Computational Models for the Formation of Protocell Structures

Linglan Edwards

November 29, 1998

Abstract

The study of the formation of membrane structures has long been an active research area. The study of pre-biological and chemical evolution suggests that, prior to the appearance of the first living cell, there was a long period of chemical abiotic synthesis of organic compounds as a necessary precondition for the appearance of the first life forms. One of the critical steps of the pre-biotic evolution is the formation of proto-cell structures prior to the origin of life, which provide the closed environment for the bio-chemical reaction to take place. Phospholipids, due to the hydrophobicity of their tails and the hydropholicity of their charged head-groups, are the main components of cell-membranes. It was discovered that phospholipids can self-assemble to form structures like micelles, mono-layers, inverted micelles and bilayer vesicles. But the underline principles and the actual process for these different formations are still not fully understood. 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, as well as mono-layer structure in a air-water interface.

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 bilayer and multi-layer aggregates. It can also be used to study the membrane-protein interactions.

Keywords. Simulation, membrane, lipid, micelle, protocell, self-assembly.

Contents

List of Figures

List of Tables

Chapter 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 [48]. 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 of how the processes could take place, constructing laboratory models, and more recently, creating computer simulations.

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

The rest of the dissertation is organized as follows: Section 2 gives an introduction to the biological and chemical background which motivated this research, as well as the basic biological theories our computational models are based on. Section 3 describes existing work of several researchers attempting to develop effective computer simulations of molecular evolution, including the theoretical underpinnings of each model, the results obtained, and drawbacks of the model used. Section 4 states the goal of this dissertation research. Section 5 contains a detailed description of the approach we propose and the justification of this approach. Section 6 describes the detailed implementation for the computational model, using micelle as an example. Section 7 presents the extension of the first simulation model, the application to reverse micelle and mono-layer structure simulations, and a test on bilayer patch. Section 8 describes the integration of the simulation system. Section 9 concludes the dissertation research.

Chapter 2

Overview of Biological Background

For a long period of time since Darwin's "The Origin of Species" came out, a lot of biologists were not convinced. While Darwin' theory described the evolution of life as a process of "natural selection", explained how life have evolved over thousands of years and are still evolving, he also left a huge gap in the evolution process unexplained — how did the first life come to exist? Darwin's theory was based on the premise that life can only originate from life. But did life just appear spontaneously, or was it also a long, gradual product of evolution? Till now, a lot of effort had been devoted to explore the early evolution of life.

This leads to the following hypothesis about the origin of life on Earth:

2.0.1 Theory of prebiological-chemical evolution

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 [15]. 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.

Because the Earth is full of complex molecules formed from a variety of self-replication mechanisms, it is necessary to get away from the Earth in order to find what the composition of molecules might have been on its surface in the first billion years of its existence. Molecules on the primordial Earth can be assumed to be at least as complex as those commonly found in interstellar space. Similar evidence can be gleaned from the composition of the gaseous planets. Laboratory experiments show how plausible mixtures of these basic molecules can be coerced into forming complex organic molecules under the right combination of temperature, pressure, and other environmental factors (ultraviolet light or electrical discharge from lightning, for instance) consistent with the conditions which might have been present on the Earth during that period.

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 2.1)[48].

The discovery of microwave techniques in 1969 led to the discovery of interstellar ammonia, water, and formaldehyde, the first organic molecule to be discovered in space. Since then, more than 80 different molecules including a significant amount of organic

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 2.1: Biochemical monomers and properties which can be derived from interstellar molecules [48].

compounds have been detected in interstellar or circumstellar clouds.

Prebiological chemical evolution is based on the premise that life evolved progressively as a result of chemical processes which occurred on the primitive earth. There were probably at least three major steps in the synthesis and evolution of organic matter on the primitive earth before the emergence of life:

First, as a result of various kinds of energy and catalytic effects (electric discharges, ultraviolet light, heat, etc.), the simple molecules from the atmosphere, hydrosphere and lithosphere reacted to form a wide variety of low molecular weight organic compounds, including the biochemical monomers, amino acids, fatty acids, sugars and the purine and pyrimidine bases. It should be noted that these are the compounds which have been found in carbonaceous chondrites, and may also be present in localized environments within comets and other bodies of the solar system such as Europa, the satellite of Jupiter which was recently confirmed to have a subsurface layer of water. Second, with the accumulation of these biochemical monomers in the lakes and ponds of the primitive Earth's surface, further condensation and polymerization reactions occurred to yield higher molecular weight compounds and polymeric products, including oligomers, polymers and catalytic molecules. Finally, the selective interaction and association of these macromolecules and condensation products resulted in the genesis of a living cell. [47]

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 [25], 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 [29].

The study of the origin of life naturally leads to many speculations — the evolution of DNA, RNA, the formation and reproduction of cells. Here we focus our attention on how the first living cell formed, specifically, how the first cell membrane formed.

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.

2.0.2 The biological membrane and its structure

Clearly we cannot hope to understand how live began without knowing something about how living things work. In order to find out how the original cells have evolved, it is necessary to start from studying the structure of contemporary cells.

Cells are typically referred as “unit of life”, because all life forms have a cellular basis regardless of their nature. Furthermore, cell structures in different organisms are remarkably similar. There are two main types of cells among living organisms, prokaryotic cells and eukaryotic cells, each representing a different level of structural and functional complexity. Also, it has been shown that, cells from almost all organisms, including humans, can be culture outside the body, “i.e., in vitro, under conditions in which they will remain alive and prosperous, long after the organism from which they were removed has died.”[35]

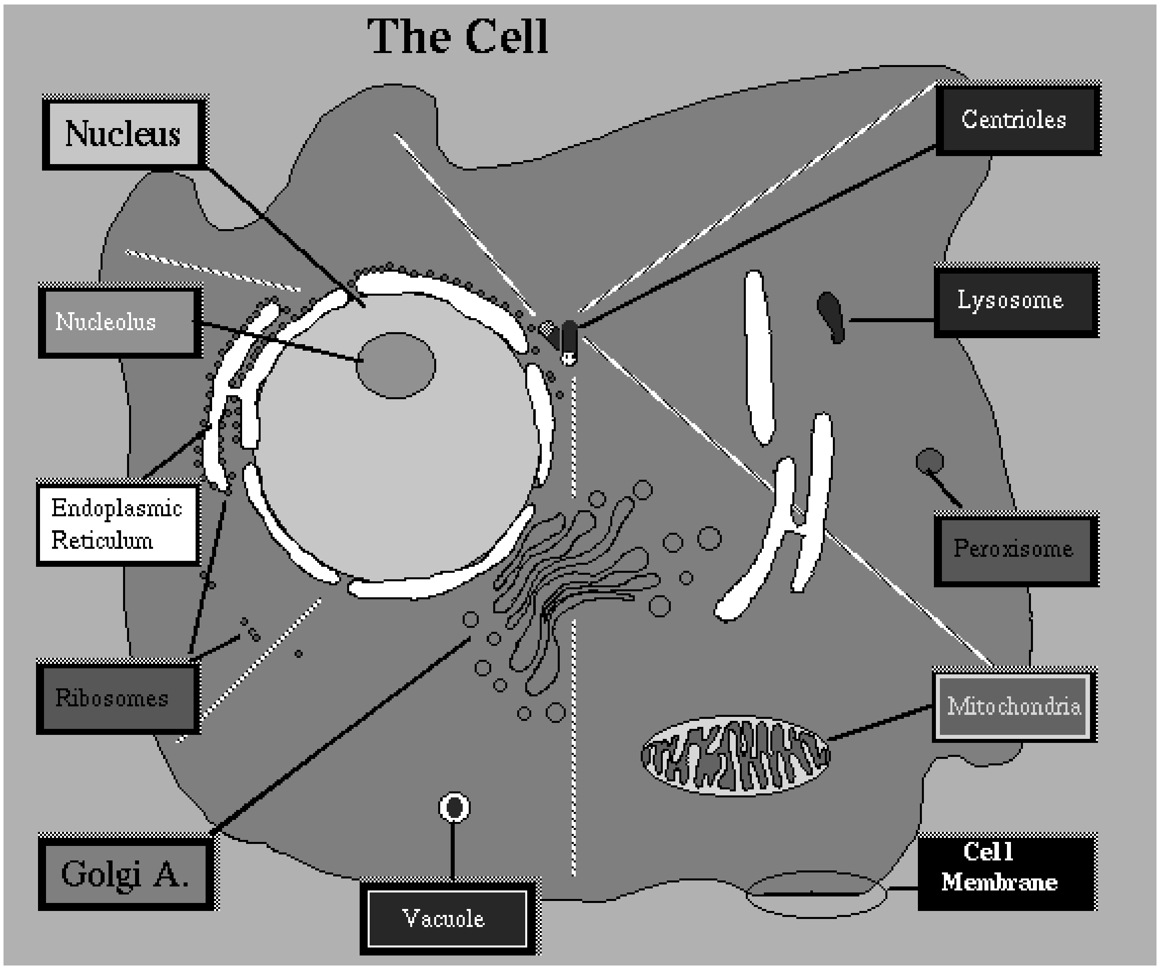


Figure 2.1: A Generic Cell Diagram.

Figure 2.1 shows a diagram of contemporary cell. Cells are 90% water. Of the remaining molecules present, the dry weight is approximately:

- 50% protein
- 15% carbohydrate
- 15% nucleic acid

- 10% lipid
- 10% miscellaneous

The cell is cytoplasm surrounded by cell-membranes. The cytoplasm is:

- Cytosol a lot of water - all except the organelles.
- Organelles(which also have membranes) in ‘higher’ eukaryote organisms:
  - Nucleus(in eukaryotes) — where genetic material (DNA) is located, RNA is transcribed.
  - Endoplasmic Reticulum (ER) — Important for protein synthesis. It is a transport network for molecules destined for specific modifications and locations. There are two types:
    - Rough ER — has ribosomes, and tends to be more in ‘sheets’.
    - Smooth ER — Does not have ribosomes and tends to be more of a tubular network.
  - Ribosomes — half are on the Endoplasmic Reticulum, the other half are ’free’ in the cytosol, this is where the RNA goes for translation into proteins.
  - Golgi Apparatus — important for glycosylation, secretion.
  - Lysosomes — Digestive sacks - the main point of digestion, these are only found in animal cells.
  - peroxisomes — Use oxygen to carry out catabolic reactions, in both plant and animals.

  - Microtubules — made from tubulin, and make up centrioles,cilia,etc.
  - Cytoskeleton — Microtubules, actin and intermediate filaments.
  - Mitochondria — convert foods into usable energy. (ATP production) A mitochondrion does this through aerobic respiration. They have 2 membranes, the inner membranes shapes differ between different types of cells, but they form projections called cristae. The mitochondrion is about the size of a bacteria, and it carries its own genetic material and ribosomes.

Since our research is based on the premise that, structures similar to cell membrane had formed as a precursor of first living cells to provide the fundamental environment that is needed for all the inter-cellular bio-chemical reaction to take place. So here we focus our attention to cell membranes.

Membranes serve many important functions: They separate a cell contents from its non-living surroundings, which allows the selective storage of metabolites and products, and the removal of waste products. The membrane surrounds the cell and controls chemical traffic into and out of the cell. It makes life possible by its ability to discriminate in its chemical exchanges with the environment. Genetic material may attach to membranes to facilitate distribution of genetic material to the cell's progeny during cell division. Membranes also create a unique structure which determines their function and solubility characteristics. [48] Gerald Karp summarized the major membrane functions as follows: [35]

- Compartmentalization Membranes are continuous unbroken sheets and as such inevitably enclose compartments. the plasma membrane encloses the contents of the entire cell. The nuclear and cytoplasmic membranes enclose various internal cellular spaces. Just as the space within a building needs to be divided so that different types of activities can proceed in their own compartments with a minimum of interference with one another, so too does the cell need to be divided. In the cell compartmentalization is particularly important since the various spaces are fluid filled, ad the indiscriminate intermixing of their contents would be very chaotic. One of the most limiting aspects of the prokaryotic condition is the relative lack of inter-cellular membranous organelles.

- Regulation of the movement of materials As just described, the plasma membrane serves as a wall or barrier between the living and non-living worlds, the intracelular membranes between the various cytoplasmic compartments. As barriers, membranes prevent the free interchange of materials from one side to the other, but at the same time membranes provide the means of communication between the spaces. Every cell requires nutrition, water, oxygen, ions, substrates, etc., from its environment, whether that environment is the blood stream of a multicellular organism or the medium in which a culture of single-celled of a multicellular organism or the medium in which a culture of single-celled organisms is growing. It is the responsibility of the plasma membrane to determine that the appropriate substances are allowed in and the inappropriate substances are kept out. In this capacity the plasma membrane can be considered as a *selectively permeable* wrapper.

  Among the substances that are allowed to enter the cell, some do so simply by diffusion in response to a concentration difference (gradient) across the membrane. In this regard the plasma membrane acts as a gate that can be opened or closed. Other substances are carried through the membrane in order to be maintained at a higher concentration on the inner side; the plasma membrane can act as a molecular pump. Still other materials, including the fluid in which they are suspended, can be brought inside by the formation of vesicles from the plasma membrane. One important consequence of the selective permeability of the plasma membrane is its capacity to separate ions and therefore establish an electrical potential difference across itself. This potential (voltage) is critical for the so-called excitable cells, the neurons and muscle cells, but may also be of great importance in the ability of all cells to respond to their environment.

The regulation of the passage of substances across membranes is not limited to movement from the outside of the cell to the inside by the plasma membrane. The various intra-cellular compartments must communicate with one another as well. No one organelle of the cell is autonomous; each depends upon the exchange of materials in both directions across the membrane by which it is bounded. The problem, in this connection, is that we know less about the various intra-cellular membranes than we do about the plasma membrane. The intra-cellular membranes are much less accessible for study and must be prepared from cells that have been broken apart and the cellular compartments separated.

