Computational Models for the Formation of Protocell Structures

Linglan Edwards and Yun Peng
Department of Computer Science and Electrical Engineering
University of Maryland, Baltimore County
1000 Hilltop Circle
Baltimore, MD 21250
email: linglan@cs.umbc.edu ypeng@cs.umbc.edu

James A. Reggia
Department of Computer Science
University of Maryland
College Park, MD 20742
email: reggia@cs.umd.edu

Abstract

There have been various attempts to simulate the self-assembly process of lipid aggregates by computers. However, due to the computationally complex nature of the problem, previous simulations were often conducted with unrealistic simplifications of the molecules' morphology, intermolecular interactions, and the environment in which the lipid molecules interact. In this paper, we present a new computational model in which each lipid is simulated by a more realistic amphiphilic particle consisting of a hydrophilic head and a long hydrophobic tail. The intermolecular interactions are approximated by a set of simple forces reflecting physical and chemical properties of lipids, e.g., hydrophobicity and electrostatic forces, which are believed to be crucial for the formation of various aggregates. With a set of carefully selected parameters, this model is able to successfully simulate the formation of micelles in an aqueous environment and reversed micelle structures in an oil solvent from an initially randomly distributed set of lipid-like particles.

This model can be used to study, at the microscopic level, the self-assembly of different protocell structures in the evolutionary process and the impact of environmental conditions on the formation of these structures. It may be further generalized to simulate the formation of other, more complex structures of amphiphilic molecules such as monolayer and bilayer aggregates.

1 Introduction

From the quest for the origins of life emerges this question: How did the first living cell form? This question leads to many speculations, including the origin of DNA, RNA, proteins, and also the origin of protocell structures such as micelles and membranes. The generally accepted assumptions are: 1) basic elements found in living systems were available on the primordial Earth; 2) as a result of various kinds of energy and catalytic effects, the simple molecules in the atmosphere, hydrosphere and lithosphere reacted to form a wide variety of small organic compounds; 3) condensation and polymerization reactions resulted in the formation of compounds of higher molecular weight and polymeric products; and finally, 4) the selective interaction and association of these macromolecules resulted in the generation of a living cell [17]. Researchers have been trying to verify these assumptions, and different approaches have been taken toward this problem from several different perspectives. They include finding clues about the processes involved in the origin of the cell, identifying underlying principles, constructing laboratory models, and more recently, creating computer simulations.

The focus of our work is the development of computational models to simulate the formation of simple protocell structures from lipid-like amphiphilic particles. In this paper we present a new approach for computer modeling of the self-assembly process of micelles and reversed micelles.

1.1 The formation of protocell structures

The general hypothesis about the origin of life on Earth is that an abundance of the simplest molecules existed in the prebiological period; these simple molecules underwent chemical reactions to form more and more complex molecules, eventually forming all of the elemental components of the first cell [4]. In order to support this hypothesis, it is necessary to show convincing evidence that all of the required molecules, as well as the conditions necessary for them to evolve into more complicated forms, existed on the early Earth.

It is reasonable to assure that molecules on the primordial Earth were at least as complex as those commonly found in interstellar space. Finding molecules that are necessary to form living systems in interstellar space suggests that such molecules also existed on prebiological Earth. While six of the most abundant elements in the universe (H, C, N, O, S, P) are precisely those necessary to form living systems, a large number of organic and inorganic molecules have also been identified in different galaxies. Twelve of the most numerous ones "can be considered as the prebiological precursors of essentially all the biochemical compounds present in living systems (shown in Table 1)[17].

Table 1 about here.

Laboratory experiments have also shown that a mixture of these basic molecules can be coerced into forming more complex macromolecular organic compounds under the right combination of temperature, pressure and other environmental factors (ultraviolet light or electrical discharge from lightning, for instance, consistent with conditions believed to have been present on the prebiotic Earth). For example, Fox, *et. al* produced a proteinoid from thermal copolymerization of amino acids in 1958 [6], and Hargreaves, *et. al* conducted experiments to synthesize various lipids from glycerol, fatty acid, aldehyde and phosphate under conditions believed to be similar to that on the prebiotic Earth [9].

If life did come from a gradual chemical and biological evolution, it is conceivable that some mechanism created a closed compartment enclosing a complex mixture of simple (organic and inorganic) and macromolecular organic compounds, prior to the formation of the first living cell. It is generally agreed that structures similar to cell membranes provided such closed environments within which bio-chemical reactions crucial for life could take place. The major components of contemporary cell membranes are lipids, primarily in the form of phospholipids. Together with protein, they divide the intracellular substance from the external environment. Lipids are water-insoluble organic substances. They may be considered amphiphilic molecules that contain nonpolar and polar regions, where the term amphiphilic means that they incorporate both a hydrophobic tail and a hydrophilic head within a single lipid molecule. The attractive interaction between nonpolar molecules in water is referred to as a hydrophobic effect. The special properties of amphiphilic molecules provide them the ability to *self-associate* or *self-assemble* into small molecular aggregates such as monolayers, micelles, bilayers, vesicles and biological membranes (Figure 1).

Figure 1 about here.

Two opposing forces, arising from the hydrophobic properties of lipids, are believed to be the main contributing factors to the self-assembly of lipid aggregates. The hydrophobic nature of the lipid tails causes them to associate, while the hydrophilic nature of the lipid headgroups tends to force these molecules to remain in contact with water, and thus away from each other [12]. However, exactly how lipid molecules formed these protocell structures in a prebiotic environment is still unknown.

1.2 Previous work on computer simulation of lipid aggregates

With the rapid development of computers and computing technology, computer simulation of biological systems has become progressively more important in the study of prebiological evolution. The formation of the first living cells as a result of possibly hundreds of millions of years of biochemical evolution is impossible to completely simulate or reproduce in a biology lab. However, fairly realistic computer simulation of inter-molecular interactions and compound formation can still provide insights into prebiotic biochemical synthesis. Moreover, since a computer simulation can be built upon our understanding of biological systems and hypotheses, it provides persuasive support for the underlying biological theories. Also, a well-defined computational model can be used as a tool to investigate the properties and dynamics of more complex biological structures.

There have been a number of attempts by researchers to develop computational models relevant to the origins of life, e.g., the "spontaneous" emergence of self-replicating structures [18] [3]. Here we focus solely on those concerned with the formation of lipid structures. Each approached the problem differently, yielding different results, and very often these simulations are very computationally expensive (requiring days to months of computer time for a simulation). Some simulated the self-assembly of lipid aggregates, while others focused on investigating the properties of such aggregates. Some of the main results are reviewed below.

