MOLECULAR MODELING OF PROTOCELLULAR FUNCTIONS

Andrew Pohorille[a,b], Christophe Chipot[b†], Michael H. New[a,b], and Michael A. Wilson[a,b]

[a] *Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, California 94143*

[b] *Planetary Biology Branch, NASA — Ames Research Center, MS 239-4, Moffett Field, California 94035-1000*

The mechanisms of three protocellular functions have been studied using molecular modeling techniques. These functions are (1) the transport of ions across membranes, (2) the formation of photoactivated proton gradient that could drive chemical synthesis in the protocell, and (3) the organization of small peptides necessary for catalytic activity. In all these processes, membranes play an essential role. The transfer of ions across the barrier formed by protocellular walls is facilitated by the formation of deep, thinning defects in the membrane. Membranes also form a barrier to charged species that allows for retaining proton gradients. These gradients can be generated by a simple transmembrane proton pump consisting of a proton source and two acceptors. The directionality of the pump is ensured by a "gate-keeping" mechanism involving a water molecule, conformational change of the primary acceptor or tautomerization of a histidine. The pump can be formed by two transmembrane helices but not one helix. They provide surfaces at which organic molecules concentrate and small peptides can organize into ordered, amphiphilic structures. In general, valuable information about the origins and evolution of protocells can be obtained from the knowledge of physical and chemical principles that govern functioning of contemporary cells.

1 Introduction

The emergence of cellular life was a central event on the evolutionary pathway from simple organic matter to present-day life forms. In this fundamental step, organic material assembled into boundary structures which acquired metabolism and the capabilities to replicate and evolve.

Any direct record of this stage of evolution is lost. At the present time, it seems unlikely that we will be able to recreate its details with a reasonable degree of certainty. This immediately raises the question of how can we reliably study the formation of cellular life and, even further, whether it is at all possible. It has been suggested by Monod[1] that life originated from a series of highly unlikely events. Since specific conditions leading to these events might

†on leave from: *Laboratoire de Chimie Théorique, U.R.A. C.N.R.S. n^0 510, Université Henri Poincaré-Nancy I, BP. 239, 54506 Vandœuvre-lès-Nancy Cedex — France*

never be known, one might conclude that the origin of life is not an appropriate subject of scientific inquiry. An alternative view holds that the emergence of life was a reasonably robust event and its main aspects can be understood, although perhaps not in detail, from basic physical and chemical principles, once we account for different, known limitations to this event.

One type of limitation arises from the environmental conditions on the early Earth. A different set of limitations is imposed by our knowledge of the only known successful "experiment" in the origin of life — the living cell. We require that the earliest precursors of cells — protocells — were capable of performing ubiquitous cellular functions, utilizing only those simple molecules which could have existed under prebiotic environmental conditions. We further restrict the processes employed by protocells to those for which a plausible evolutionary pathway into contemporary cellular functions can be postulated. This condition is motivated by the continuity argument [2] that the evolution of cellular structures progressed without undergoing discontinuous transitions.

The assumptions of robustness and continuity are essential for methodological purposes. It is possible that certain steps in the formation of protocells required highly specific environmental conditions the existence of which cannot be proven. It is also possible that protocellular evolution produced dead ends or intermediate structures which left no trace in contemporary cellular functions. However, abandoning the assumptions of robustness and continuity could easily lead to a slippery path of speculations, not amenable to testing. On the other hand, by adhering to these assumptions we can base studies of protocells on the firm foundations formed by our broad knowledge of the structure and functions of contemporary cells.

In this paper we discuss how we can shed light on the functioning of a primitive cell through the application of these concepts in large–scale, molecular–level computer simulations. Computational methods have a unique role in studies of the origin of life. Their goal is to identify the structural and energetic conditions emerging from the fundamental principles of physics and chemistry that successful models of protocellular functions must fulfill [3]. Thus, they establish physico–chemical boundaries for these models which, in turn, provide guidance for laboratory experiments.