- Access to information Another important aspect of the communication process mediated by membranes, particularly the plasma membrane, is the control of information received by and transmitted across from one side to the other. The plasma membrane, and possibly intra-cellular membranes as well, has receptor molecules that have the capacity to combine with specific molecules of the appropriate configuration, just as an enzyme combines with the appropriate substrate. Different types of cells have different types of receptors and, therefore, combine with different substances. Substances that combine with receptors are termed *ligands*. The best-studied ligands are hormones, growth factors, and neurotransmitters, all of which attach to the membrane but do not pass through it. The membrane to generate a new signal (e.g., cyclic AMP) directed to the internal compartment of the cell. In other words, the interaction between a membrane receptor and an external agent may cause a new informational stimulus to be produced by the membrane. In this way it appears that the plasma membrane has responsibility for the regulation of many different internal events. In some cases, the signal generated by the membrane tells the cell to divide or to differentiate into a new type of cell, while in the case of bacteria it may tell the cell to move toward food. Most of these processes are understood at a very superficial level or not at all; however in the case of a few hormones (particularly insulin and glucagon) quite a bit of the sequence of events has been elucidated.

- Inter-cellular interaction At the outer edge of the living cell, the plasma membrane is responsible for mediating the interactions between cells that occur continually in a multicellular organism. Organs are usually composed of many different types of cells that must acquire and maintain a certain orientation with one another and must work together to perform an overall function. The plasma membrane allows cells to recognize one another, to adhere when appropriate, and to exchange materials and information regardless of whether the cells are fixed in place, as in a tissue, or moving around individually within the plant or animal.

- Locus for biochemical activities Membranes provide a means to organize cellular activities. One of the consequences of a solution of reactants is that their relative positions cannot be stabilized; instead they remain at the mercy of random collisions. The association of enzymes catalizing successive reactions greatly facilitates the reaction sequence since each enzyme is automatically in the right place at the right time. Membranes provide the cell with a large-scale framework or scaffolding within which components can be ordered for effective interaction. Within a membrane many different processes are going on simultaneously, apparently without interference with one another. In the inner mitochondrial membrane, for example, the various components of the electron-transport chain and oxidative phosphorylation must work together in a coordinated manner. The organization of these proteins into respiratory assemblies within the structure of the membrane allows electrons to be passed from one carrier to the next in an orderly fashion. As a result, the energy of the electrons can be captured in the formation of ATP. If one considers all of the enzymatic machinery within a cell, much of it would be found associated with the various membranes.

- Energy transduction The conversion of one type of energy into another(transduction) is vital in many different ways, and membranes are intimately involved in these processes. Basic to all life is the ability of plant cells to convert light energy from the sun into the chemical energy contained in carbohydrates. Conversely, it is equally important for both plants and animals to convert the chemical energy of these nutrients into that of ATP and other "high energy" compounds. The machinery for energy capture is contained within the membranes of the mitochondria and chloroplasts. Light, thermal, and mechanical energy are converted by our sensory receptors into electrochemical energy for communication with the central nervous system. The underlying mechanisms involved in these transductions are not well understood, but membranes are known to be centrally involved.

- Miscellaneous Membranes act in a variety of miscellaneous capacities. For example, they provide mechanical strength, they are involved in cell movement and secretion, they act as electrical insulators, as sites for ribosome attachment, and as conductors of nervous impulses.

Figure 2.2 shows a diagram of a patch of typical cell membrane. While the composition of membranes varies with their sources, they generally contain about 40% of their dry weight as *lipids* and 60% as protein, held together in a complex by non-covalent interactions. Usually carbohydrate is present to the extent of $1 - 10\%$ of the total dry weight. This is covalently bounded either to lipids or to protein, with carbohydrate-containing molecules being considered lipids or proteins as appropriate. In addition to the above components, membranes contain some 20% of their total weight as water, which is tightly bound and essential to the maintenance of their structure. The density of isolated membranes is directly proportional to their protein content. [31]

The special properties of cell membranes are provided by their major component: lipids. In the next section, we discuss the structure and properties of lipid molecules.

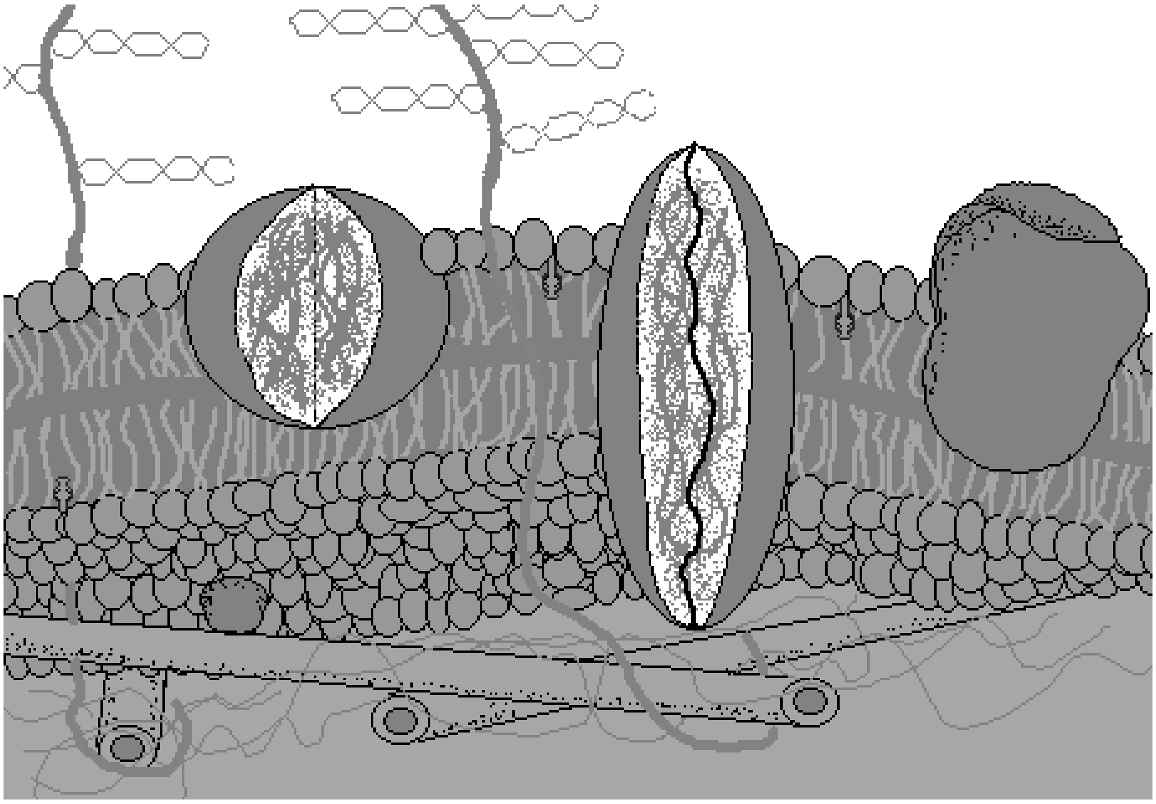


Figure 2.2: Biological membrane structure.

2.0.3 Lipids

The major components of contemporary cell membranes are lipids, primarily in the form of phospholipids. Together with protein, they divide the intra-cellular substance from the external environment.

Lipids are water-insoluble organic substances, defined operationally as actual or potential derivatives of fatty acids and their metabolites. They may be considered amphiphilic molecules that contain non-polar 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 non-polar portion of lipids in bio-membranes is mostly made up of polymethylene chains. Typically, most of the lipid components present in an organism are found predominantly in membranes. The possible number of structural and conformational isomers of lipids in membranes is very large.

There are 10 main types of different lipids in cell membranes. Different types of cell have different percentages of lipids, protein, and carbohydrates. The main types of lipids are Cholesterol, Glycolipids, Phosphatidylcholine, Sphingomyelin, Phosphatidylethnolamine, Phosphatydilinositol, Phosphatidylserine, Phosphatidylglycerol, Diphosphatidylglycerol (Cardiolipin), and Phosphatidic acid.

Phospholipids refer to those lipids that have phosphorus in their molecule structures, including the phosphatidyl choline, phosphatidyl ethanolamine, phosphatydilinositol, phosphatidyl serine, Phosphatidylglycerol, Diphosphatidylglycerol and phosphatidic acid just mentioned. Phospholipids are amphipathic molecules of very distinctive structure: they have a hydrophilic[1] head and a hydrophobic[2] tail (Figure 2.3).

2.0.4 Lipid aggregates

The attractive interaction between non-polar 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 mono-layers, micelles, bilayers, vesicles and biological membranes (Figure 2.4). Different structures form under different environmental conditions. For example, mono-layers tend to form at the

---

[1] Hydrophilic - “likes water” (hydro = water; philic = like) Meaning that a hydrophilic molecule or portion would be attracted to water.

[2] Hydrophobic - “fears water” (hydro = water; phobic = fear). Meaning that a hydrophobic molecule or portion would be repulsed/push-away a water molecule. Examples: oils, fatty acids (i.e. the “tails” of phospholipids), cholesterol.

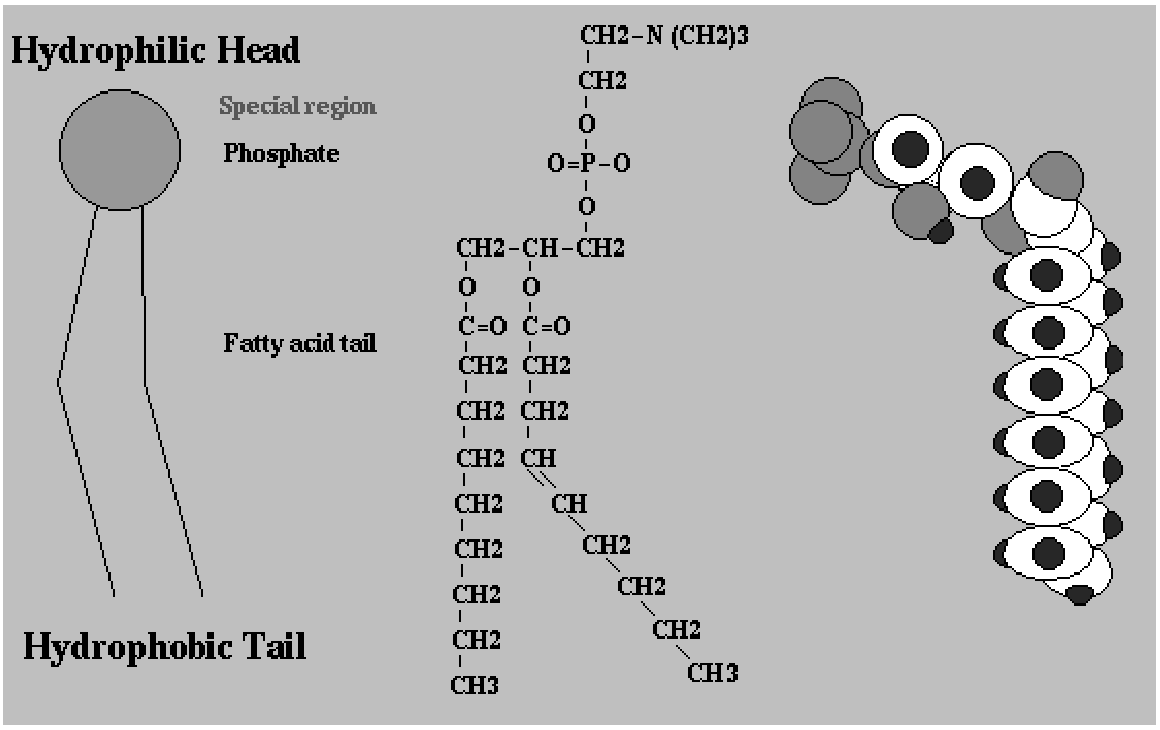


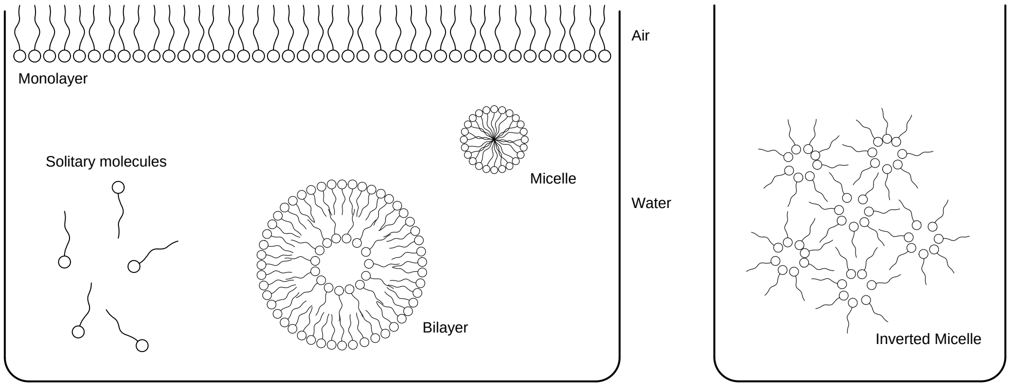
Figure 2.3: Molecular structure of phospholipid.

surface between air and water, or between oil and water, with their “heads” in the water environment and “tails” stay in the air or oil; micelles usually form ball-shaped structure in an aqueous environment, with the lipids’ heads stay in contact with water and tails pointing toward inside.

Current findings about the lipid aggregates are significant but far from complete. Below we briefly discuss the biological facts and existing theories.

Structures and properties

Amphiphilic molecules such as lipids can *self-associate* or *self-assemble* to small molecular aggregates such as mono-layers, micelles, bilayers, vesicles and biological membranes. (Figure 2.4). These structures and the systems they form—sometimes collectively referred



(a) Water (b) Oil

Figure 2.4: Formations of lipids in water and oil environments [34].

to as *association colloids* or *complex fluids*—are different from the conventional colloidal particles in one important respect: unlike solid particles or rigid macromolecules such as DNA, they are soft and flexible, i.e., *fluid-like*. This is because the forces that hold amphiphilic molecules together in micelles and bilayers are not due to strong covalent or ionic bonds but arise from weaker van der Waals, hydrophobic, hydrogen-bonding and screened electrostatic interactions.

Most amphipaths with a long polymethylene chain form a thin film at the air-water interface. Such an arrangement avoids energetically unfavorable interactions between acyl chains and polar solvents like water, and it also avoids unbalanced cohesive forces between the water molecules at the interface. As there is considerable difference between the polarity of the two media, amphipathic molecules align themselves such that the polar groups face the water and the acyl chains extend into the air. An interface between water and a non-polar solvent like benzene also offers an ideal environment for the formation of mono-layers where the acyl chains are in the organic medium.