B. Smit *et al.* developed a computational model to simulate the dynamics of an oil-water-surfactant system [19]. This model works with two kinds of particles: oil (**o**) and water (**w**). A water molecule in the water region consists of one **w** particle, and an oil molecule in the oil region consists of one **o** particle. Each surfactant molecule is represented as a chain of 2 **w** particles followed by 5 **o** particles, each bound to its neighbor in the chain by a strong harmonic force. All surfactants are initially randomly distributed in the water region. The interaction between particles (**o**'s and **w**'s in water and oil regions, and in surfactants) is determined by the Lennard-Jones potential:

$$\Phi = 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^{6}] \tag{1}$$

The **o**-**o** and **w**-**w** interaction is truncated at a much larger distance (2.5σ) than that of **o**-**w** ($2^{1/6}\sigma$). As a result, interactions between the same types of particles are attractive (within the truncation distance) but almost completely repulsive between different types of particles. The simulation, using a Monte Carlo method, with 39,304 particles, yields formation of micelle-like structures in the water region and a monolayer along the water-oil interface between the two regions. However, the model fails to produce the reversed micelles in the oil region that one would have expected in such a system. Another drawback of this model is that it does not take into account the actual biological structure of the lipid molecules, and the use of **w** and **o** particles to form surfactants cannot accurately capture the interactions of lipid-like molecules with water and oil solvents. Also, the micelles the simulation generates are somewhat different from the ones seen in actual biological experiments. Due to the geometrical packing properties of lipid molecules, lipids tend to form micelle structures when they are cone shaped, i.e. when they are single-chained lipids (surfactants) with large head-group areas and relatively thin tails [12]. However, in these simulation experiments, the surfactants have large tails and small head-group areas, and thus should form inverted micelles.

J. M. Drouffe *et al.* developed a computational model that simulates three-dimensional particles which self-assemble to form membrane-like objects [5]. In this model, each particle, represented as a ball, is a combination of two lipid molecules with the two tails combined into a intermediate layer across the center of a ball. A set of interaction forces is used to represent the inter-particle interactions. Starting from a randomly distributed set of 1,962 particles, the simulation ends with a membrane-like structure representing a minimum energy state. The main purpose of this model is to analyze the large-scale universal properties of membranes and vesicles, rather than their realistic *formation*. From the point of formation of protocell structures, several assumptions on which this model is based are questionable. First, the basic particle structure used in this model is not an actual stable biological structure. Such particles could not be present in abundance in the Prebiotic Earth. Also this model neglects the effect of the shape and structure of individual lipid molecules, which is generally considered essential to the self-assembly of lipid aggregates. Another problem is that there is no biological evidence to suggest the existence of an "anisotropic" interaction, which is used in this model as the major interaction forcing the molecules to align in planar configurations.

H. Heller *et al.* performed a molecular dynamics simulation of a rectangular patch of a lipid bilayer structure (membrane) in a water environment [11]. The initial state of the system consists of a hand-constructed patch of a bilayer membrane with 200 molecules of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylcholine (each containing 134 atoms) and of 5483 water molecules covering the head groups on each side of the bilayer. The simulation was done at the atomic level, with a total number of about 38,000 atoms. The simulation represented a time span of 263 ps ($1 \times 10^{-12} seconds$) of the chemical dynamic process. However, adopting this approach (i.e., simulating the molecular dynamics at the atomic level) to the formation of protocell structures such as membranes is computationally intractable. This is because the simulation of the formation of a membrane (or a patch of it) is computationally several orders of magnitude more complex than the simulation of an already-formed membrane reaching thermal equilibrium (which itself took months of time on a supercomputer).

Finally, B. Mayer *et al.* introduced a different type of self-assembly simulation, the lattice molecular automaton (LMA) model [14]. As a cellular automata model, each cell, representing a particle, changes its state according to the state transition rules (representing the inter-particle interaction) and the current states of its hexagonal neighbors. This model is able to simulate formation of small polymer clusters of lipid-like pentamers in a polar environment. One advantage of this model is its low computational complexity. Unlike other models where interaction is computed for every pair of particles at each computing step, each cell in the LMA model only interacts with its six neighbors, ignoring potentially important larger range interactions. In order to simulate systems of larger scale, such as micelles and membranes, the level of description of automata would have to be redefined to make the computation more realistic.

2 A Molecule-Level Computational Model

A good computational model for simulating lipid aggregates should have the ability to simulate: 1) different types of protocell structures, 2) formation of structures under different environmental conditions (different pH levels, presence of salt, etc.), 3) fairly realistic and detailed lipid molecule structures and their interactions, and 4) protocell formation in a computationally efficient manner. Deciding what level of detail the model should represent is crucial for a good compromise among these conflicting objectives.

Based on these considerations, we adopted an approach different from the simulation models summarized in the previous section. Our model is centered on what is believed to be the most important property of lipid molecules responsible for the formation of lipid aggregates, namely, the different hydrophobicity of the lipid's head and tail. First, to better reflect this amphiphilic property, each lipid is modeled as a *structured* particle having a large head and a long, thin tail. Second, the inter-molecular interactions are approximated as a set of simple forces with heads and tails of particles playing different roles in defining these interactive forces. Each of these forces reflects some physical or chemical property of the lipid particles believed to be crucial to the formation of various aggregates. They, together with properly choosing parameters, capture in some detail the interactions occurring during actual biochemical process. These forces are then used to determine the linear and rotational movements of individual particles according to classical Newtonian mechanics.

For example, our approach differs from existing models such as that of Smit *et al.* ([19]) where molecular interactions were coarsely modeled by an energy function (Lennard-Jones potential) and an energy-minimizing approach was used to determine the motion of particles. Our approach is also different from the model of Drouffe *et. al* [5], where structurally unstable particles were used as the basic elements of simulations. While we do not attempt to model events at the level of individual atoms, we believe that Newtonian mechanics provides a reasonable approximation for the forces influencing the motion of the molecular particles in the computer simulation.

2.1 Basic model and the interacting forces

In an aqueous environment, hydrophobic effects drive phospholipid particles closer to each other, with their tails pointing inward, forming structures like micelles where water has been "squeezed out". In order to simulate hydrophobic effects, we define each particle in our model to be an amphiphilic structure having a polar hydrophilic head group and a non-polar hydrophobic tail. These lipid-like particles are the basic elements of our computational model. The head of a particle is defined as a sphere of radius r. The tail is represented as a thin, inflexible rod of length L (Figure 2).

