoles remain the same in both bulk and aqueous systems, which we relate to the associations with no particular ordering of imidazoles observed from the VMD rendering of the trajectories. This general association of the imidazoles with each other is likely controlled by hydrogen bonding which is stronger than any stacking interactions.

For benzene ring pair distribution profile of the aqueous VC-C22 system, in addition to the sharp first peak at 0.44 nm and the broad second peak at 0.64 nm, there is a distinct third peak at 0.86 nm. We note here that the distinct peaks in the pair distribution profile of the benzene rings in the aqueous system, which becomes less pronounced in the bulk system profile, indicate that a long-range correlation among benzene rings exists only in the aqueous system. This reduction in the peak height which is particularly large for the first peak agrees well with the visual observations made with VMD; VMD renderings of the trajectory reveal that the benzene rings stack in aqueous systems, while those in bulk systems associate less frequently and only in pairs with no particularly preferred ordering. These distinctly different behaviors are likely due to the intricate balance of the hydrophilic and hydrophobic interactions of chromanols with water. While the hydroxyl group of chromanols is hydrophilic and tends to pull them toward the surface of the self-assemblies, the benzene rings are hydrophobic and are apt to stack in order to minimize the contact with water.

To gain additional insights into the aromatic rings ordering we calculated the percentage of the number of each type of aromatic rings participating in stacking or general association out of the total number of the rings. Our findings indicate that a lower percentage of imidazoles participate in associations in the aqueous system (65.7%) than in the bulk system (77.6%). This reduction of association of imidazoles in the aqueous system can be attributed to the strong hydrophilicity of imidazoles, which can drive some imidazoles away from the associations with other imidazoles into their association with water. On the other hand, even with the hydrophilicity of chromanols, there is a much higher percentage of benzene rings associating in the aqueous system (26.3%) than in the bulk system (17.6%). This supports the previous findings about the stacking of benzene rings being promoted by having them exposed to water at the interface.

As shown in Fig. 11, in our simulations we find mostly face-to-face or parallel-displaced stacking of two to four benzene rings in the VC-C22 aqueous systems. However, no stacks of five or more benzene rings are found. It is worth exploring potential reasons behind this limit of the stack size. To clearly show the location of the stack of four benzene rings relative to the whole body of the VC-C22 self-assembly, we include a schematic drawing in Fig. 11. In the stack of four benzene rings shown in Fig. 11(c), the geometrical centers of the rings fall on a sphere with a negative radius of curvature of about 89 nm. While this negative curvature may prohibit a larger number of rings from stacking in the present study of the early stages of the assembly process, it may cause a spiral or twisted macromolecular structure in the aqueous

VECAR molecular systems to form at much larger size and time scale as was found in other studies of bolaamphiphile assemblies [37]. We note here that the stacking of benzene rings is also observed in VC-C4 and VC-C13 aqueous systems but having only two or three rings per stacking sequence as opposed to four rings per stacking sequence found in VC-C22.

Figure 10. Pair distributions of rings: (a) imidazole-imidazole rings and (b) benzene-benzene rings. Solid and dotted curves represent data for bulk and aqueous systems of VC-C22, respectively.

Figure 11. Snap shots and schematic drawing depicting the stacking of the benzene rings formed in the VC-C22 aqueous systems: (a) a top view of two rings in face-to-face stacking, (b) a side view of three rings in parallel-displaced packing, and (c) a side view of four rings in face-to-face stacking with negative curvature as depicted in the schematic drawing.

4. CONCLUDING SUMMARY AND OUTLOOK

Aqueous solutions of novel asymmetric bolaamphiphilic molecules, VECAR, have been studied using MD simulations. Starting from VECAR randomly placed in a box and then hydrated, the evolution of each system has been observed. Initially, there is rapid formation of many small aggregates, and those eventually fuse into one, final aggregate. We find that, as the chain length increases: (1) the aggregation occurs at a slower pace, (2) the final assembly shape is less branched, (3) well-developed micelle structure is formed for small assemblies if the chain length is long enough (in VC-C22 and VC-C13, but not for VC-C4 systems), (4) the elliptical cross-section of assemblies becomes more spherical, (5) the mass density of assemblies decreases, and (6) the common value found to limit the size of one the semi-axes of the elliptical cross-section of assemblies increases. We also find stacking of the benzene rings of the chromanol groups in aqueous systems, while no ordered stacking is found between the aromatic rings in the imidazole groups in aqueous systems. Ordered stacking of either type of aromatic rings is not observed in the bulk VECAR systems where there is no water. We also find that the benzene rings form parallel or parallel-displaced stacking of as many as four rings. We believe that this limit of stacking size may be due to the short time of the simulations and is likely also a function of the assembly size. In conclusion, the kinetic and structural characteristics of the self-assembly process found in the present study demonstrate the distinctly different behavior of the chromanol and carnosine derivative head groups, and the effects of varying the length of the hydrophobic alkyl spacer. These findings are a valuable new addition to the current understanding of the self-assemblies formed by asymmetric bolaamphiphiles.