Amphipaths with a relatively large head group form aggregated structures called micelles in which the polar groups are on the surface and the chains are segregated from the aqueous phase. In organic phases containing small amounts of water, the energetically favorable organization is inverted; in such inverted micelles the polar groups are oriented toward the center and enclose a small droplet of water (Figure 2.4b). Depending on the structure of the monomer and the environment, the size of micelles ranges from a sphere of about 5 nm for dihexanoyl- and diheptanoyl-PC to oblate cylinders of several hundred nanometers in length for dioctanoyl-PC.[34]

Double chained lipids that cannot pack into micellar structures due to their small head-group area or because their hydrocarbon chains are too bulky to fit into such small aggregates tend to form bilayer structures. They can appear in vesicle form or planar bilayer form. Bilayer is the basic structure of cell membranes, incorporated properties such as flexibility, selective permeability, which are also the most important features of membranes.

Conditions under which these structures are formed

To discuss this subject, we need to start with the fundamental thermodynamic equations of self-assembly.

Equilibrium thermodynamics requires that in a system of molecules that form aggregated structures in solution the chemical potential of all identical molecules in different aggregates be the same. This can be expressed as

$$\mu = \mu_1^0 + kT\log X_1 = \mu_2^0 + \frac{1}{2}kT\log\frac{1}{2}X_2 = \mu_3^0 + \frac{1}{3}kT\log\frac{1}{3}X_3 = ...$$

or

$$\mu = \mu_N = \mu_N^0 + \frac{kT}{N} \log\left(\frac{X_N}{N}\right) = constant, N = 1, 2, 3, ...,$$

where μ_N is the mean chemical potential of a molecule in an aggregate of aggregation number N, μ_N^0 is the standard part of the chemical potential(the mean interaction free energy per molecule) in aggregates of aggregation number N, and X_N the concentration of molecules in aggregates of number N ($N = 1$, μ_1^0 and X_1 correspond to isolated molecules in solution). The conservation relation for the total solute concentration C is:

$$C = X_1 + X_2 + X_3 + ... = \sum_{N=1}^{\infty} X_N$$

Depending on how the free energies μ_1^0, μ_N^0 are defined the dimensionless concentrations C and X_N can be expressed volume fraction or mole fraction units.

Aggregates form only when there is a difference in the cohesive energies between the molecules in the aggregates and the dispersed states. If the molecules in different sized aggregates all experience the same interaction with their surroundings, the value of μ_N^0 will remain constant in different aggregates. The necessary condition for the formation of large stable aggregates is that $\mu_N^0 < \mu_1^0$ for some value of N. The *critical micelle concentration* (CMC) is the concentration at which further addition of solute molecules results in the formation of more aggregates while leaving the monomer concentration more or less unchanged at the CMC value. Simple molecules can self-assemble into linear aggregates, discs and spheres, whereas more complex amphiphilic molecules can assemble into more complex shapes such as vesicles, interconnected rods or even three-dimensional “periodic” structures.

The major forces that govern the self-assembly of amphiphiles into well-defined structures such as micelles and bilayers derive from the hydrophobic attraction at the hydrocarbon-water interface, i.e., from the hydrophobic property of the lipid tails. The forces derived from the hydrophobic effect induces the molecules to associate, and the hydrophilic, ionic or steric repulsion of the headgroups, which imposes the opposite requirement that they remain in contact with water. These two interactions compete to give rise to the idea of two "opposing forces" acting mainly in the interfacial region: one tending to decrease and the other to increase the interfacial area a_0 per molecule (the effective headgroup area) exposed to the aqueous phase (Figure 2.5).[33]

Geometrical properties of different lipid structures.

Let's look at lipid formations from another perspective—its geometrical properties.

Lipids tend to form different kind of aggregates depending on different lipid molecule structures, i.e., they depend on their optimal head group area a_0, the volume v of their hydrocarbon chain or chains, which is assumed to be fluid and incompressible, and the maximum effective length that the chains can assume. There is a limit on how far the chains can extend, usually called the *critical chain length*, l_c. Given the optimal area a_0, hydrocarbon chain volume v and critical length l_c for a molecule, one may ascertain which structures the molecules can pack into consistent with these geometric constraints. For lipids of optimal area a_0, hydrocarbon volume v and critical chain length l_c, the value of the packing parameter v/a_0l_c, will determine whether they will form spherical micelles ($v/a_0l_c < \frac{1}{3}$), non-spherical micelles ($\frac{1}{3} < v/a_0l_c < \frac{1}{2}$), vesicles or bilayers ($\frac{1}{2} < v/a_0l_c < 1$), or "inverted" structures ($v/a_0l_c > 1$) [33].

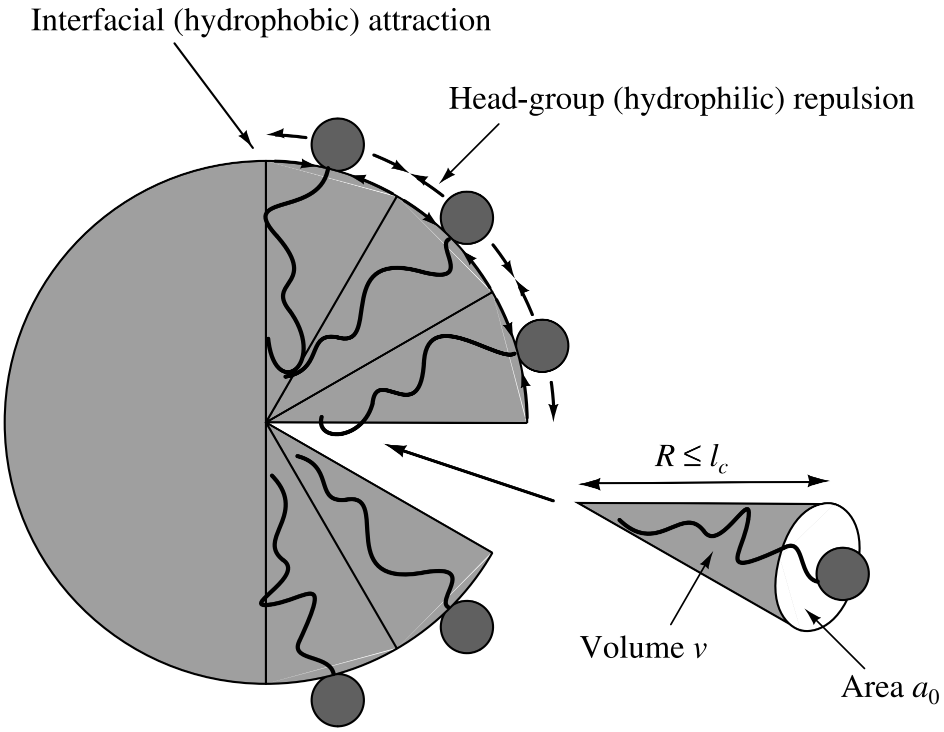


Figure 2.5: Geometric properties of lipid aggregates.

- Spherical micelles

For molecules to assemble into spherical micelles, their optimal surface area a_0 must be sufficiently large and their hydrocarbon volume v sufficiently small that the radius of the micelle R will not exceed the critical chain length l_c. From simple geometry we have, for a spherical micelle of radius R and mean aggregation number M (Figure 2.5),

$$M = 4\pi R^2/a_0 = 4\pi R^3/3v,$$

namely,

$$R = 3v/a_0,$$

so that only for

$$\frac{v}{a_0 l_c} < \frac{1}{3}$$

will the amphiphiles be able to pack into a spherical micelle with their headgroup areas equal to a_0 and with the micelle radius R not exceeding l_c. An example of such micelle-forming amphiphiles is the 12-carbon chain sodium dodecyl sulphate surfactant (SDS) in water.

- Non-Spherical and Cylindrical Micelles

  Most lipids that form spherical micelles have charged headgroups. Because charged headgroups usually leads to large headgroup areas a_0. Addition of salt partially screens the electrostatic inter-headgroup repulsion and thereby reduce a_0. Those lipids that possess smaller headgroup areas such that $\frac{1}{3} < v/a_0 l_c < \frac{1}{2}$ cannot pack into spherical micelles but can form cylindrical (rod-like) micelles. Falling into this category are single-chained lipids possessing charged headgroups in high salt (e.g., SDS, CTAB) or those possessing uncharged, non-ionic or zwitterionic headgroups(e.g., $C_{12}E_5$, lysolecithin).

- Bilayers

  Lipids that form bilayers are those that cannot pack into micellar structures due to their small head-group area a_0 or because their hydrocarbon chains are too bulky to fit into such small aggregates while maintaining the surface area at its optimal value. For bilayer-forming lipids, the value of v/a_0l_c must lie close to 1, and requires that for the same headgroup area a_0 and chain length l_c, their hydrocarbon volume v must be about twice that of micelle-forming lipids (for which v/a_0l_c is in the range $\frac{1}{3}$ to $\frac{1}{2}$). Therefore, lipids with *two* chains are likely to form bilayers. For example, single-chained lysolecithins form small but non-spherical micelles while lecithins with two alkyl chains, such as $di - C_{14}$ and $di - C_{16}$ lecithins, form bilayers. (The doubling of the chains also affects other aggregate properties — it increases the hydrophobicity of the lipids, which in turn drastically lowers their CMC, also it increases the lifetimes or "residence times", τ_R, of the molecules within aggregates.)

- Vesicles

  Under certain conditions it becomes more favorable for closed spherical bilayers (vesicles) to form rather than infinite planar bilayers. This arises since in a closed bilayer the energetically unfavorable edges are eliminated at a finite, rather than infinite, aggregation number, which is also entropically favored. Thus, so long as the lipids in a curved bilayer can maintain their areas at their optimal value, vesicles should be the preferred structures.

- Inverted Micelles

  For lipids with very small optimal headgroup areas or with bulky polyunsaturated chains (large v, small l_c), the value v/a_0l_c will exceed unity. In this case, "inverted" micellar structures are formed. [33]

Chapter 3

Previous work on computer simulation of lipid aggregates

We discussed in the previous section that, 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 [33]. However, exactly how lipid molecules formed these protocell structures in a prebiotic environment is still unknown.

The research necessary to gather all this evidence naturally involves astronomy and chemistry, but also relies heavily on the use of computer models and simulation to verify theories and generate supporting evidence which may be tedious or impossible to reproduce in the laboratory.

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 [49] [7]. 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.

3.0.5 B. Smit's Model of Micelles and Monolayer Structures

B. Smit *et al.* developed a computational model to simulate the dynamics of an oil-water-surfactant system [6]. 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 ($wwooooo$), 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:

$$\Phi = 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^{6}] \tag{3.1}$$

The Lennard–Jones potential is a potential that describes the interaction between two uncharged molecules or atoms. It is mildly attractive as two uncharged molecules or atoms approach one another from a distance, but strongly repulsive when they approach too close (Figure 3.1).

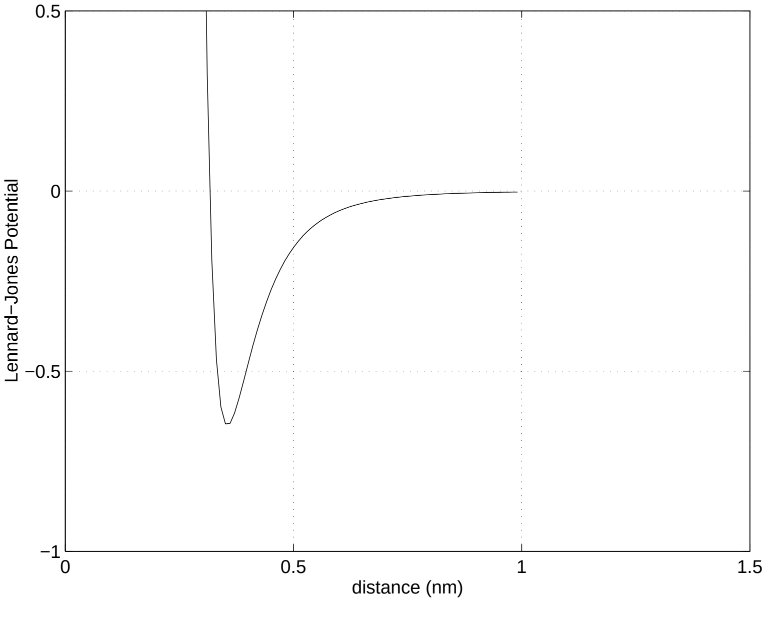


Figure 3.1: Lennard-Jones potential for water molecules.

σ and ϵ are the specific Lennard–Jones parameters, different for different interacting particles. For water, the values are: $\sigma = 0.316555 nanometers$ and $\epsilon = 0.6501696 kJ/mol$.

The strong close-in repulsion between molecules results from mutual deformation of their structures (meaning, one molecule cannot diffuse through another). The mild attraction at larger distances results from *induced dipole–dipole moment*[1] *interaction* of the particles. At equilibrium, the pair of atoms or molecules tends approach a separation corresponding to the minimum of the Lennard–Jones potential. σ and ϵ are the specific Lennard–Jones parameters, different for different interacting particles. When the separation r is very small, the $1r^{12}$ term dominates, and the potential is strongly positive. Hence the $1r^{12}$ term describes the short range repulsive potential due to the distortion of the electron clouds at small separations. In contrast the $1r^{6}$ predominates when the separation increases in magnitude. Hence the $1r^{6}$ term describes the long–range attractive tail of the potential between two particles.

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

---

[1] *Dipole* means “having two poles”, a positive and a negative one. A dipole is an electrical structure in which charges of equal magnitude but opposite sign are separated along a line. A *dipole moment* is a measure of the “strength” of a dipole.

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 [33]. However, in these simulation experiments, the surfactants have large tails and small head-group areas, and thus should form inverted micelles.

3.0.6 J. M. Drouffe's Model for Bilayer Structures

J. M. Drouffe *et al.* developed a computational model that simulates three-dimensional particles which self-assemble to form membrane-like objects [16]. 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(see Figure 3.2).