Figure 2 about here.

Based on hydrophobicity and other physical and chemical properties of lipids, we define seven inter-particle interaction forces (Figure 3). These forces, described below from a top-level, intuitive viewpoint, exist between every individual pair of particles in the system. In the force definitions the particles are treated as structured rigid body rather than single points, and heads and tails play different roles in these definitions.

Figure 3 about here.

1. The hydrophobic effect is the major driving force in this model. To model hydrophobicity in the lipid tails and hydrophilicity in the lipid heads, we introduce the following group of forces. All forces in this group have relatively long range.

- f_1: head-head attraction. The heads of the particles are hydrophilic. Therefore they tend to move toward the direction of an aqueous environment. We interpret this as an attraction force between the heads of each pair of particles. The direction of f_1 is depicted in the figure, and the point of action is the center of the head.

- f_5: tail-tail attraction. The tails of particles are hydrophobic. Therefore the tails will also tend to move toward each other in an aqueous environment, forming a structure to create an environment that does not contain water. In other words, they will tend to squeeze water out of the compartment. We represent this effect as an attraction force between the tails of each pair of lipid particles. Strictly speaking, since tails are long rods, every part of a tail is attracted by every part of the other tail. To simplify the force computation, we treat this force as between the ends of the two tails, as depicted in the figure. An alternative, treating this force as between the centers of the two tails, did not work well in the computer simulations.

- f_3 and f_4: tail-head and head-tail repulsion. The tails and heads of the lipid have opposite hydrophobicity. Therefore the head of a particle tends to repel the tail of another particle, and vice versa. We represent this property as a repulsive force between heads and tails of different particles. Again, as depicted in the figure, these forces are between the center of the head and the end of the tail.

2. Forces based on electrostatic charge and hydrophilicity.

- f_2: head-head repulsion. It is known that the heads of lipids are electrically charged but tails are not. The electrostatic charge will prevent the heads from becoming too close to each other. Also, if two heads do become too close to each other, their hydrophilicity will tend to increase their distance, ensuring a water environment among them. These two interactions are captured by the repulsive force f_2. This force has relatively short range.

3. Forces from incompressibility of particles.

- f_6 and f_7: a pair of repulsion forces. These forces exist between two particles due to the incompressibility of the molecules when they are close to each other. In our current simulation model, these forces are simplified by only considering the endpoints (heads and the ends of tails) of the particles (Figure 3b). Specifically, consider a pair of particles i and j. If an endpoint of particle j is too close to i (measured by the minimum distance from that endpoint of j to any part of i), then the repulsive force f_6 takes effect. Force f_7 is the reacting force of f_6, applied on j by i. For each pair of particles there are four pairs of forces f_6 and f_7, one for each of the four endpoints. These forces have very short range.

In summary, for any particle i, its head will be subject to forces of f_1, f_2, and f_3 from particle j, while its tail will be subject to f_4 and f_5 from j. Moreover, when i and j are sufficiently close to each other, j may impose two f_6 type of forces on i, one from j's head, the other from j's tail. Conversely, i is subject to two f_7 forces from j, which are reactions of the f_6 forces that i's head and tail impose on j. Therefore, a total of nine forces will be applied to particle i from another particle j, although some of them may have negligible magnitudes when i and j are not close. These forces together provide a more detailed and comprehensive representation of the inter-particle interactions relevant to protocell development than previous models.

2.2 Computing the forces

The directions and the points of actions of forces f_1 through f_7 are shown in Figure 3. The magnitudes of the forces f_k depend on their types and their respective distances d_k between the two entities they acted on. For example, f_1 depends on the distance between the centers of the heads of two particles, while f_3 depends on the distance between the center of the head of particle i and the endpoint of the tail of particle j. The functional relationships between f_k and d_k are different for individual forces, and some of them are complicated. To simplify the computation, we adopt the following relationship between f_k and d_k for all forces involved. Force f_k maintains a constant maximum magnitude as long as d_k is within a given fixed range, and drops quickly toward its minimum value when d_k is beyond that range. To further simplify the computation, instead of using a smooth, differentiable function (e.g., the radial function used in [5] or the Lennard-Jones potential in [19]), ramp functions (see Figure 4) are used in this model to implement the functional relationships between f_k and d_k. A ramp function for force f_k can be specified by the following parameters:

- a_k: the maximum magnitude of f_k;
- b_k: the minimum magnitude of f_k;
- r_k: the radius of the range within which force f_k maintains the maximum magnitude (e.g., if $d_k \leq r_k$ then $f_k = a_k$);
- r'_k: the radius of the range beyond which force f_k maintains the minimum magnitude (e.g., if $d_k \geq r_k$ then $f_k = b_k$).

When d_k is in between r_k and r'_k, f_k decreases linearly as d_k increases.

Figure 4 about here.

Two types of ramp functions are used in the simulation. The first has $r'_k = 2r_k$, and thus allows a relatively slow reduction of f_k as d_k exceeds r_k. This type of ramp function is used for the long-range forces f_1, f_3, f_4, and f_5. The second type of ramp function has $r'_k = r_k + \epsilon$, where $\epsilon \ll r_k$ is a constant, and thus causes a more abrupt reduction of f_k as d_k exceeds r_k. This type of ramp function is used for the short-range forces. The magnitude of each force f_k is then determined by the distance d_k and the three parameters r_k, a_k, and b_k that define its ramp function. Given distance d_k, the force magnitude f_k is then calculated via the function

$$f_k = ramp(a_k, b_k, r_k, d_k). \tag{2}$$

Each force produces a torque relative to the center of mass of the particle by the function

$$\tau_k = f_k * \sigma_k * sin(\alpha_k), \tag{3}$$

where σ_k is the distance from the point of action of force f_k to the geometric center of the particle, and α_k is the angle between f_k and the tail of the particle.