In the next section we present the main ideas underlying the concept of the protocell. Then, we briefly describe the computational methods and molecular models used in our studies. The next three sections are devoted to transport of ions across membranes, the formation of proton gradients that can be utilized as an energy source, and the organization of small peptides at membrane interfaces for primitive catalysis. The paper closes with conclusions and suggestions for future research.

2 A Model of the Protocell

What computer simulations of protocellular functions are relevant? The answer to this question depends upon the conceptual model of the protocell. Indeed, for different models of protocells, different sets of laboratory experiments and computer simulations are significant. To define a model that is consistent with both prebiotic condition and contemporary cellular life we have to address several fundamental questions:

What were the environmental conditions of early Earth? The early Earth between 4.4 and 3.9 Ga was the subject of bombardment by large size impactors (comets, meteorites) which delivered sufficient energy to evaporate early oceans. It is likely that life, even if it had formed during that period, could not have survived these impacts. On the other hand, there is compelling evidence of bacterial life from sediments 3.5-3.57 Ga old [4]. These estimates provide time boundaries for the formation of life.

Besides their detrimental effects, early impacts also benefited the origin of life by delivering water, volatiles and organic molecules to Earth [5]. Volatiles from extraterrestrial sources augmented the Earth's atmosphere, which was mostly composed of N_2, CO_2 and CO but did not contain any significant amounts of O_2. This atmosphere was not very conducive to the formation of compounds considered to be good precursors of biological molecules, such as NH_3 and HCN. In an atmosphere of this composition, the greenhouse effect kept surface temperatures above the current level, over 40°C and perhaps as high as 80-100°C [6].

In summary, according to current theories, life formed relatively quickly on geological time scales, perhaps as fast as 10^5-10^7 years at elevated temperatures in seawater or at aqueous interfaces and in a neutral or, possibly, reducing atmosphere [7].

Was the formation of cellular boundary structures possible in prebiotic conditions? In analogy with contemporary cells, ancestral cells must have been closed structures with an aqueous interior separated from the environment by walls built of amphiphilic molecules. These molecules were assembled into bilayer membranes such that their polar parts were in contact with water whereas their nonpolar portions were buried in the membrane interior. Such structures have several features that would have been highly desirable in prebiotic conditions. Amphiphilic molecules spontaneously accumulate at water–air and water–oil interfaces and, at sufficient concentrations, self–assemble into boundary structures, called vesicles, by agitation or cycles of wetting and drying [8]. Vesicles can grow and divide by acquiring additional amphiphilic material. The concentration of amphiphilic material, and the self–assembly and stabil-

ity of vesicles are robust phenomena which occur over a fairly broad range of environmental conditions and molecular compositions.

While efficient synthetic pathways to obtain amphiphilic molecules under prebiotic conditions have not been established (a difficulty common to all cellular components) it has been demonstrated that highly heterogeneous mixtures of amphiphiles extracted from the Murchison meteorite can form vesicles [9]. This points to extraterrestrial infall as one possible source of the membrane–forming material and underscores the potential protobiological significance of vesicles.

What were the functions of protocells? The simplest, "minimal" cell [10] must have performed several essential functions, such as (a) capturing and transducing energy, (b) sequestering organic matter and ions from the environment, (c) catalyzing the synthesis of its components from the captured material, (d) protecting organic matter accumulated in its interior from dilution in the surrounding water, and (e) self–replication.

In contemporary cells, enzymatic catalysis and bioenergetics are accomplished mostly by proteins inside the cell or embedded in the cell membrane whereas genetic information is transferred during replication by nucleic acids which, in turn, are synthesized by proteins. The discovery that nucleic acids also possess some catalytic activity led to a hypothesis that the current division of functions between proteins and nucleic acids was preceded by an "RNA world" [11]. However, synthetic pathways for nucleic acids are among the most difficult to postulate under protobiological conditions. This raises a distinct possibility that protocells initially represented the "pre-RNA world" in which cellular functions were performed, probably with low efficiency and specificity, by simple molecules, including possible precursors to proteins (*i.e.* peptides) and nucleic acids (see for instance Wächterhäuser [12]). Finding *how* these functions might have been accomplished and what molecules might have been involved is one of the main challenges in studies of protocellular life.