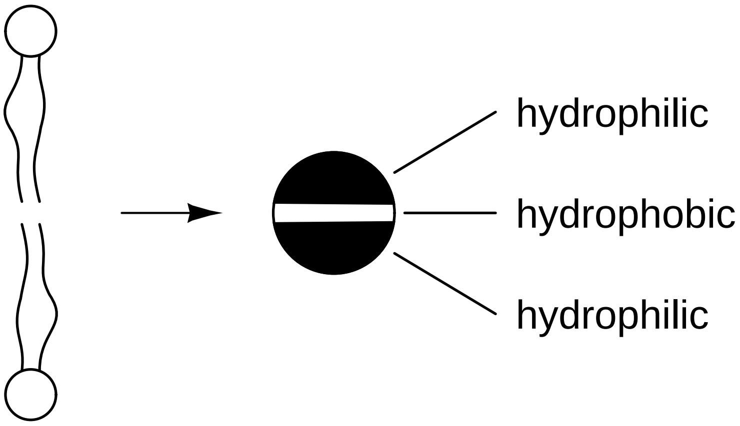


Figure 3.2: The basic particle structure in Drouffe's model.

Drouffe used a set of interaction forces to represent the inter-particle interactions (Figure 3.3):

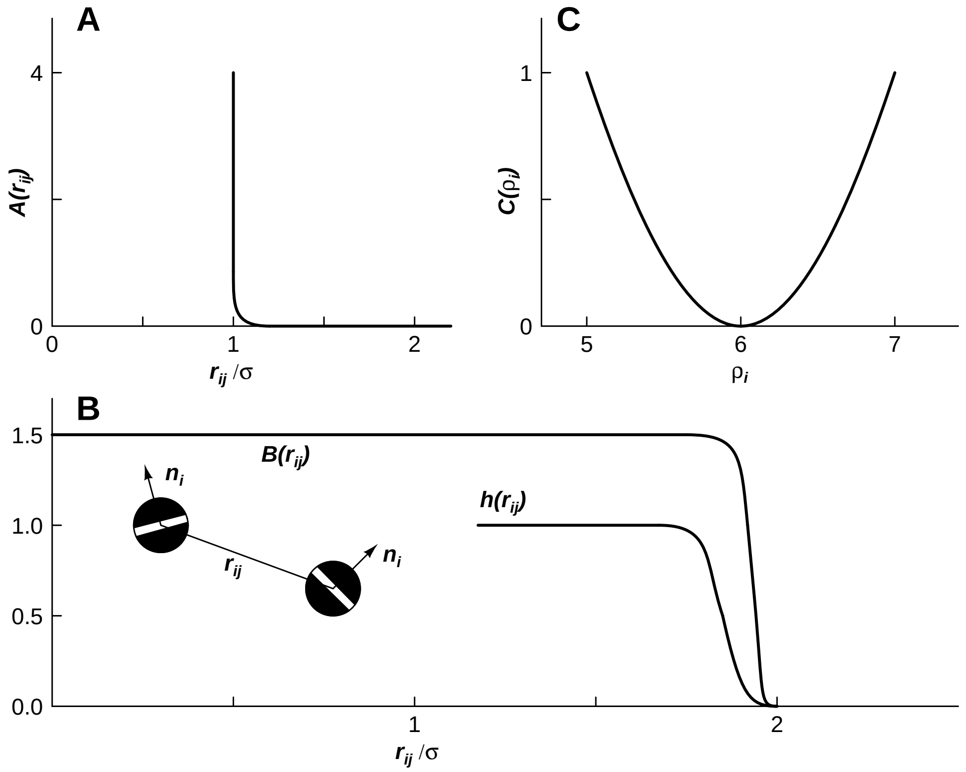


Figure 3.3: The interactions between the particles used in Drouffe's model.

- A hard-core repulsive interaction $A(r_{ij})$ of range σ, ensuring the incompressibility of the molecules at high density.

- An anisotropic interaction, $B(r_{ij})\{(n_i \cdot n_j)^\alpha + g[(n_i \cdot r_{ij})^2] + g[(n_j \cdot \hat{r}_{ij})]^2\}$ between pairs of particles i and j, where n_i and n_j are unit vectors expressing the orientation of the particles and $\hat{r}_{ij}$ is the unit vector between the particles. $g(x) = ax^3 + bx + c$. In this simulation, $a = 0.75$, $b = 0.25$, and $c = 0.8$.

- A multibody "hydrophobic" potential that expresses the large increase in free energy occurring when water is in contact with the hydrocarbon parts of the amphiphiles.

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.

3.0.7 Dynamics of a Membrane Patch by H. Heller

H. Heller *et al.* performed a molecular dynamics simulation of a rectangular patch of a lipid bilayer structure (membrane) in a water environment [32]. 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. It included all atoms explicitly except non-polar hydrogen atoms (Figure 3.4).

They built a 60 node multiple instructions, multiple data (MIMD) type parallel computer with double ring architecture. Each node of this machine consists of a T800 Transputer processor with a 4 Mbyte dynamic random access memory (DRAM). The computer

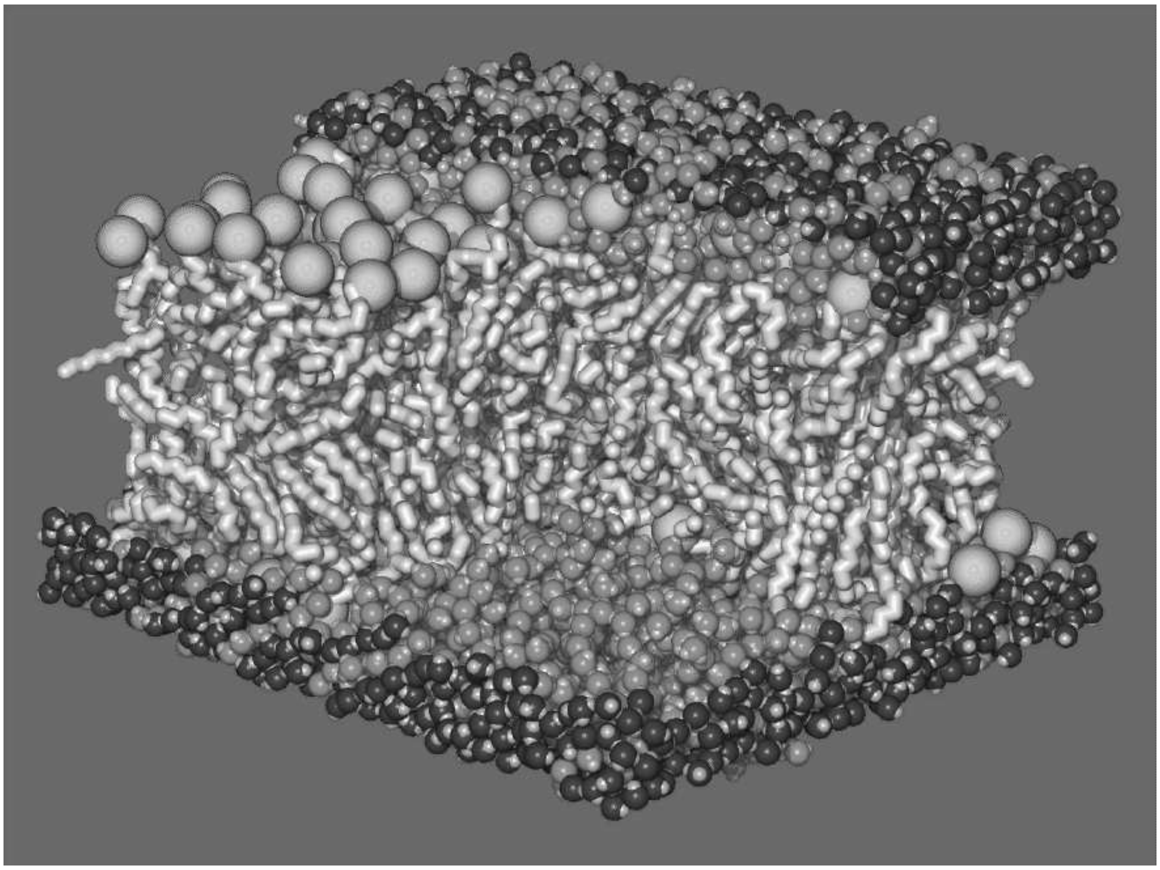


Figure 3.4: The membrane-patch generated in Heller's Model.

achieves Cray 2 processor speed for molecular dynamics calculations. The computations for the respective molecular dynamics simulations consumed a total of $14,640$ hours or 20 months.

This is undoubtly an extraordinary effort. 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).

3.0.8 Lattice Molecular Automaton by B. Mayer

Finally, B. Mayer *et al.* introduced a different type of self-assembly simulation, the lattice molecular automaton (LMA) model [41]. 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.

3.0.9 Problems in Existing Models

It is not difficult to see the problems among the existing computational models. While different models were designed for slightly different purposes, these various problems prevent these models to be generalized to simulate the self-assembly process. For example, Smit's model is a very abstract representation of the actual biological system. Although the system can model formation of micelles and monolayer structures, it does not provide much insight on how and why different system forms. Drouffe's model used a two-lipid combined structure as the basic unit for simulation, besides the fact that the anisotropic force the model used was not justified, while it can model the formation of bilayer structures, such model cannot be used to simulate the formation of other lipid aggregates such as monolayers, micelles, and reversed micelles. Heller's model provide a fairly accurate representation of the actually biological system, however, since it was modeling the system at an atomic level, the computational complexity prevent it to be extended to any larger system. Heller's model simulated a bilayer patch consists of 200 lipid molecules and simulated the actual time span of 263 ps ($1 \times 10^{-12} seconds$), while the actual formation of lipid aggregates usually takes seconds to minutes.

Realism and efficient become essential to the successful modeling of the lipid self-assembly process.

Chapter 4

Research Goal — A Molecule-Level Computational Model

Our research goal is to build a computational model to simulation the formation of self-assembled protocell structures, by analyzing and simulating the interactions between lipid molecules and the conditions under which various structures are formed. While we want to borrow many of the useful ideas adopted in the research described above, we also want to avoid the problems presented to provide a more realistic and yet computationally feasible model.

Although many people had attempted simulations along different directions, no one has gotten very far. The computation involved is very intensive. In simulation of self-assembly of pre-cell formations, deciding to what level of detail should the model simulate is very critical. As we have seen from the previous section, if the model is too coerce, the conditions under which the structures are formed cannot be well represented. However if we build the simulation model on a too detailed level, like the simulation of a membrane-patch presented by Heller, *et. al* [32], then the computation is so intense that we would not be able to go very far. Based on these considerations, we propose the following abstract level for the simulation:

4.0.10 Molecule-level computational models

In this model, we want to reflect the biological features that are believed to be important to the proto-cell formation. It should be realistic yet computationally feasible. The model should be able to be used to simulate the actual self-assembly process.

We think that, 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.

In order to balance the realistism and computational feasibility, to avoid the problems in the computational models described in the previous section, we choose to build the computational model on a molecule-level — use the macro-molecule as a basic unit in the simulation, define the interactions between these molecules, and furthermore, simulate the self-assembly process.

4.0.11 Basic structures of the model — forces and justification

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.* ([6]) 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* [16], 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.

Chapter 5

Basic model and interacting forces

We now demonstrate the model using the simulation of micelle structures as an example.

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". Here we want to model a system in which initially random distributed lipid-like particle self-assemble into micelle structures under the interactions.

5.0.12 Basic lipid structure

We discussed in earlier sections that the major driving force for lipid aggregates is believed to derived from the hydrophobic effects. And the amphiphilic molecules obtained the distinct hydrophobic properties because their unique molecular structures — each amphiphilic molecule has a hydrophilic head that favors an aqueous environment and a long hydrophobic tail which favors a none-water environment. We think that modeling a 'structured' particle is essential to a realistic modeling of the system. Therefore, in order to simulate the 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 5.1).

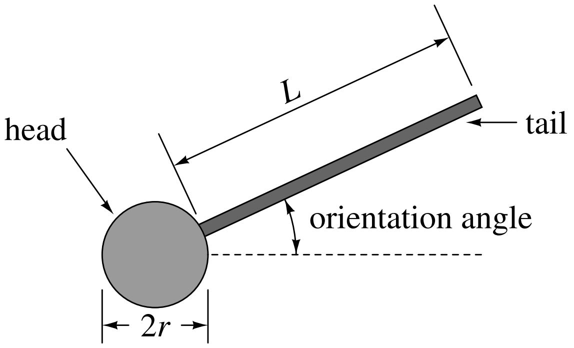
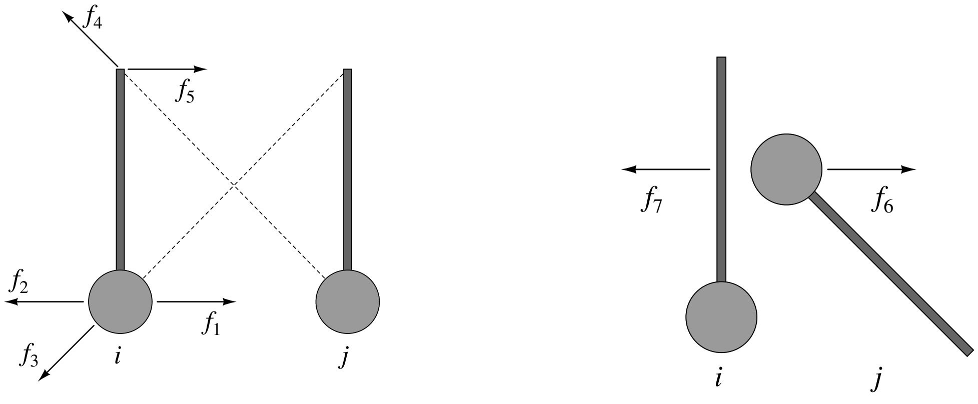


Figure 5.1: Lipid particle — The basic element in the simulation.

5.0.13 Interaction forces: definition and justification

Based on hydrophobicity and other physical and chemical properties of lipids, we define seven inter-particle interaction forces (Figure 5.2). 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.

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.



(a) (b)

Figure 5.2: The interaction forces defined in the model.

- 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 5.2b). 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.