2.3 Overall movement influences

Now we consider the movement of each particle when subjected to the interactive forces from all other particles. The linear movement of particle i is represented by the movement of the center of mass of i. This movement is determined by the compound force F_i on i, which is computed by

$$F_i(t) = \left(\sum_{j \neq i} \sum_k f_k(t) \right) + rand_1(t) - c_1 * v_i(t). \tag{4}$$

The first term in the compound force F_i is the vector summation of all interactive forces to i from all other particles j in the system, as defined above. The second term in F_i, $rand_1(t)$, is a small random force placed on particle i, representing thermal effects. The third term, $c_1 * v_i(t)$, is the "friction" (resistance to i's movement), where $v_i(t)$ is the current velocity of particle i, and $c_1 > 0$ is the friction coefficient.

Assuming each particle has a unit mass, and using Euler's method to approximate Newton's second law of motion, we then have,

$$\begin{aligned} \Delta v_i(t + \delta t) &= F_i(t) \cdot \delta t && (5) \\ \Delta S_i(t + \delta t) &= \frac{1}{2} F_i(t) \delta t^2 + v_i(t) \delta t && (6) \end{aligned}$$

where ΔS_i is the linear displacement for the time step δt.

The rotational movement of particle i can be determined similarly. The compound torque of particle i can be computed by

$$T_i(t) = \left(\sum_{j \neq i} \sum_{k} \tau_k(t) \right) + rand_2(t) - c_2 \omega_i(t) \tag{7}$$

where k ranges over all forces from j to i, and

- $\tau_k(t)$ is the torque generated by force f_k at time t,
- $rand_2$ is a small random torque,
- $\omega_i(t)$ is the current rotational velocity, and
- $c_2 > 0$ is rotational friction coefficient.

Assuming each particle to be an uniform thin rod of length $2r + l$, the rotational inertia I can be computed by

$$I = \frac{1}{12}(2r + l)^2. \tag{8}$$

The rotational movement will then be

$$\Delta\omega(t + \delta t) = \frac{T_i(t)}{I} \delta t, \tag{9}$$

$$\Delta\theta_i(t + \delta t) = \frac{1}{2} \frac{T_i(t)}{I} \delta t^2 + \omega_i(t) \delta t \tag{10}$$

where $\Delta\theta_i$ is the angular displacement for the time step δt.

3 Implementation and Simulation Results

Based on the computational model described in the previous section, we have built a system to simulate the interactions between the lipid particles in an aqueous environment, and used it to examine the formation of micelle structures from free-floating lipid molecules. A major difficulty with the computer simulation, as with the previous models developed by others that we summarized earlier, is the computational complexity involved. As discussed in the previous section, every particle is subject to nine forces in its interaction with every other particle. Therefore, to determine one-step movements of all particles during a single iteration for a system of N particles, we need to conduct $4.5 * N(N-1)$ individual force computations. Moreover, hundreds of thousands of simulation steps (iterations) are needed for the system to reach a relatively stable state when N is at the level of a few hundred. When N is too small, the system would not have sufficient lipid concentration to yield satisfactory results. We made various efforts to reduce the computation. In developing the computational model, we use the endpoint of a tail rather than the entire tail rod to define forces f_3, f_4 and f_5. Also, instead of smooth functions such as a radial function, we use the computationally simple ramp functions to calculate the magnitudes of interactive forces. To further reduce the computational complexity, we restrict the simulations to two spatial dimensions, and calculate inter-particle interactions only for those particle pairs whose distance is within a certain range.

Another major practical issue is the adjustment and final selection of the parameters a_k, b_k and r_k, which define the ramp function for computing the force magnitudes. Qualitatively, f_1 and f_4, the head-head and tail-tail attractions, are responsible for pulling particles toward each other; the head-head repulsion f_2 ensures that the distances between heads are greater than those between tails; and f_3 and f_4, the tail-head and head-tail repulsions, move tails away from heads, and thus align the particles in such as way that their heads are in one direction and tails in another direction. Superficially, it appears conceivable that these forces would be able to pull particles together and form a circle-like structure where the tails of these particles all point to the center of the circle. However, the actual movement of the particles, and thus the evolution of the system, depends on the magnitudes of these forces, defined by the parameters a_k, b_k and r_k. If these parameters are not selected properly, micelle-like structures do not form. For example, when f_3 and f_4 are too small, the circumferences of circles formed by the particles become zigzag. Only when we increase these forces, circles with smooth circumferences were formed. We also found in the experiment that "friction" or resistance to motion for both linear and rotational movements ($c_1 v_i(t)$ and $c_2 \omega_i(t)$ in Eqs. 4 and 7.) are important for the system to reach a stable state.

The results shown in this section are simulations of a system in an aqueous environment. The size of the environment is 900 $\times$ 900 (normalized) spatial units. Each lipid particle has a sphere head of radius 5 units, and a tail of length 30 units. There are 200 lipid-like particles randomly distributed in the initial state of the system. Two representative simulations are reported here. They started from different randomized initial states, and use slightly different parameter sets. The parameter set of Simulation 1 is given in Table 2.

Table 2 about here.

A series of snapshots for different stages of Simulation 1 is shown in Figure 5. After 260,000 iterations, the system reaches a relatively stable state in which all lipids stop move. Micelle-like structures can be seen forming eventually during the process.

Figure 5 about here.

Simulation 2 uses the same set of parameters, except that the ranges of repulsive forces f_3 and f_4 are reduced 30% from those used in Simulation 1. As the result, particles in a larger neighborhood can be drawn together to form larger micelles. The snapshots at stages in Simulation 2 are given in Figure 6. It can be observed that larger sized micelle structures are indeed formed in Simulation 2.

Figure 6 about here.

Simulation 1 was performed on an *SGI R4400* platform. The total running time for 260,000 iterations was about 76 hours. Simulation 2 was performed on an *SGI R8000* platform, with a total running time of about 57.5 hours for the same number of iterations. Actual simulations can be viewed more clearly using an interactive animated display software we developed for this task. This software reads the data file containing the locations and orientations of all particles at every simulation time step and graphically displays the movement of lipid particles. Moreover, the user is allowed to change the display speed, to pause and resume the display at any time.

4 Formation of Reversed Micelles

Reversed micelles, which can be formed by lipids in an oil environment, are also believed to exist in prebiotic Earth. It is interesting to see if our computational model for micelle formation can be modified to simulate the formation of reversed micelles. The environment we choose for this investigation is an oil environment that contains numerous small water drops and randomly distributed lipid particles. Intuitively, in such an environment, due to the hydrophobic nature of the tails and hydrophilic nature of the heads, the lipid particles, if close to a water drop, tend to have their heads moving into the drop but keeping the tails outside the drop, thus forming a reversed micelle.