One should be aware of the limitations of the atomistic simulations employed in the present work. Although we studied relatively large-scale (over 500,000 atoms) aqueous systems over a time scale of 150 to 300 ns, the results presented in this work characterize only the early stage of the self-assembly process. It remains to be seen if the structural and kinetic properties as found here continue to apply in the mesoscopic and macroscopic systems. Of course, atomistic simulations of such systems are many orders of magnitude more demanding in size and time than those of present study. We are currently undertaking the extension of the present molecular dynamics simulation studies using coarse-grained force field of VECAR molecules.

Finally, the small-size micellar assemblies of VC-C22 and VC-C13 found in the current study are expected to be stable at low concentration and thus can be used in a study of micelle-lipid interactions for their drug delivery applications. Efforts in that direction are currently being made using aggregates from the VC-C22/338 simulation.

ACKNOWLEDGMENT. Research reported in this publication was supported by the Louisiana BOR-RCS fund [grant number LEQSF(2012-15)-RD-A-19] (Kim), by the National Institute of General Medical Sciences of the National Institutes of Health [grant number P20GM103424] (Kim), and in part by the National Science Foundation [grant numbers EPSCoR LA-SiGMA EPS-1003897 and CBET-1067583] (Novak and Moldovan). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Computational resources were provided by the Louisiana Optical Network Initiative (<http://www.loni.org>).

REFERENCES

- [1] C. Tanford, *The Hydrophobic Effect – Formation of Micelles and Biological Membranes*, second ed., John Wiley and Sons, New York, 1980.
- [2] J.H. Jordan, B.C. Gibb, Molecular containers assembled through the hydrophobic effect, *Chem. Soc. Rev.* 44 (2015) 547-585.
- [3] T.A. Taton, C.A. Mirkin, R.L. Letsinger, Scanometric DNA array detection with nanoparticle probes, *Science* 289 (2000) 1757-1760.
- [4] E. Gazit, Mechanisms of amyloid fibril self-assembly and inhibition: Model short peptides as a key research tool, *FEBS J.* 272 (2005) 5971-5978.
- [5] J.N. Israelachvili, D.J. Mitchell, B.W. Ninham, Theory of self-assembly of hydrocarbon amphiphiles into micelles and bilayers, *J. Chem. Soc., Faraday Trans. II* 72 (1976) 1525-1568.
- [6] J.N. Israelachvili, S. Marcelja, R.G. Horn, Physical principles of membrane organization, *Quart. Rev. Biophys.* 13 (1980) 121-200.
- [7] J.N. Israelachvili, *Intermolecular and surface forces*, second ed., Academic Press, Amsterdam, 2006.
- [8] W.M. Gelbart, D. Roux, A. Ben-Shaul, *Modern ideas and problems in amphiphilic sciences*, Springer, Berlin, 1993.
- [9] H.-J. Butt, K. Graf, M. Kappl, *Physics and chemistry of interfaces*, second ed., Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2006.
- [10] M.M. Kozlov, D. Lichtenberg, D. Andelman, Shape of phospholipid/surfactant mixed micelles: cylinders or disks? Theoretical analysis, *J. Phys. Chem. B* 101 (1997) 6600-6606.
- [11] M. Corti, V. Degiorgio, *Physics of amphiphiles: micelles, vesicles and micromulsions*, North-Holland, Amsterdam, 1985.
- [12] D. Langevin, J. Meunier, N. Boccara, *Physics of Amphiphilic Layers*, Springer, Berlin, 1987.
- [13] W.M. Gelbart, A. Ben-Shaul, D. Roux, *Micelles, membranes, microemulsions, and monolayers*, Springer, New York, 1994.
- [14] S.A. Safran, *Statistical thermodynamics of surfaces, interfaces and membranes*, Addison Wesley, Reading, MA, 1994.
- [15] R. Lipowsky, E. Sackmann, *Structure and dynamics of membranes*, Elsevier, Amsterdam, 1994.
- [16] B.J. Chun, J.I. Choi, S.S. Jang, Molecular dynamics simulation study of sodium dodecyl sulfate micelle: water penetration and sodium dodecyl sulfate dissociation, *Colloids and Surfaces A* 474 (2015) 36-43.
- [17] C.E. Astete, D.S. Meador, D. Spivak, C. Sabliov, Synthesis of vitamin E-carnosine (VECAR): new antioxidant molecule with potential application in atherosclerosis, *Synthetic Commun.* 43 (2013) 1299-1313.
- [18] M. Sammalkorpi, M. Karttunen, M. Haataja, Structural properties of ionic detergent aggregates: a large-scale molecular dynamics study of sodium dodecyl sulfate, *J. Phys. Chem. B* 111 (2007) 11722-11733.
- [19] S. Tsonchev, K.L. Niece, G.C. Schatz, M.A. Ratner, S.I. Stupp, Phase diagram for assembly of biologically-active peptide amphiphiles, *J. Phys. Chem. B* 112 (2008) 441-447.
- [20] S. Tsonchev, A. Troisi, G.C. Schatz, M.A. Ratner, On the structure and stability of self-assembled zwitterionic peptide amphiphiles: A theoretical study, *Nano Lett.* 4 (2004) 427-431.
- [21] C.A. Dreiss, Wormlike micelles: where do we stand? Recent developments, linear rheology and scattering techniques, *Soft Matter*, 3 (2007) 956-970.
- [22] S. Manet, A. Cuvier, C. Valotteau, G.C. Fadda, J. Perez, E. Karakas, S. Abel, N. Baccile, Structure of bolaamphiphile sophorolipid micelles characterized with SAXS, SANS, and MD simulations, *J. Phys. Chem. B* 119 (2015) 13113-13133.
- [23] E. Lindahl, B. Hess, D. van der Spoel, GROMACS 3.0: a package for molecular simulation and trajectory analysis, *J. Mol. Model.* 7 (2001) 306-317.
- [24] H.J.C. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: A message-passing parallel molecular dynamics implementation, *Computer Phys. Commun.* 91 (1995) 43-56.
- [25] A.K. Malde, L. Zuo, M. Breeze, M. Stroet, D. Poger, P.C. Nair, C. Oostenbrink, A.E. Mark, An automated force field topology builder (ATB) and repository: version 1.0, *J. Chem. Theory Comput.* 7 (2011) 4026-4037.