5.0.14 Computing the forces

The directions and the points of actions of forces f_1 through f_7 are shown in Figure 5.2. 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 [16] or the Lennard-Jones potential in [6]), ramp functions (see Figure 5.3) 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.

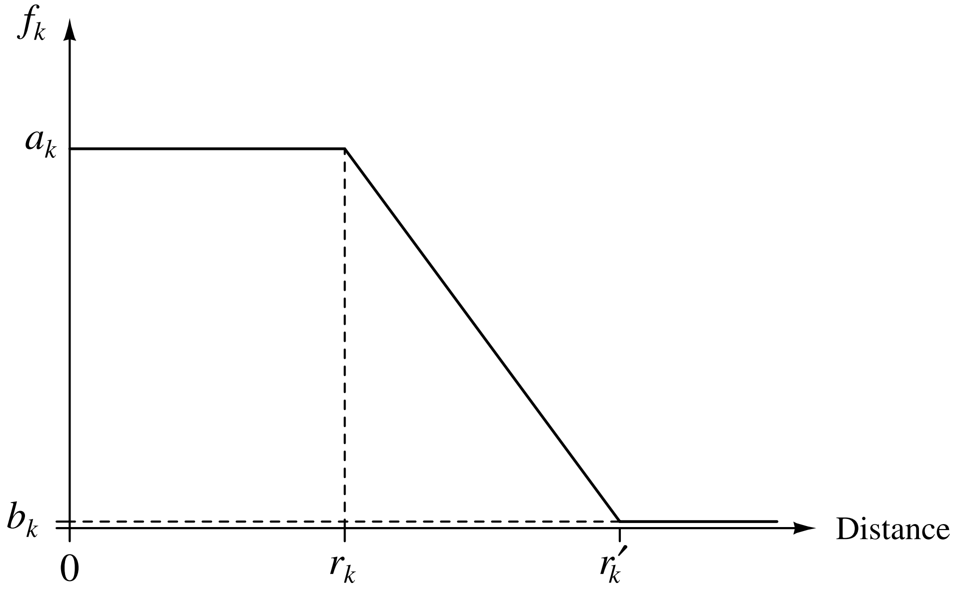


Figure 5.3: Ramp function for the magnitudes of all interacting forces.

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

$$f_k = a_k * ramp(b_k, r_k, d_k). \tag{5.1}$$

The reason we adopt the ramp function instead of using a differentiable smooth function is to reduce the computational cost.

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

$$\tau_k = f_k * \sigma_k * sin(\alpha_k), \tag{5.2}$$

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.

5.0.15 Overall movement influences

Now we consider the movement of each particle when subjected to the interactive forces from all other particles. In the system, each particle has a linear movement and a rotational movement. The linear movement of the particle decide the change of the particle position while the rotational movement decides the change of each particle's orientation angle.

Linear Movement

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

$$F_i(t) = \left(\sum_{j \neq i} \sum_{k} f_k(t)\right) + rand_1(t) - c_1 * v_i(t). \tag{5.3}$$

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.

Assume m is the mass of a particle, $\mathbf{a}$ the acceleration, from Newton's second law, we have

$$\mathbf{F}_i(t) = m\mathbf{a}$$

therefore,

$$\mathbf{a} = \mathbf{F}_i(t)/m$$

To simplify the computation, we assume each particle has a unit mass, then:

$$\mathbf{a} = \mathbf{F}_i(t)$$

Assume v is the velocity of a particle, using Euler's method to approximate Newton's second law of motion, we then have,

$$\Delta v_i(t + \delta t) = F_i(t) \cdot \delta t \tag{5.4}$$

$$\tag{5.5}$$

The linear displacement for the time step δt will be:

$$\Delta S_i(t+\delta t) \;=\; \frac{1}{2}\mathbf{a}\delta t^2 + v_i(t)\delta t \tag{5.6}$$

therefore,

$$\Delta S_i(t+\delta t) \;=\; \frac{1}{2}F_i(t)\delta t^2 + v_i(t)\delta t \tag{5.7}$$

Rotational Movement

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

$$T_i(t) = \left(\sum_{j\neq i}\sum_{k}\tau_k(t)\right) + rand_2(t) - c_2\omega_i(t) \tag{5.8}$$

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

$$I = \frac{1}{12}(2r+l)^2. \tag{5.9}$$

The rotational movement will then be

$$\Delta\omega(t+\delta t) \;=\; \frac{T_i(t)}{I}\delta t, \tag{5.10}$$

$$\Delta\theta_i(t+\delta t) \;=\; \frac{1}{2}\frac{T_i(t)}{I}\delta t^2 + \omega_i(t)\delta t \tag{5.11}$$

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

Chapter 6

Implementation of Micelle Formation

We first use micelle as an example to implement the computational model describe in last section. The system is divided into three major parts: Random system generator, simulation of interaction and animated display.

6.0.16 Definition of environment

We implemented the computational model in two-dimensional system. Each particle in the system therefore has three parameters to decide its location: x, y and a, where x and y are the artesian coordinates and a is the orientation angle of the particle.

6.0.17 Random system generation

The part of the implementation is straight-forward. We define a bounding box for the system. The set of random particles are generated using the following algorithm:

```
1. for $i = 1$ to $n$
```

```
    while $good - position == 1$
        genervate a random position for node $i : (x, y, a)$
        caculate the position of particle tail
        for $j = 1 to i$
            if node $i$ overlaps node $j$
                good-position = $0$
            end-if
        end-for
    end-while
end-for
```

The program outputs a position data file that includes the position and orientation angle of each particle, and an initial zero velocity for each particle (which includes v_x, v_y, v_a, where v_x and v_y are velocities on x and y directions, v_a is the rotational velocity.

Figure 6.1 shows an example of a randomly generated initial system: the size of the bounding box is $900x900$, each particle has a head radius of 5, tail length of 30, the system consists of 300 particles.

6.0.18 Simulation of the interactions

This is the major part of the simulation. 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

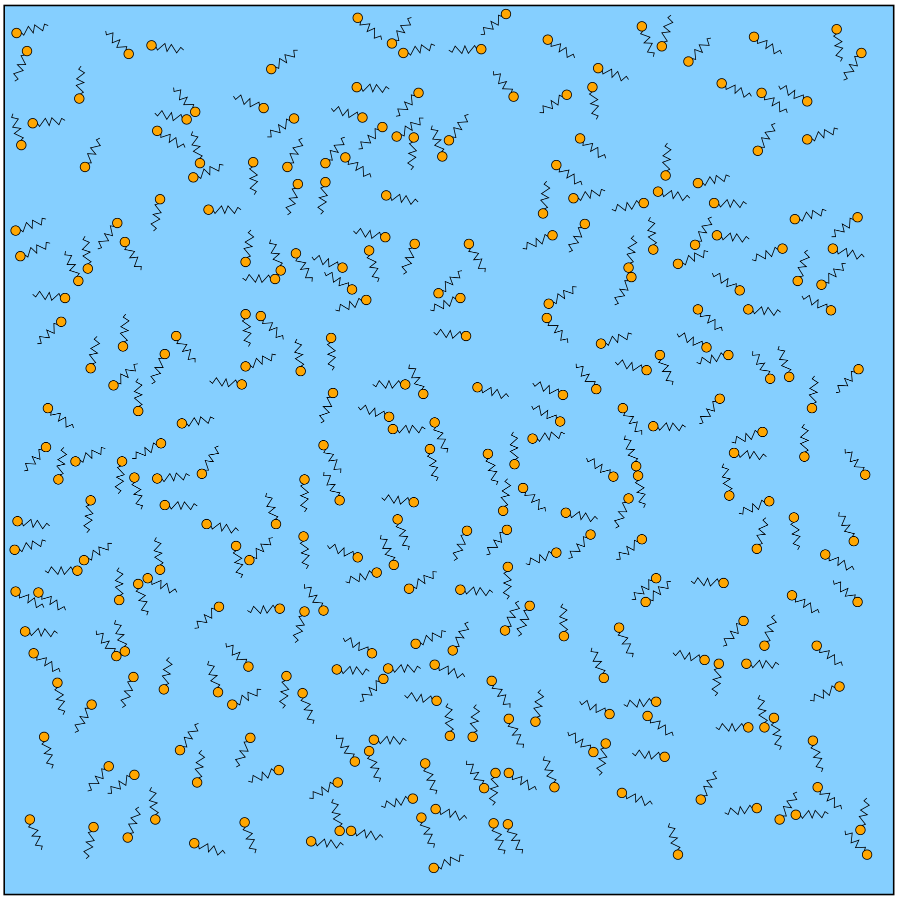


Figure 6.1: A randomly generated initial system of 300 particles.

structures from free-floating lipid molecules. The program takes input the position data generated from the random system generator, or from position data files generated via previous simulations. It then computes for each time step the interactions between each pair of particles in the system, then calculate the movement and new position for each particle.

The program output the positions of particles at each selected output-interval: because the simulation usually involves large number of time steps, the positions are only recorded at certain output time step intervals. The size of the output interval is defined by the user: it should be selected base on the consideration that it should be small enough to observe the assembly process of the particles, but large enough so that it does significantly affect the simulation speed, also ensure the output file to be small enough for ease of use.

The program generates three output files: the first one writes the positions and orientation angles of particles with the same file name and format as the input file: so that the new positions can be used to further simulation; The second file records all the location and orientations of all particles at each output interval, to provide the input data for displaying program; The third file logs the parameters used in the simulation.

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. 5.3 and 5.8.) are important for the system to reach a stable state.

6.0.19 Displaying method

This part of the project is fairly straight-forward: provide a graphics interface to observe the results of the simulations (Figure 6.2). It is implemented in C with Motif. The result of the simulation (described in previous section), including the position of each particle at every output time step, is saved in a data file. The display program retrieve the data from this file, and display for each time step the position of every particle of the system. The double-buffering technique is used to provide a smooth animation effect. The displaying program includes the following menu options:

Figure 6.2: Displaying program for the simulation.

- **File** Handle data file input and program termination.
  - **Open** Open the data file generated by the simulation program.
  - **Quit** Exit the displaying program.
- **Simulation** Control the display of the assembly-process.

  - **Start** Start the animation.
  - **Stop** Pause the animation.
  - **Restart** Go back to the initial state of the current data-file.
  - **Faster** Increase the speed of the assembly animation.
  - **Slower** Decrease the speed of the assembly animation.
- Print
  - **Write PS** The program writes the current frame into a Postscript file specified by the user.

The motion can also be controlled by mouse: each click of the right-mouse button cause the animation to move one time step forward. Although in the actually simulation we treat the particle tails as a long thin rod to simplify the computation, in the display program, we draw it as a zig-zag line to mimic the actual molecule structures.

Chapter 7

Simulation Results of Formation of Micelles

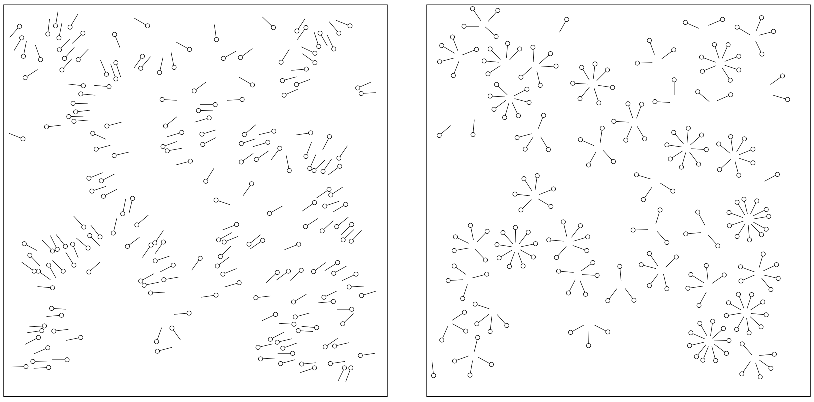
The results shown in this section are simulations of a system in an aqueous environment. The size of the environment is 900 × 900 (normalized) spatial units. Each lipid particle has a sphere head of radius 5 units, and a tail of length 30 units. There are 200 lipid-like particles randomly distributed 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 7.1.

A series of snapshots for different stages of Simulation 1 is shown in Figure 7.1. 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.

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

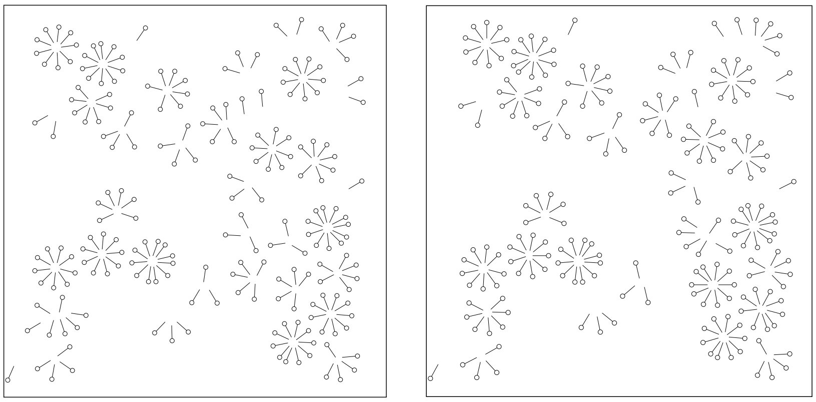
Force	Description	Magnitude a_k	Range r_k	Residual b_k
f_1	head-head attraction	40	15	0.4
f_2	head-head repulsion	30	30	0
f_3	tail-head repulsion	30	30	0.6
f_4	head-tail repulsion	30	30	0.6
f_5	tail-tail attraction	40	40	0.4
f_6	(active) repulsion due to incompressibility	$1,600$	15	0
f_7	(reactive) repulsion due to incompressibility	$1,600$	15	0

Table 7.1: Parameters used in simulation 1.