To simulate this interaction, we made the following modifications to the model developed in the previous sections. Note that the opposing hydrophobicity of lipid heads and tails stems from the interaction between the lipids and water molecules. Therefore, hydrophobicity, which plays the central role in forming micelles in aqueous environment, becomes less obvious in the oil environment except when the lipid particles are sufficiently close to the water drops. In other words, the interaction between lipid particles and water molecules in nearby water drops, rather than between the lipids themselves, plays the dominant role in the formation of the reversed micelles. Consequently, all inter-particle forces modeling hydrophobicity in the previous model are eliminated or reduced. These forces include f_1 for head-head attraction and f_5 for tail-tail attraction, and f_3 and f_4 for head-tail repulsion.

Two new forces are introduced to model the interaction between the heads and tails of lipid particles and the water molecules in the nearby water drops (Figure 7):

- Force f_8: tail to water repulsion. When a lipid tail is sufficiently close to a water drop, the hydrophobic effect will drive the tail away from water. A repulsive force from the water drop, acting on the tail of particle, is defined to approximate this property. This force is of very short range. The ramp function used for this force is the same as the ramp function shown in Figure 4. It has $a_k = 40, b_k = 0$, $r_8 = R + L, r'_8 = R + L + 2r$, where R is the radius of the water drop, L is the length of the lipid tail, r is the radius of lipid heads.

- Force f_9: head to water attraction. Due to its hydrophilicity, a lipid head will tend to move toward a close water drop. We model this property by an attractive force between the water drop and the lipid head.

Force f_9 is of relatively long range. It will move a lipid particle's head toward the closest water drop until the head of a lipid is completely submerged in the water. At that time, f_9 on that particle is reduced to zero. Also note that the head-to-water attraction is an electrostatic force, which comes from the net surface charge formed by the alignment of water molecule dipoles at the surface. The water drop thus can be assumed to induce a charge proportional to its own surface area (in the two dimensional case, its circumference). Whenever a lipid head is drawn into the water drop, the electric field lines close between the water dipoles and the molecule head, reducing the total surface charge of the water drop by a constant amount. That is, when a reversed lipid micelle forms around the water drop, the micelle "shields" the outside environment from this attraction force of the water drop. In other words, the ability of a water drop to attract lipid heads diminishes when more and more lipid heads are drawn into it. This effect is incorporated into f_9 with its magnitude set to be proportional to $2\pi R - 2rm$ where m is the number of lipids submerged in the water drop.

A test simulation was performed on an *SGI R10000* platform. It simulated a system of 16 water drops and 200 lipid particles. In this simulation, we increased the time-step by 10 fold. The total running time was about 2.5 hours when a relatively stable state was reached after $100,000$ steps. A series of snapshots for this simulation is shown in Figure 8, where reversed micelles are formed around water drops.

5 Discussion

In this paper we presented a new computational model for simulations of the self-assembly of lipid-like particles into well-formed micelle-like protocell structures in an aqueous environment. By adopting a structural particle model and using a set of simple forces as an approximation of the inter-molecular interactions between lipid molecules, this model provides a reasonable reflection of protocell and reversed micelle formation while remaining reasonably tractable computationality. The effectiveness and efficiency of this model was demonstrated both by successful simulations of the formation of such as micelles from a pool of randomly distributed lipid-like particles, and the formation of reversed micelle structures in an oil environment containing small water drops.

While there are very promising results, much work remains to be done. Our next objective is to simulate the formation of monolayer structures in an oil-water interface environment, and develop computer programs for a three-dimensional simulation. This model may also be used to study, at the microscopic level, the self-assembly of different protocell structures in the evolutionary process and the impact of environmental conditions on the formations of these structures. It may also be useful as a tool to investigate the effects of changes in the environment (such as pH values and temperature in the solvent, the density of the lipids, etc.) on the formation of lipid aggregates, as well as the dynamic (fluid-like) property of those aggregates. This model may be further generalized to simulate the formation of other, more complex structures of amphiphilic components. Along this line, we are currently exploring the possibility of further extending this model to simulate the formation of membrane-like bilayer structures. We hope that further progress of this research will provide substantial insight into the process of protocell structure formations, and provide a useful tool for such studies.

Acknowledgment: This work was supported in part by NASA award NAGW-2805.

References

[1] P. A. Bachmann, P. L. Luisi, and J. Lang. Autocatalytic self-replicating micelles as models for prebiotic structures. *Nature*, 357:57–59, 5 1992.

[2] P. A. Bachmann, P. Walde, P. L. Luisi, and J. Lang. Self-replicating micelles: Aqueous micelles and enzymatically driven reactions in reverse micelles. *J. Am. Chem. Soc.*, 113(22):8204–8209, 1991.

[3] H. Chou and J. Reggia. Emergence of self-replicating structures in a cellular automata space. *Physica D*, 110:252–276, 1998.

[4] D. W. Deamer and J. Oro. Role of lipids in prebiotic structures. *BioSystems*, 12:167–175, 1980.

[5] J. M. Drouffe, A.C. Maggs, and S. Leibler. Computer simulations of self-assembled membranes. *Science*, 254:1353–1356, 12 1991.

[6] S. W. Fox and K. Harada. Thermal copolymerization of amino acids to a product resembling protein. In David W. Deamer and Gail R. Fleischaker, editors, *Origins of Life, The Central Concepts*, page 189. Jones and Bartlett Publishers, Boston, 1994.

[7] J.M. Haile. *Molecular Dynamics Simulation — Elementary Methods*. John Wiley & Sons, Inc., 1992.

[8] W. R. Hargreaves and D. W. Deamer. Origin and early evolution of bilayer membranes. In D. W. Deamer, editor, *Light Transducing Membranes, Structure, Function and Evolution*, 23–59. Academic Press, Inc., New York, 1978.

[9] W. R. Hargreaves, S. J. Mulvihill, and D. W. Deamer. Synthesis of phospholipid and membrane in prebiotic conditions. In David W. Deamer and Gail R. Fleischaker, editors, *Origins of Life, The Central Concepts*, 205–207. Jones and Bartlett Publishers, Boston, 1994.

[10] R. Harrison and G. G. Lunt. *Biological Membranes, Their Structure and Function*. Blackie, second edition, 1980.

[11] H. Heller, M. Schaefer, and K. Schulten. Molecular dynamics simulation of a bilayer of 200 lipids in the gel and in the liquid-crystal phases. *J. of Phys. Chem.*, 97(31):8343–8360, 1993.