- [26] C. Oostenbrink, T.A. Soares, N.F. van der Vegt, W.F. van Gunsteren, Validation of the 53A6 GROMOS force field, *Eur. Biophys. J.* 30 (2005) 273-284.
- [27] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, J. Hermans, Interaction models for water in relation to protein hydration, in: A. Pullman (Ed.), *Intermolecular Forces*, Springer, The Netherlands, 1981, pp. 331-342.
- [28] L. Martinez, R. Andrade, E.G. Birgin, J.M. Martinez, PACKMOL: A package for building initial configurations for molecular dynamics simulations, *J. Comput. Chem.* 30 (2009) 2157-2164.
- [29] R.W. Hockney, S.P. Goel, J. Eastwood, Quiet high-resolution computer models of a plasma, *J. Comp. Phys.* 14 (1974) 148-158.
- [30] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, *J. Chem. Phys.* 126 (2007) 014101.
- [31] H.J.C. Berendsen, J.P.M. Postma, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, *J. Chem. Phys.* 81 (1984) 3684-3690.
- [32] W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics, *J. Molec. Graphics* 14 (1996) 33-38.
- [33] B.L. Bhargava, M.L. Klein, Initial stages of aggregation in aqueous solutions of ionic liquids: molecular dynamics studies, *J. Phys. Chem. B* 113 (2009) 9499-9505.
- [34] J.A. Myers, *Handbook of equations for mass and area properties of various geometrical shapes*, vol. 7827, US Naval Ordnance Test Station, 1962.
- [35] H. Lee, J.R. Baker, R.G. Larson, Molecular dynamics studies of the size, shape, and internal structure of 0% and 90% acetylated fifth-generation polyamidoamine dendrimers in water and methanol, *J. Phys. Chem. B* 110 (2006) 4014-4019.
- [36] D.R. Fattal, D. Andelman, A. Ben-Shaul, The vesicle-micelle transition in mixed lipid-surfactant systems: a molecular model, *Langmuir* 11 (1995) 1154-1161.
- [37] X. Liu, T. Wang, M. Liu, Interfacial Assembly of a Series of Cinnamoyl-Containing Bolaamphiphiles: Spacer-controlled packing, photochemistry, and odd-even effect, *Langmuir* 28 (2012) 3474-3482.

Graphical Abstracts (5cm H x 13cm W)