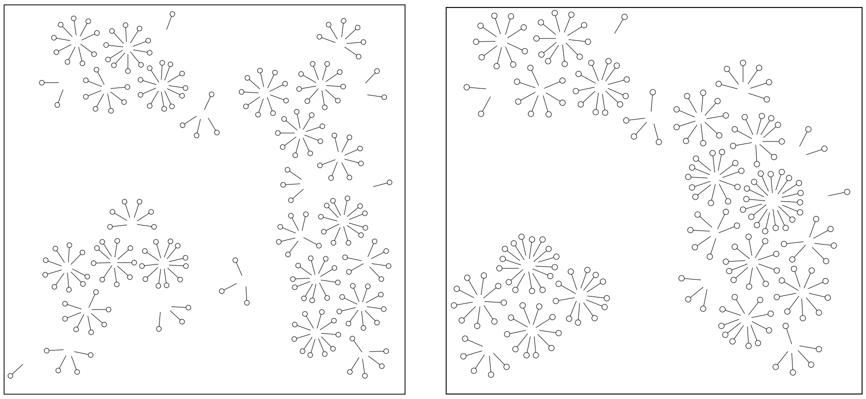
(a) initial state

(b) after 20,000 iterations



(c) after 40,000 iterations

(d) after 60,000 iterations



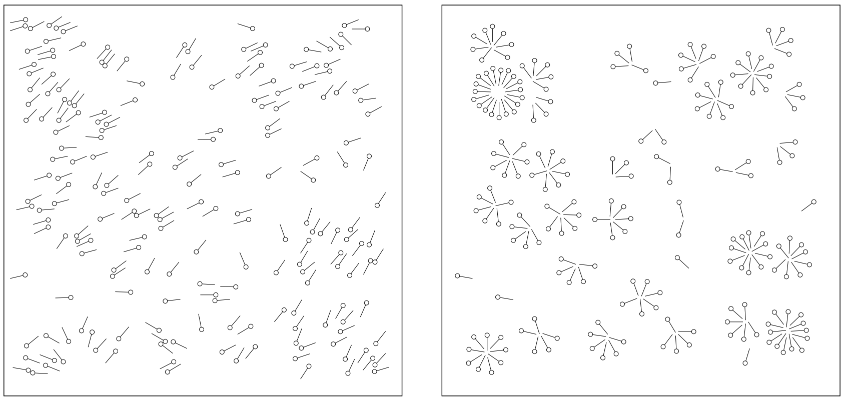
(e) after 100,000 iterations

(f) after 260,000 iterations

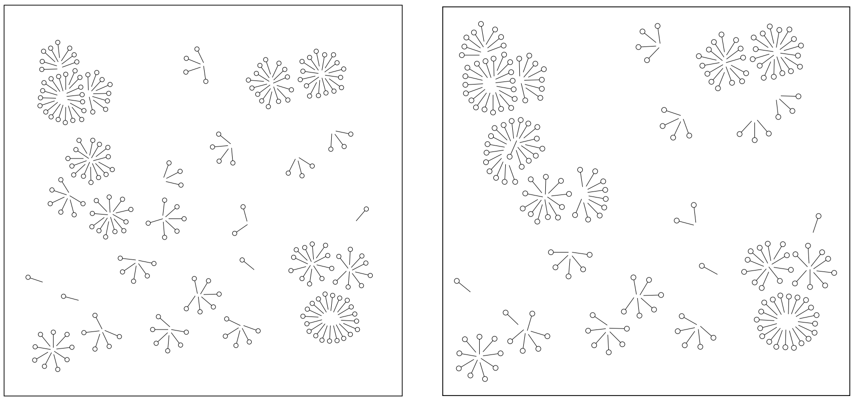
Figure 7.1: System of 200 lipid particles, Simulation 1.

stages in Simulation 2 are given in Figure 7.2. It can be observed that larger sized micelle structures are indeed formed in Simulation 2.

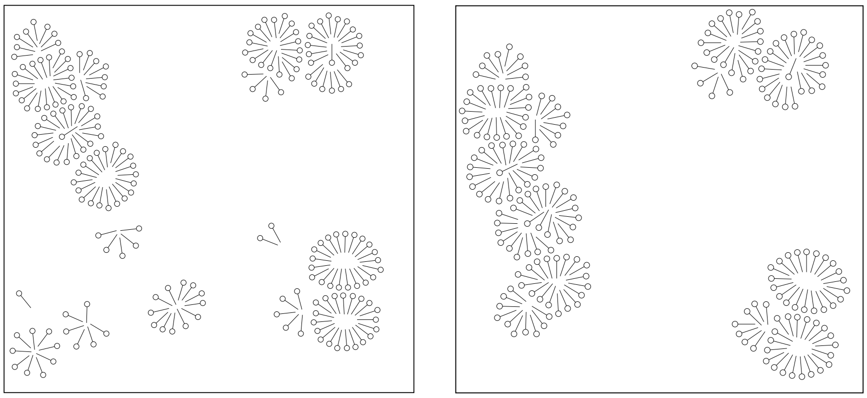
Simulation 1 was performed on an *SGI R4400* platform. The total running time for 260,000 iterations was about 76 hours. Simulation 2 was performed on an *SGI R8000* platform, with a total running time of about 57.5 hours for the same number of iterations. Actual simulations can be viewed more clearly 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.



(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 7.2: System of 200 lipid particles, Simulation 2.

Chapter 8

Formation of Other Structures

The next step is to see if the same computation model can be used to generate other type of lipid aggregates. We start with reversed micelle structures.

8.0.20 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.

The Model

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 8.1):

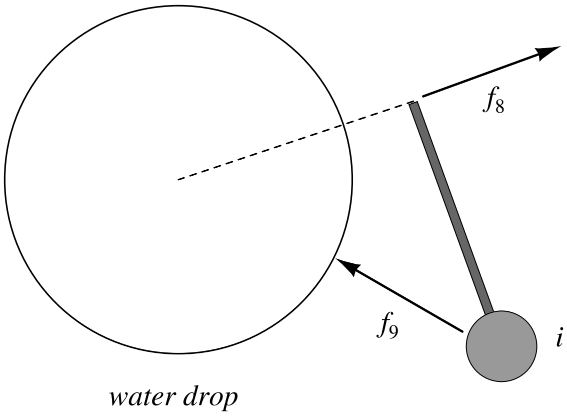


Figure 8.1: Interaction forces defined for reversed micelle simulation.

- 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 5.3. 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

Force	Description	Magnitude a_k	Range r_k	Residual b_k
f_2	head-head repulsion	30	30	0
f_3	tail-head repulsion	30	30	0.02
f_4	head-tail repulsion	30	30	0.02
f_6	(active) repulsion due to incompressibility	40	10	0
f_7	(reactive) repulsion due to incompressibility	40	10	0
f_8	tail-water repulsion	40	R_w	0
f_9	head-water attraction	2.0	$r_k > R_w - 2r$	0

Table 8.1: Parameters used in reversed micelle simulations.

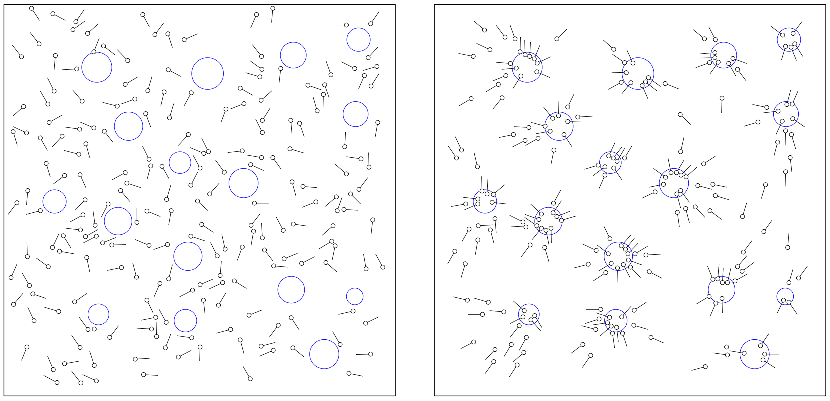
drop.

Implementation and Results

All previous implementation for the simulation of micelle structures are modified accordingly.

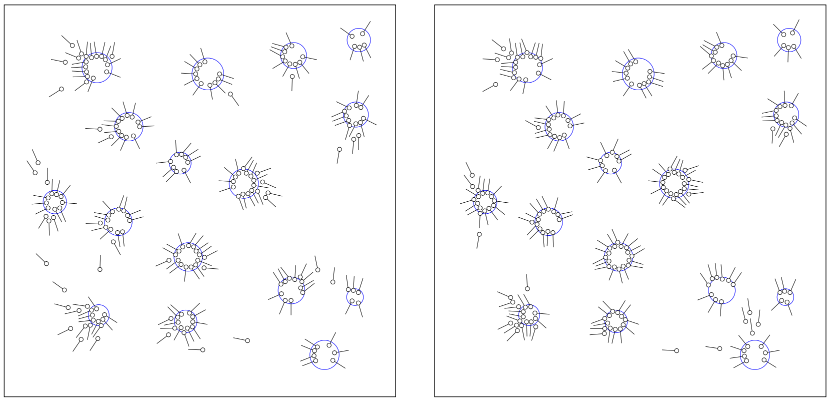
The initial system has the following components: the environment of the system is oil-based. In the system, there are randomly distributed water drops, displayed in the diagram as large, blue circles; the rest are randomly distributed lipid particles, which have the same structures as defined in previous simulation. An example of initial state system is shown in Figure 8.2(a).

A test simulation was performed on an *SGI R10000* platform. It simulated a system of



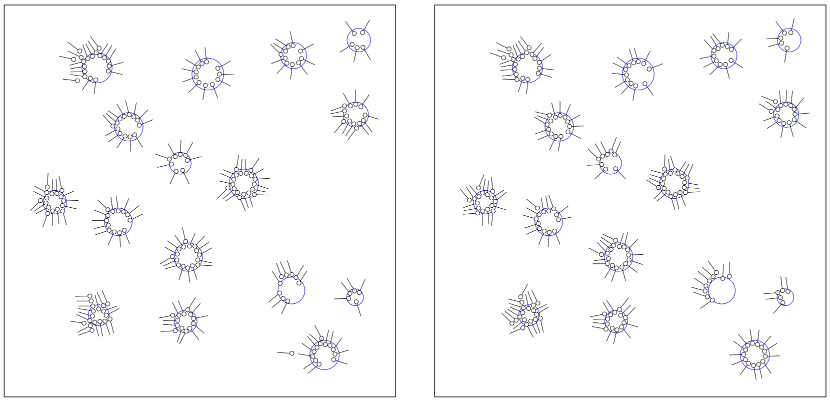
(a) Initial State

(b) Formation after 1,600 iterations



(c) Formation after 6,000 iterations

(d) Formation after 14,600 iterations



(e) Formation after 32,400 iterations

(f) After 100,000 iterations

Figure 8.2: System of 16 water drops, 200 lipid particles.

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 8.2, where reversed micelles are formed around water drops. The parameters used in this simulation are shown in Table 8.1.

8.0.21 Formation of Mono-Layers

The next structure we want to model is the formation of monolayers at the air-water surface.

Monolayer structure usually forms at the air-water or oil-water interface. The hydrophobic effect plays the major role in the formation process — the hydrophobic tails of lipid particles will try to stay outside the aqueous environment while the hydrophilic heads of the particles will try to remain in contact with water.

This type of structure is of great interest to surface chemists. The biological laboratory experiment is usually done by placing a solution of amphiphilic molecules such as lipids in alcohol on a water surface, after the alcohol is quickly dissolved in the water, the molecules of polar lipid arrange themselves at the air-water interface with the polar portion in the water and the hydrophobic tails projecting into the air.[1]

The Model

We simulate the formation of mono-layer structure at the air-water interface. Because two different environmental conditions are present in this model, molecules that are close the the air-water surface will be under the influence of two additional interaction forces:

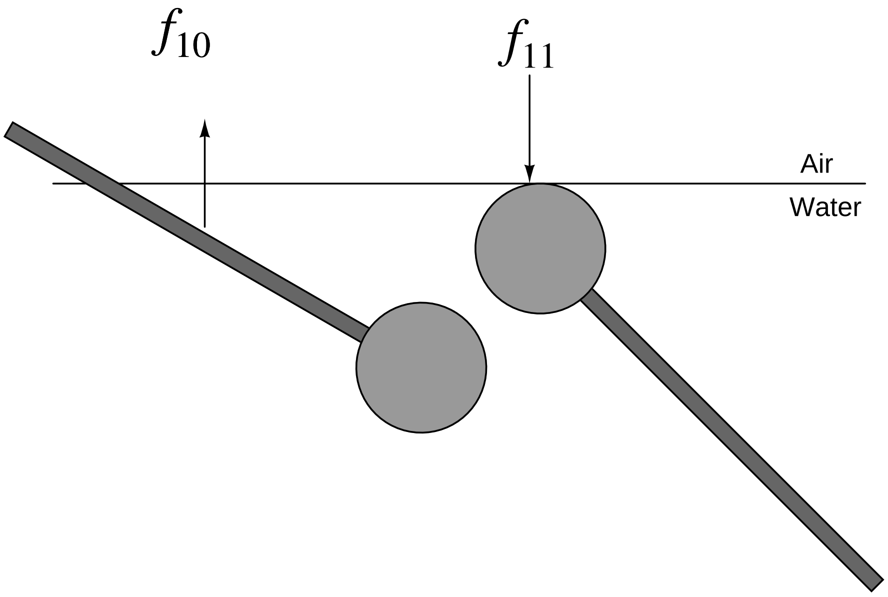


Figure 8.3: Interaction Forces for Monolayer Structures.

- F_{10} Hydrophobic Force — air-tail attraction. When a particle is at the surface of water, because its molecule tail is hydrophobic, the molecule will tend to stay in the position so that its tail will stay in the air, away from the water surface. This is emulated by a force that is placed on the center of the tail of the particle, pointing upward toward the air (Figure 8.3.

- F_{11} Hydrophilic Force — air-head repulsion. Because the heads of the particles are hydrophilic, when a particle is close to the surface of the water, it will tend to stay in a position to ensure that its molecule head stays in contact with water. This is denoted by a force that is placed on the center of the molecule head, point downward.(Figure 8.3.