[12] J. N. Israelachvili. *Intermolecular and Surface Forces*. Academic Press, 1991.

[13] M. K. Jain. *Introduction to Biological Membranes*. John Wiley & Sons, 1988.

[14] B. Mayer, G. Köhler, and S. Rasmussen. Simulation and dynamics of entropy-driven, molecular self-assembly processes. *Physical Review E*, 55(4):4489–4499, 4 1997.

[15] J. Oró and A. Lazcano. A minimal living system and the origin of a protocell. *Adv. Space Res.*, 4(12):167–176, 1984.

[16] J. Oró, S. L. Miller, and A. Lazcano. The origin and early evolution of life on earth. *Annu. Rev. Earth Planet. Sci.*, 1990.

[17] J. Oró, E. Sherwood, J. Eichberg, and D. Epps. Formation of phospholipids under primitive earth conditions and the role of membranes in prebiological evolution. In D. W. Deamer, editor, *Light Transducing Membranes, structure, function and evolution*, 1–21. Academic Press, Inc., New York, 1978.

[18] A. Pargellis. The evolution of self-replicating computer organisms. *Physica D*, 98:111–127, 1996.

[19] B. Smit, P.A.J. Hilbers, K. Esselink, L.A.M. Rupert, N.M. van Os, and A.G. Schlijper. Computer simulations of a water/oil interface in the presence of micelles. *Nature*, 348:624–625, 12 1990.

List of Figures

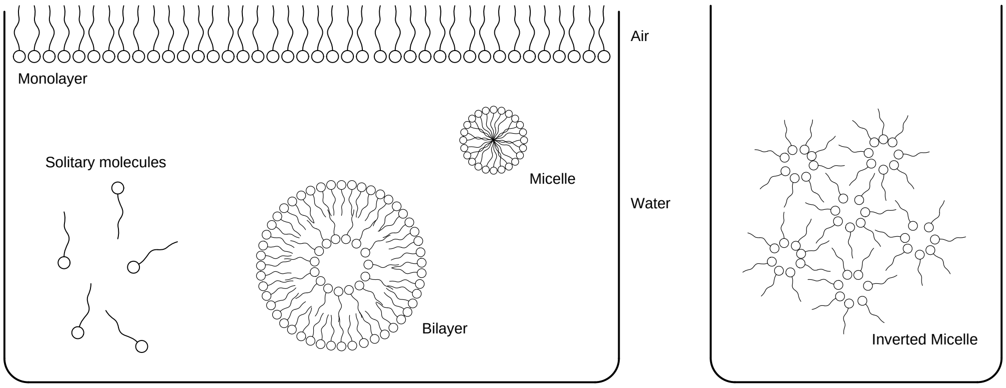
Interstellar Molecule	Formula	Biochemical Monomers and Properties
1. Hydrogen	H_2	Reducing Agent, Protonation
2. Water	H_2O	Universal Solvent, Hydroxylation
3. Ammonia	NH_3	Base Catalysis, Amination
4. Carbon Monoxide	$CO(H_2)$	Hydrocarbons and Fatty Acids
5. Formaldehyde	CH_2O	Monosaccharides (ribose) and Glycerol
6. Acetaldehyde	CH_3CHO	Deoxypentoses (deoxyribose)
7. Aldehydes (HCN)	$RCHO$	Amino Acids
8. Thioformaldehyde	CH_2S	Cysteine and Methionine
9. Hydrogen Cyanide	HCN	Purines (adenine, guanine) and Amino Acids
10. Cyanacetylene	HC_3N	Pyrimidines (cytosine, uracil, thymine)
11. Cyanamide	H_2NCN	Polypeptides, Polynucleotides and Lipids
12. Phosphine (Jupiter)	PH_3	Phosphates and Polyphosphates

Table 1: Biochemical monomers and properties which can be derived from interstellar molecules [17].

Force	Description	Magnitude a_k	Range r_k	Residual b_k
f_1	head-head attraction	40	15	0.4
f_2	head-head repulsion	30	30	0
f_3	tail-head repulsion	30	30	0.6
f_4	head-tail repulsion	30	30	0.6
f_5	tail-tail attraction	40	40	0.4
f_6	(active) repulsion due to incompressibility	1600	15	0
f_7	(reactive) repulsion due to incompressibility	1600	15	0

Table 2: Parameters used in simulation 1.

List of Figures



(a) Water (b) Oil

Figure 1: Formations of lipids in water and oil environments [13].

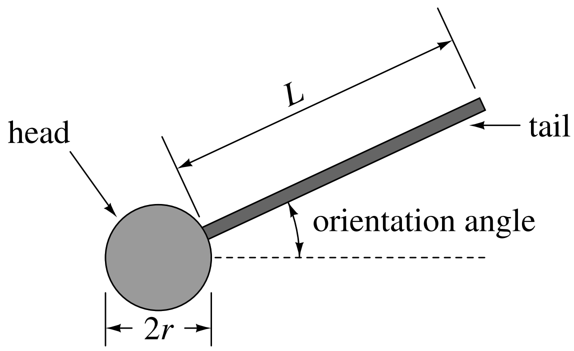
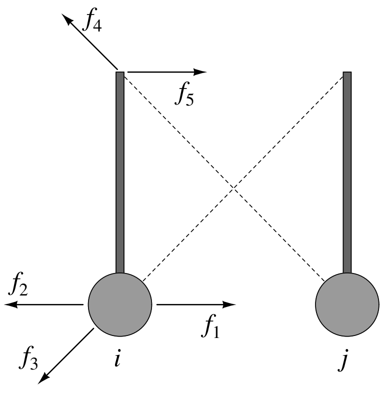
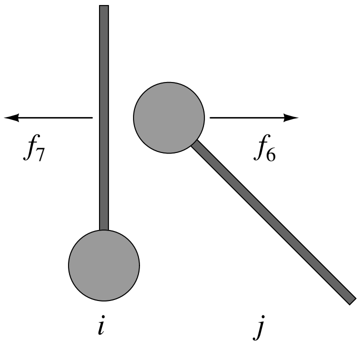


Figure 2: Lipid particle — The basic element in the simulation.



(a)



(b)

Figure 3: The interaction forces defined in the model.