Where did protocellular functions evolve? Studies of protocellular evolution can be based on two fundamentally different hypotheses. According to one hypothesis, cellular structures and functions evolved *simultaneously* in the protocell. This can be contrasted with an alternative hypothesis that these structures and functions evolved in different environments, such as mineral surfaces, and were incorporated into vesicles at some later stage of evolution, forming a functioning cell. Intermediate scenarios are also possible — some functions were always associated with protocells while others initially evolved separately. The two extreme hypotheses can be distinguished by observing that both structures and functions sensitively depend on the molecular environment. For example, contemporary bioenergetics is so closely connected

with membranes that it would be difficult to imagine its evolution in a different environment. Conversely, any possible alternative mechanism for acquiring and utilizing energy that might have developed outside the cell would have to remain functioning once encapsulated in the protocell, despite environmental change. This argument favors the hypothesis about cellular evolution of at least some functions. A corollary to this argument is that special attention should be focused on unique properties of protocells that distinguish them from other environments and, in particular, on the role of membranes.

3 Methods and Models

The most direct method to simulate protocellular functions at a molecular level is molecular dynamics [13]. In this method, Newton's equations of motion for all the atoms in the system are solved as a function of time. Exactly the same approach is commonly used to study computationally large systems in chemistry and structural biology.

The first task is to define a protobiologically relevant, yet computationally tractable, model system. As a suitable model for membrane forming material we selected glycerol monooleate (GMO). A GMO molecule is composed of a glycerol head group linked by an ester bond to a hydrocarbon tail containing 18 carbon atoms with a double bond in the middle. Undoubtedly, the GMO bilayer does not faithfully represent the composition of protocellular membranes which, most likely, were built of highly heterogeneous amphiphilic material that contained both charged and uncharged molecules. However, in contrast to GMO membranes, such heterogeneous systems cannot be reliably modeled by computational methods. Even though the GMO bilayer has a homogeneous composition, it still retains important features expected of primitive membranes, namely structurally simple head groups and a highly fluid interior. In these respects it may be a better protocellular model than membranes built of phospholipids that form the walls of contemporary cells. This point is further reinforced by difficulties in identifying sufficient sources of phosphate on the early Earth [14] to allow for the formation of significant quantities of phospholipids.

Present computational resources do not permit simulating a whole vesicle in an aqueous environment. Instead, we considered only a part of this system — a fragment of the membrane consisting of 72 GMO molecules and covering an area of 37×37 Å^2, surrounded by 2300 water molecules evenly distributed on both sides of the bilayer. Undesirable discontinuities at the edges of the system were eliminated by applying periodic boundary conditions. This system is sufficiently large to study such phenomena as transport, organization

and chemical reactions of molecules at water–membrane interfaces, and the formation of transmembrane proton gradients.

The equations of motion describing the system were solved numerically on a step–by–step basis using a finite difference method. From dynamical information about the system at time t, we obtained the positions and velocities of all the atoms at time $t + \delta t$ that, in turn, were used to calculate these quantities at $t + 2\delta t$, *etc.* This procedure was repeated many times, resulting in a complete, microscopic description of the system as a function of time (called a trajectory). Macroscopic quantities can also be obtained from this trajectory. In our simulations, δt was set between 2×10^{-15} and 5×10^{-15} sec and the length of trajectories varied between 10^{-9} and 10^{-8} sec. Thus, each trajectory required approximately 10^6 time steps.

In each step, the forces acting on each atom in the system have to be calculated. These forces are derivatives of the potential energy function with respect to the atomic coordinates. The potential energy function was represented as a sum of contributions from electrostatic and van der Waals interatomic interactions as well as terms describing intramolecular bond and angle vibrations and changes in the dihedral angles formed by three consecutive bonds. Electrostatic contributions were evaluated as a sum of Coulomb energies between partial point charges assigned to atoms. For water, the TIP4P potential energy function [15] was used. Potential energy functions for GMO were developed by Wilson and Pohorille [16] and for peptides by Cornell *et al.* [17].