Implementation and Results

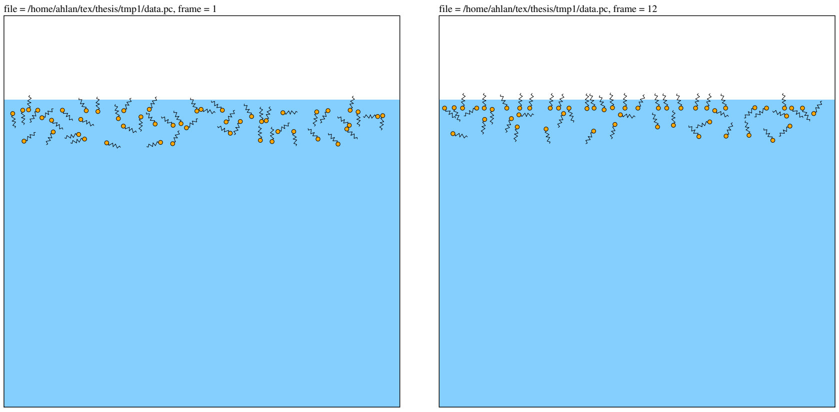
The simulation environment is defined as following: an air-water surface presents in the environment. Below the surface is water, above the surface is air. A set of initially random distributed particles are generated, placed below the surface of the system Figure 8.4(a). Then the simulation program compute the interactions among the particles and their environment at each time step. Figure 8.4 shows the simulation results of a system of 50 particles. The simulation was performed on an *194 MHZ SGI R10000*, the simulation time for $200,000$ time steps is about 36 minutes. The parameters used in this simulation are shown in Table 8.2. Figure 8.5 shows the simulation results of a system of 60 particles using the same set of parameters as in Figure 8.4. The simulation was performed on an *194 MHZ SGI R10000*, the simulation time for $200,000$ time steps is about 52 minutes.

8.0.22 Patch of Bilayer Structure

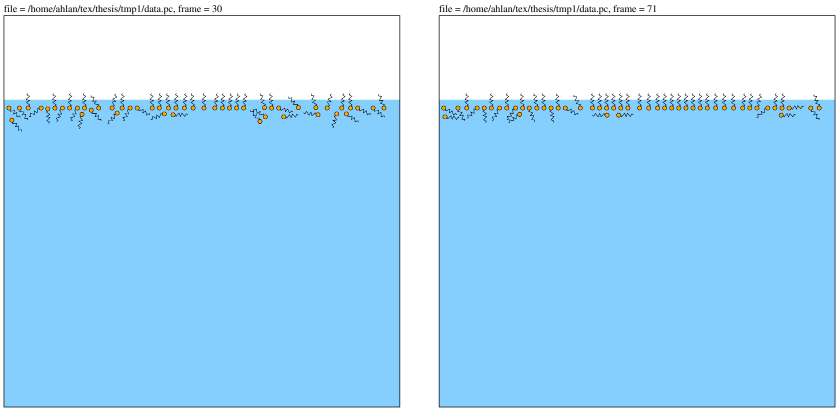
Bilayers are more complicated structures among the lipid aggregates. How the bilayer structure can form under certain conditions is still not well-understood.

We think that a first step to study the bilayer structure is to study the properties of a bilayer patch constructed by lipid particles. So we decided to do the following experiment: similar to what H. Heller *et al.* described in [32], we pre-construct a bilayer patch of lipid structure, and test the system under the interaction forces and observe the stability of the system.

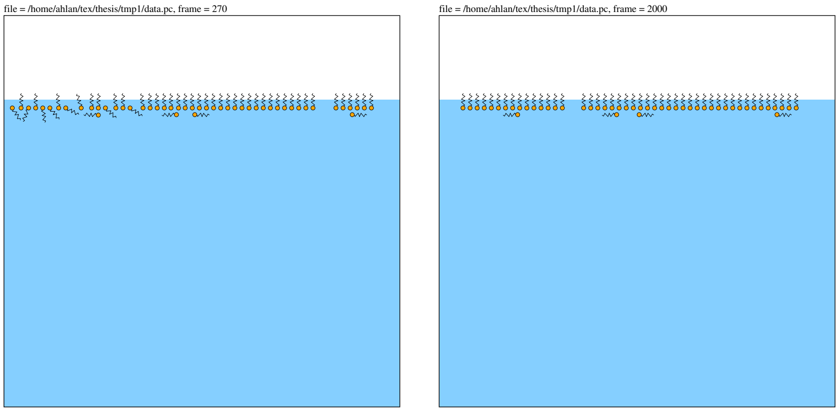
Following we describe the system in which we model the interactions between lipid particles within a pre-constructed bilayer patch.



(a) initial state (b) after $1,200$ iterations



(c) after $3,000$ iterations (d) after $7,000$ iterations

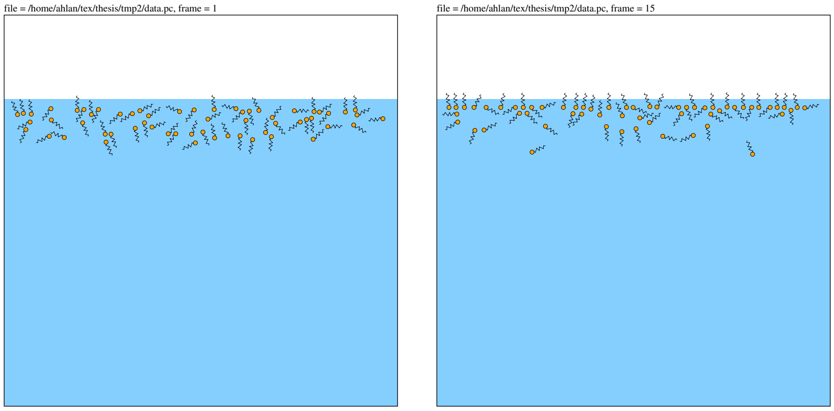


(e) after $27,000$ iterations (f) after $200,000$ iterations

Figure 8.4: Simulation of formation of monolayer structures

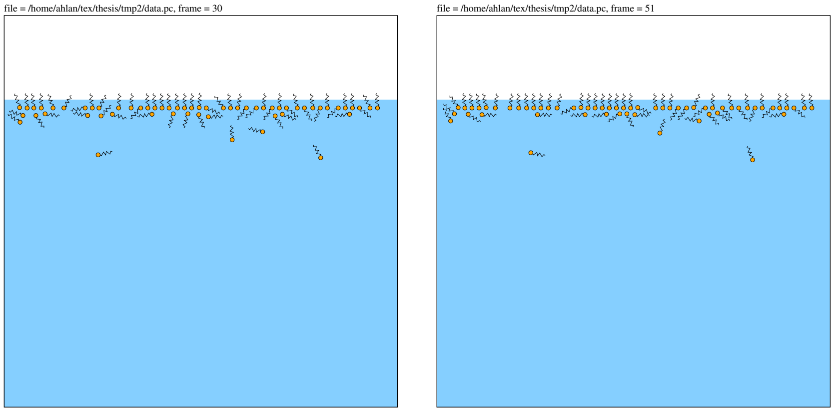
Force	Description	Magnitude a_k	Range r_k	Residual b_k
f_1	head-head attraction	30	30	1.0
f_2	head-head repulsion	20	20	2.0
f_3	tail-head repulsion	50	50	0.5
f_4	head-tail repulsion	50	50	0.5
f_5	tail-tail attraction	30	15	0.1
f_6	(active) repulsion due to incompressibility	800	15	2.0
f_7	(reactive) repulsion due to incompressibility	800	15	2.0
f_{10}	hydrophobic force	320	6	0.0
f_{11}	hydrophilic force	320	20	2.0

Table 8.2: Parameters used in Figure 8.4



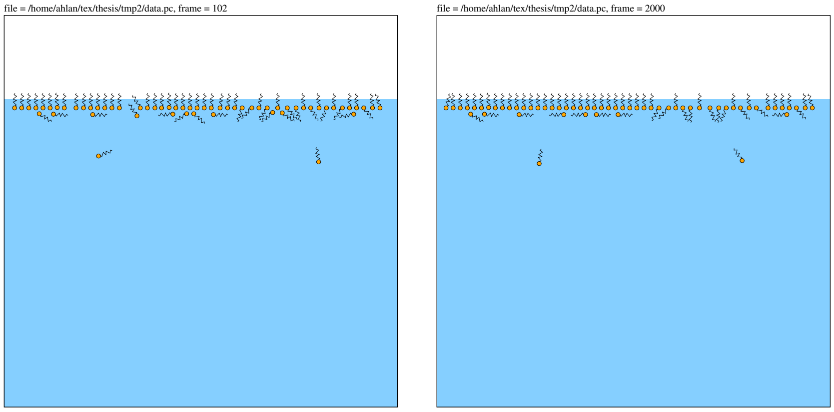
(a) initial state

(b) after $1,500$ iterations



(c) after $3,000$ iterations

(d) after $5,100$ iterations



(e) after $102,000$ iterations

(f) after $200,000$ iterations

Figure 8.5: Simulation of formation of monolayer structures

The Model

The lipid particle, which is the basic unit of the simulation model, is the same as described in previous sections. Each particle is a structured molecule with a hydrophilic head and a long hydrophobic tail. Because the bilayer structures usually formed in an aqueous environment, similar to the environment that micelles would form in, we define the same environment for the bilayer structure as for the micelle structure. Therefore, every interaction force in this simulation is the same as described in the model we defined for the formation of micelles. The system models the bilayer patch structure in an aqueous environment.

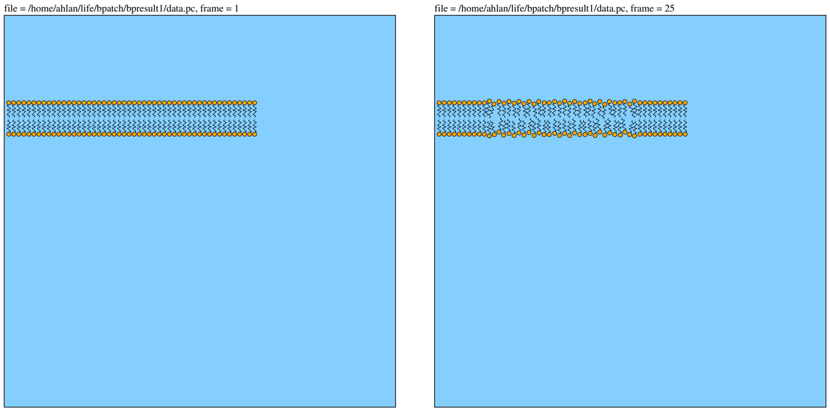
Implementation

The initial state of the system is a bilayer patch: two rows of lipid particles, each consists of same number of lipids. All the lipid particles in the top row have their hydrophilic heads on top, staying in contact with water, and hydrophobic tails pointing downward, away from water, while all the lipid particles in the bottom row have their heads on bottom and tails pointing upward.

Because we are modeling a bilayer patch instead of a whole bilayer vesicle, we fix the position of the particles on both ends of the bilayer patch, assuming that the same structure extends on both ends.

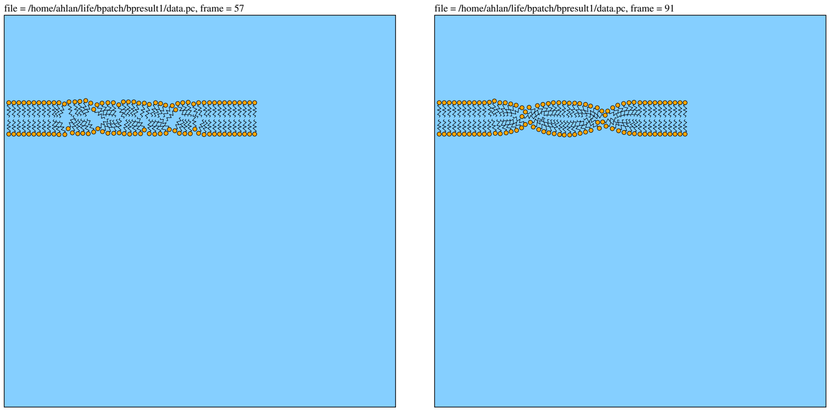
Simulation Results

Figure 8.6 shows the simulation results on a pre-constructed bilayer structure consists of 140 particles.



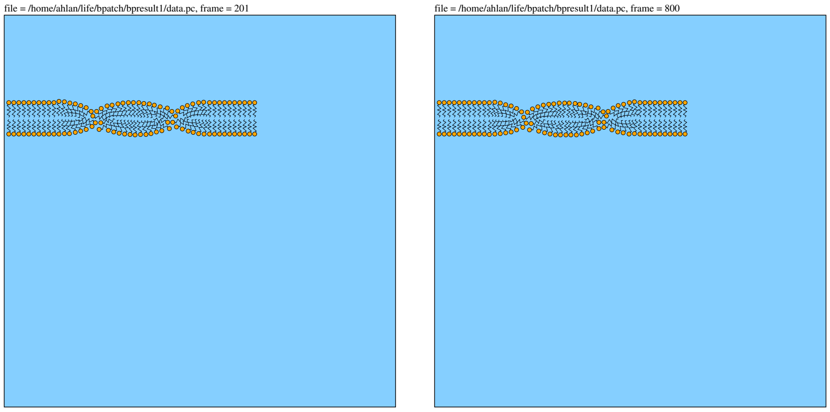
(a) initial state

(b) after $1,200$ iterations



(c) after $2,800$ iterations

(d) after $4,500$ iterations



(e) after $10,000$ iterations

(f) after $40,000$ iterations

Figure 8.6: System of 140 lipid particles, pre-constructed as a bilayer patch

Force	Description	Magnitude a_k	Range r_k	Residual b_k
f_1	head-head attraction	80	40	1.0
f_2	head-head repulsion	20	10	2.0
f_3	tail-head repulsion	20	20	0.5
f_4	head-tail repulsion	20	20	0.5
f_5	tail-tail attraction	30	25	0.1
f_6	(active) repulsion due to incompressibility	800	11	2.0
f_7	(reactive) repulsion due to incompressibility	800	11	2.0

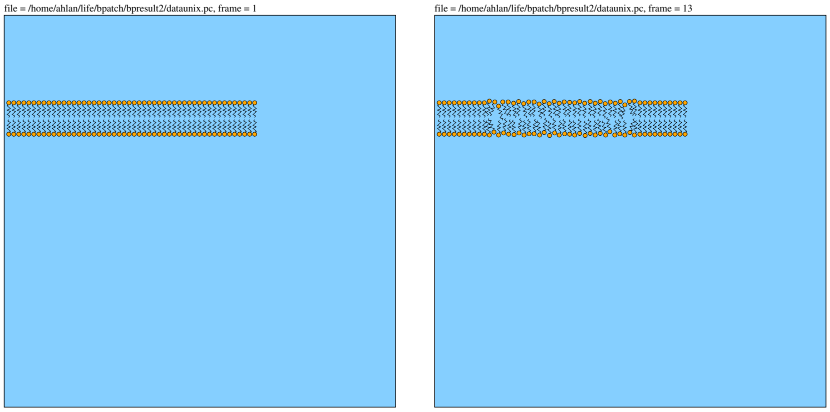
Table 8.3: Parameters used in Figure 8.6

Simulation shown in Figure 8.6 was done on a *Pentium 400*. The $40,000$ iterations took about 50 minutes. The parameters used in this simulation is shown in Table 8.3.