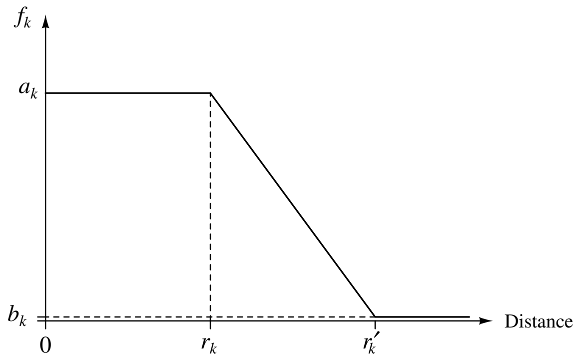
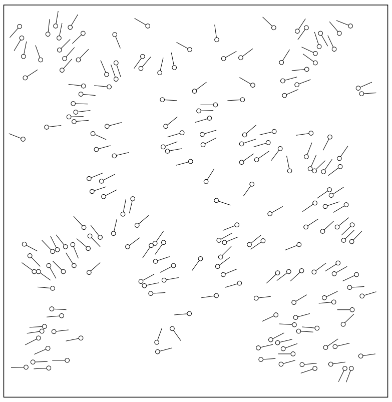
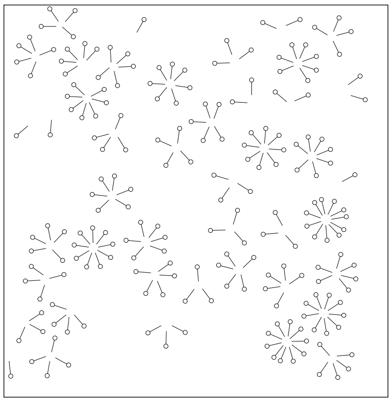


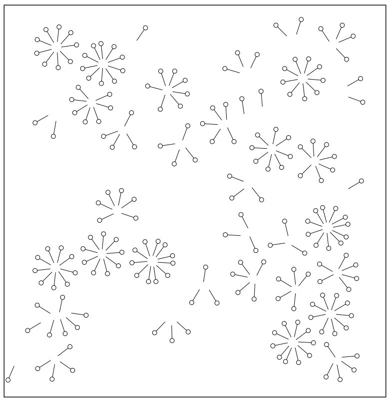
Figure 4: Ramp function for the magnitudes of all interacting forces.



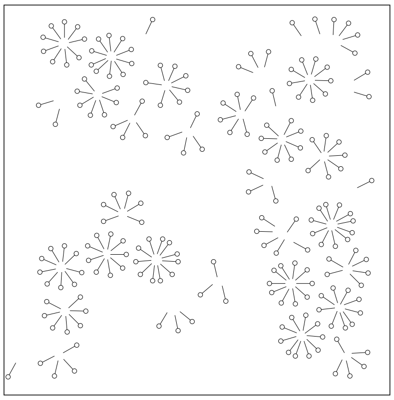
(a) initial state



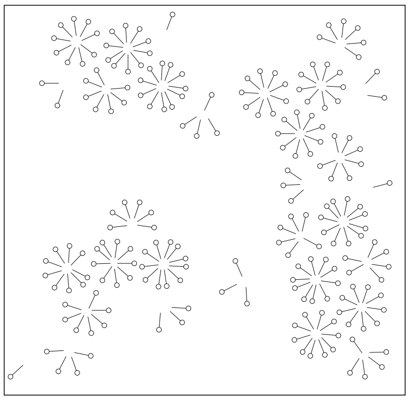
(b) after 20,000 iterations



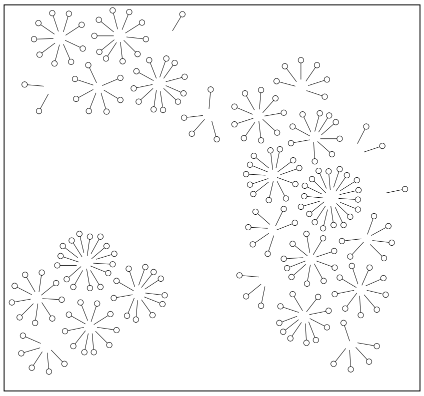
(c) after 40,000 iterations



(d) after 60,000 iterations

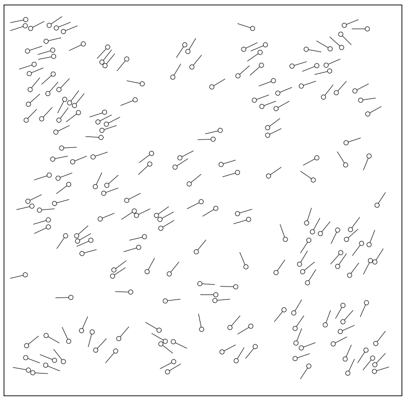


(e) after 100,000 iterations

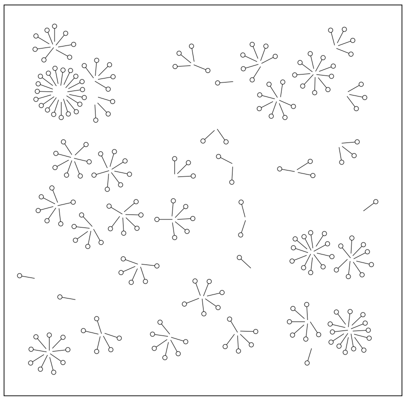


(f) after 260,000 iterations

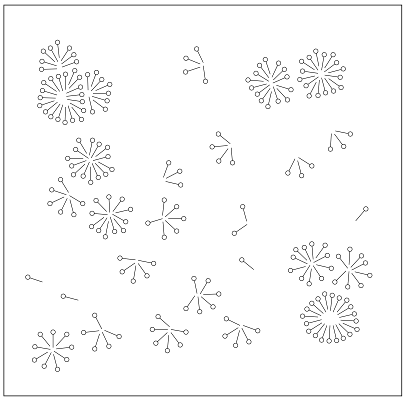
Figure 5: System of 200 lipid particles, Simulation 1.