Figure 8.7 shows a simulation using the exactly same set of parameters and initial data, but on a *194 MHZ SGI R10000*. The running time for $40,000$ iterations is about 40 minutes. The difference in the results is caused by different random numbers in the system. This also shows how small changes in the parameters can affect the delicate balance of the system.

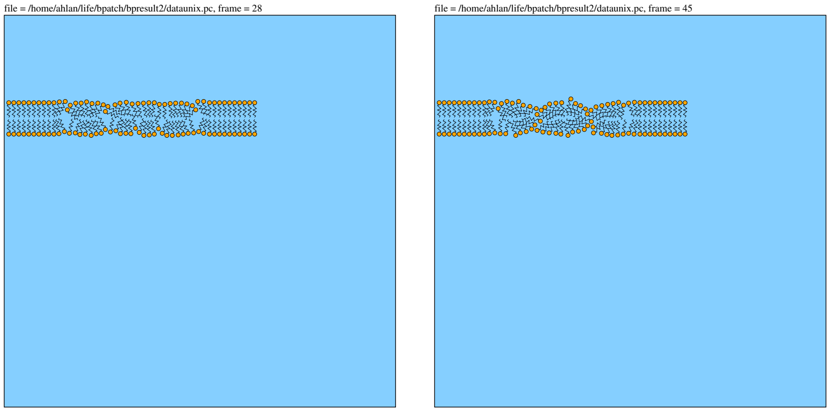
Result analysis

Since this experiment is very similar to what Heller has done in [32], we would like to compare our approach with Heller's. We think that the test model we perform will have the



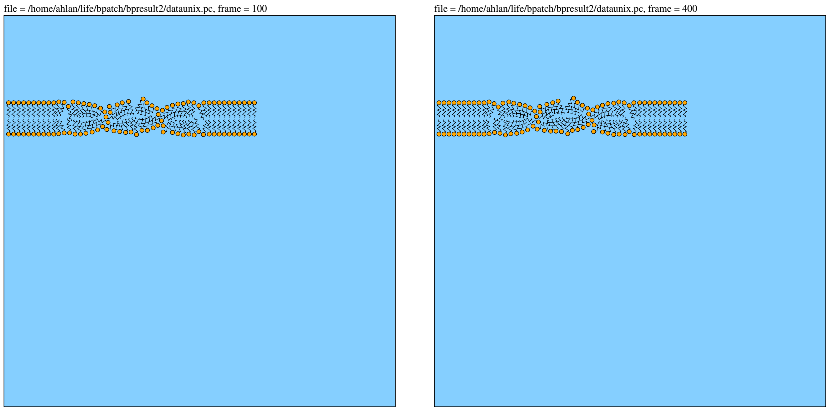
(a) initial state

(b) after $1,200$ iterations



(c) after $2,800$ iterations

(d) after $4,500$ iterations



(e) after $10,000$ iterations

(f) after $40,000$ iterations

Figure 8.7: System of 140 lipid particles, pre-constructed as a bilayer patch

following advantages over Heller's model:

- The model is built based on the structured macro molecules instead of individual atoms, the computational complexity is largely reduced.
- Because the system we build is in two dimension instead of three dimension, the computational cost is much smaller, this afford us to study the structure of larger size patch of bilayer structures, and also simulate a much longer actual time period.
- Heller's did in [32] simulate the patch of bilayer structures of 200(100 on each side) lipid particles in three dimensions, which is equivalent to a 20 particle bilayer patch(10 particles on each side) system. Since actual membrane sizes are usually very large, we think that the study of a system composed of more lipid particles can provide more insight of bilayer structure properties.
- Heller's model simulated the actual time span of 263 ps ($1 \times 10^{-12} seconds$), which we think is not adequate to claim the actual system is stable under the simulated conditions. The system we construct was able to simulate a much larger time scale, and also proved to converge to a stable state.

Chapter 9

Interactive Simulation Tool

We think that a very useful aspect of the project is to provide a simulation tool that can users the ability to modify conditions for different systems. Currently we are developing a interactive graphics user-interface program to provide a more convenient simulation tool for future development of this project and for other people to use. This is mainly an effort of graphics interface programming. The program provides an interface for user to set and adjust the parameters of the interactions (Figure 9.1), the program displays the ramp function from the parameters accordingly. The final selected parameters are saved in a file that can be read in by the simulation program.

Figure 9.1: A user-interface program for parameter selection.

Chapter 10

Conclusion and Future Directions of the Research

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, as well as the formation of monolayer structure at an air-water interface. The test of the model on a pre-constructed bilayer membrane patch is a successful first step toward further study of membrane properties such as membrane-protein interactions. For example, we can model protein structure in a similar fashion, and simulate the interactions between protein and lipid bilayer structures.

This model may 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.

Bibliography

[1] E. J. Ambrose and Dorothy M. Easty. *Cell Biology*. Thomas Nelson and Sons Ltd., 36 Park Street, London WIY 4DE, 1970.

[2] Pascale Angelica Bachmann, Pier Luigi Luisi, and Jacques Lang. Autocatalytic self-replicating micelles as models for prebiotic structures. *Nature*, 357:57–59, 5 1992.

[3] Pascale Angelica Bachmann, Peter Walde, Pier Luigi Luisi, and Jacques Lang. Self-replicating micelles: Aqueous micelles and enzymatically driven reactions in reverse micelles. *J. Am. Chem. Soc.*, 113(22):8204–8209, 1991.

[4] Isabel Baeza, Miguel Ibá nez, J. Carlos Santiago, Carlos Argüello, Carlos Wong, and J. Oró. Diffusion of mn^{2+} ions into liposomes mediated by phosphatidate and monitored by the activation of an encapsulated enzymatic system. *Journal of Molecular Evolution*, 1990.

[5] Michael W. Berns. *Cells*. Saunders College Publishing, 383 Madison Avenue, New York, NY 10017, second edition, 1983.

[6] 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.

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

[8] C.I.Simionescu, F. Dénes, and M. Totolin. The appearance of protobiopolymers and protomembrances in accordance with the "cold theory" of the origins of life. In Y. Wolman, editor, *Origins of Life*, pages 263–269. D. Reidel Publishing Company, 1981.

[9] Mark Dalton. Www cell biology course.

[10] Charles Darwin. *The Origin of Species*. Random House, 1993.

[11] François David. Random surfaces and the statistics of membranes. *Physics Reports*, pages 221–227, 1989.

[12] Christian de Duve. *A Guided Tour of The Living Cell*, volume 1. Scientific American Library, 1984.

[13] Christian de Duve. *A Guided Tour of The Living Cell*, volume 2. Scientific American Library, 1984.

[14] David W. Deamer and Gail L. Barchfeld. Encapsulation of macromolecules by lipid vesicles under simulated prebiotic conditions. *Journal of Molecular Evolution*, 1982.

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

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

[17] Linglan Edwards and Yun Peng. Computational models for the formation of protocell structures. In Christoph Adami, Richard K. Belew, Hiroaki Kitano, and Charles E. Tylor, editors, *Artificial Life VI, proceedings of the Sixth International Conference on Artificial Life*. Massachusetts Institute of Technology, 1998.

[18] Linglan Edwards and Yun Peng. Computer simulation of self-assembly of lipid aggregates. In C. Wilke, S. Altmeyer, and T. Martinetz, editors, *Third German Workshop on Artificial Life*. Verlag Harri Deutsch, Frankfurt am Main, 1998.

[19] Linglan Edwards, Yun Peng, and James A. Reggia. Computational models for the formation of protocell structures. *Artificial Life*, 1998.

[20] T.N. Evreinova, V.N. Karnaukhov, E.V. Melnikova, and B.L. Allahverdov. Coacervate drops—primitive forms of precells. *Origins of Life*, pages 369–374, 1978.

[21] T.N. Evreinova, T. W. Mamontova, V.N.Karnauhov, S.B. Stephanov, and U.R. Hrust. Coacervate systems and origin of life. *Origins of Life*, pages 347–353, 1974.

[22] J. Doyne Farmer and Alletta d'A. Belin. Artificial life: The coming evolution. *Artificial Life II, SFI Studies in the Sciences of Complexity*, 1991.

[23] Sidney W. Fox. The evolutionary significance of phase-separated microsystems. *Origins of Life*, pages 49–68, 1976.

[24] Sidney W. Fox. The emergence of living functions from molecules. In B. Gruber and J. H. Yopp, editors, *Symmetries in Science iV*, pages 53–66. Plenum Press, 1990.

[25] Sidney W. Fox and Kaoru Harada. Thermal copolymerization of amino acids to a product resembling prote in. In David W. Deamer and Gail R. Fleischaker, editors, *Origins of Life, the central concepts*, page 189. Jones and Bartlett Publishers, One Exeter Plaza, Boston, MA 02116, 1994.

[26] Oliver Gillie. *The Living Cell*. Funk & Wagnalls, A Division of Reader's Digest Books, Inc., 1971.

[27] David J. Griffiths. *Introduction to Electrodynamics*. Prentice-Hall, Inc., Englewood Cliff, New Jersey 07632, 1981.

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

[29] 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*, pages 205–207. Jones and Bartlett Publishers, One Exeter Plaza, Boston, MA 02116, 1994.

[30] William R. Hargreaves and David W. Deamer. Origin and early evolution of bilayer membranes. In D. W. Deamer, editor, *Light Transducing Membranes, structure, func-*

tion and evolution, pages 23–59. Academic Press, Inc., 111 Fifth Avenue, New York, New York 10003, 1978.

[31] Roger Harrison and George G. Lunt. *Biological Membranes, their structure and function*. Blackie, second edition, 1980.

[32] Helmut Heller, Michael Schaefer, and Klaus 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.

[33] Jacob N. Israelachvili. *Intermolecular and Surface Forces*. Academic Press, 1991.

[34] Mahendra Kumar Jain. *Introduction to Biological Membranes*. John Wiley & Sons, 1988.

[35] Gerald Karp. *Cell Biology*. McGraw-Hill Book Company, 1979.

[36] K.L.Gladilin, A.F. Orlovsky, D.B.Kirpotin, and A.I. Oparin. Coacervate drops as a model for precellular structures. In *Origins of Life*, pages 357–362. 1978.

[37] Stanislas Leibler. An introduction to the statistical mechanics of fluctuating membranes. *Physics Reports*, pages 265–270, 1989.

[38] Stanislas Leibler. Computational physics of micelles and membranes. *Nature*, 348:586–587, 12 1990.

[39] Reinhard Lipowsky. The conformation of membranes. *Nature*, 349, 2 1991.

[40] Harvey Lodish, David Baltimore, Arnold Berk, S. Lawrence Zipursky, Paul Matsudaira, and James Darnell. *Molecular Cell Biology*. Scientific American Books, third edition, 1995.

[41] 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.

[42] Harold J. Morowitz. *Beginnings of Cellular Life*. Yale University Press, 1992.

[43] Hiroyuki Okihana and Cyril Ponnamperuma. Functions of the coacervate droplets. *Origins of Life*, pages 347–353, 1982.

[44] L. E. Orgel. *The Origins of Life, Molecules and Natural Selection*. John Wiley & Sons, Inc., 1973.

[45] J. Oró and J. M. Berry. Comets and life. *Adv. Space Res.*, 7(5):23–32, 1987.

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

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

[48] J. Oró, E. Sherwood, J. Eichberg, and D. Epps. Formation of phospholipids under primitive earth conditions and the role of membranes in prebiological evolution, 1978.

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

[50] C. Ponnamperuma, Y. Honda, and R. Navarro-González. Asymmetry and the origin of life. In B. Gruber and J. H. Yopp, editors, *Symmetries in Science iV*, pages 193–203. Plenum Press, 1990.

[51] M. G. Rutten. *The Geological Aspects of the Origin of Life on Earth*. Elsevier Publishing Company, 52 Vanderbilt Avenue, New York, New York 10017, 1962.

[52] M. G. Rutten. *The Origin of Life, by Natural Causes*. Elsevier Publishing Company, 52 Vanderbilt Avenue, New York, New York 10017, 1971.

[53] Raymond A. Serway. *Physics For Scientists & Engineers with Modern Physics*. Saunders Golden Sunburst Series, third edition, 1990.

[54] Phillip Sheeler and Donald E. Bianchi. *Cell Biology, Structure, Biochemistry, and Function*. John Willey & Sons, Inc., second edition, 1983.

[55] Qing Sheng and Klaus Schulten. Molecular dynamics simulation of immobilized artificial membranes. *J. of Phys. Chem.*, 99(27):11018–11027, 1995.

[56] S.J. Singer. The proteins of membranes. *Journal of Colloid and Interface Science*, 58(3), 3 1977.

[57] Charles B. Thaxton, Walter L. Bradley, and Roger L Olsen. *The Mystery of Life's Origin: Reassessing Current Theories*. Lewis and Stanley, 1994.

[58] Richard S. Young. Morphology and chemistry of microspheres from proteinoid. In S. Fox, editor, *The Origins of Prebiological Systems and of Their Molecular Matrices*, pages 347–357. Academic Press, 1965.

[59] Feng Zhou and Klaus Schulten. Molecular dynamics study of a membrane—water interface. *J. of Phys. Chem.*, 99(7):2194–2207, 1995.