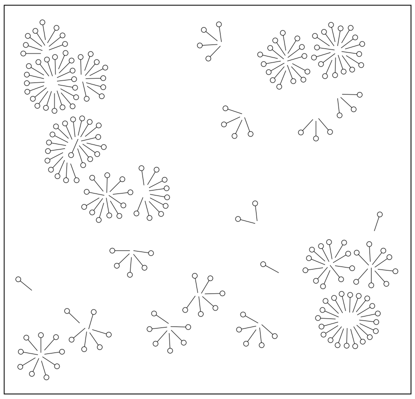
(a) initial state



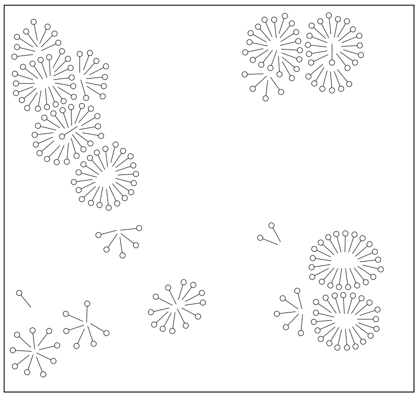
(b) after 20,000 iterations



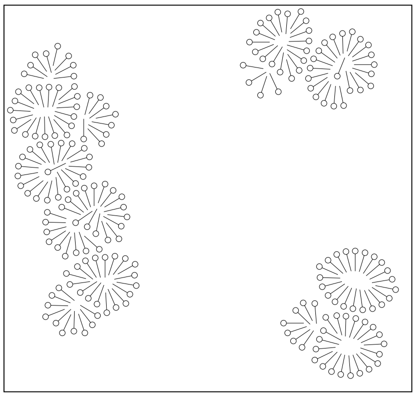
(c) after 40,000 iterations



(d) after 60,000 iterations



(e) after 100,000 iterations



(f) after 260,000 iterations

Figure 6: System of 200 lipid particles, Simulation 2.

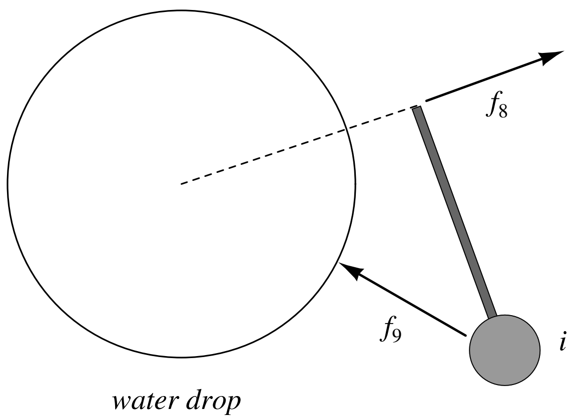
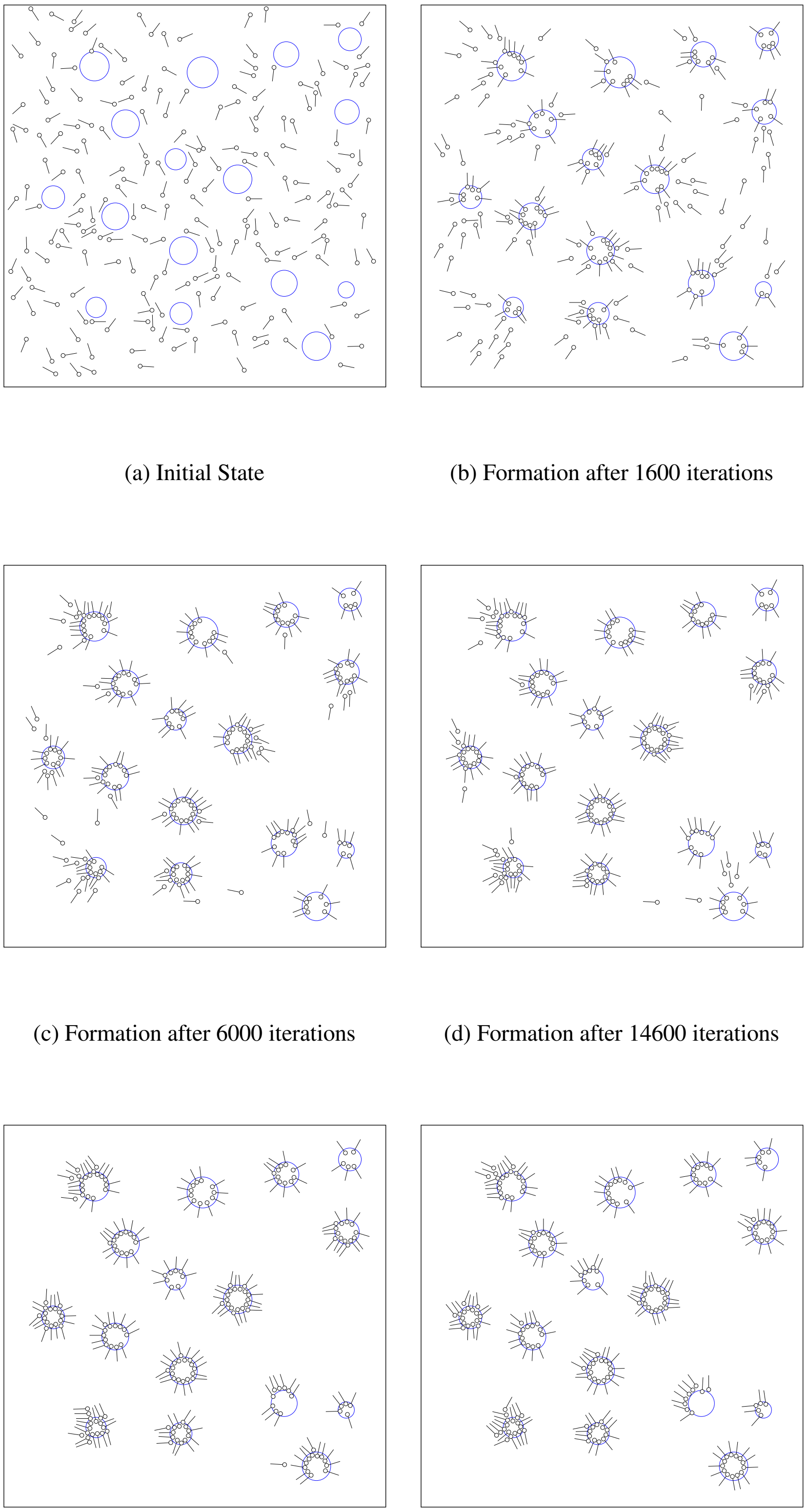


Figure 7: Interaction forces defined for reversed micelle simulation.



(a) Initial State

(b) Formation after 1600 iterations

(c) Formation after 6000 iterations

(d) Formation after 14600 iterations

(e) Formation after 32400 iterations

(f) Formation after 10000 iterations

Figure 8: System of 16 water drops, 200 lipid particles.