Despite large computational effort, the time scale covered by molecular dynamics simulations remains quite short. The probability of observing processes that typically occur at considerably longer time scales in such simulations is very low. However, reliable structural and energetic information about these processes can still be obtained by dividing them into several consecutive stages that are simulated separately. For example, solute transport across the membrane can be represented as a series of stages in which the solute is progressively moved across the membrane in the direction perpendicular to the bilayer (the z–direction). At each stage, the solute is constrained to lie within a "window" along z. The free energy change at each stage, $\Delta A(z)$, as a function of the parameter z can be obtained by observing the probability, $P(z)$, of finding the solute at z:

$$\Delta A(z) = -k_b T \, lnP(z) \tag{1}$$

where k_b is the Boltzmann constant and T is the temperature of the system. If the ranges of z explored by the solute in consecutive stages overlap, the dependence of ΔA on z for the whole process can be obtained from the condition that the free energy must be a continuous function of z. Exactly the

same procedure can be used to study other processes of interest described by different parameters.

4 Unassisted Ion Transport Across Membranes

Protocellular walls must have been permeable ions. Ions were needed for such cellular functions as bioenergetics based on charge separation across cell walls, stabilization and self–assembly of a variety of molecular structures, and chemical catalysis. Also, ion transport stabilized the protocell against osmotic pressure.

Transport of ions requires that a charged species be moved from a polar, aqueous environment into the nonpolar interior of the membrane. This process is associated with a large activation barrier. In contemporary cells, ion transport is aided by specialized, complex molecules (ion channels and ion carriers) which help to lower this barrier. However, these molecules are usually too complex to have been present in protocells, at least at the earliest stage of their evolution. Therefore, less complicated mechanisms of ion transport must be examined, among which unassisted transport is the simplest.

The activation barrier and, subsequently, ionic permeability can be readily estimated from the dielectric continuum model in which both the water and the membrane are described as continuous media characterized by their dielectric constants ϵ_w and ϵ_m, respectively. The bilayer is represented as a rigid layer of a fixed width, d. Then, the free energy, ΔA of moving a spherical ion of radius a and charge q from bulk water into the center of the bilayer is [26]:

$$\Delta A = \frac{q^2}{2}\left[\frac{1}{\epsilon_b a} - \frac{1}{\epsilon_w a}\right] - \frac{q^2}{\epsilon_b d}\, ln\left(\frac{2\epsilon_w}{\epsilon_w + \epsilon_b}\right). \qquad (2)$$

For $\epsilon_w = 77.4$, $\epsilon_b = 2$ [27,28], $q^2 = 1$, $a = 1.68$ Å (the radius of Na^+) and $d = 35$ Å, $\Delta A = 45$ kcal/mol. This yields a permeability of 10^{-27} cm sec^{-1}. These results are insensitive to the width of the bilayer; ΔA varies from 42 to 47 kcal/mol for d between 20 and 80 Å.

The permeability predicted by the simple dielectric continuum model is extremely low, approximately 13-15 orders of magnitude lower than the values measured for small ions permeating model phospholipid bilayers [29] This large difference cannot be eliminated by reasonable adjustments of the parameters in Eq. 2. Instead, it appears that the mechanism of ion transport is incorrectly described in the model. Several alternative mechanisms have been proposed focusing on ion hydration [29], defects in the membrane [32,31,33] and ordering of hydrocarbon tails [34].

Computer simulations of the transfer of Na^+ and Cl^- through the GMO bilayer [35] reveal the actual mechanism of unassisted ion transport. As the ion moves into the membrane, polar head groups on the incoming side of the GMO bilayer follow by tilting inwards, thereby creating a thinning defect in the membrane filled with water. Once the ion crosses the mid–plane of the bilayer, the defect on the incoming side of the membrane disappears and, instead, a similar deformation is formed on the outgoing side. This is illustrated in Fig. 1.

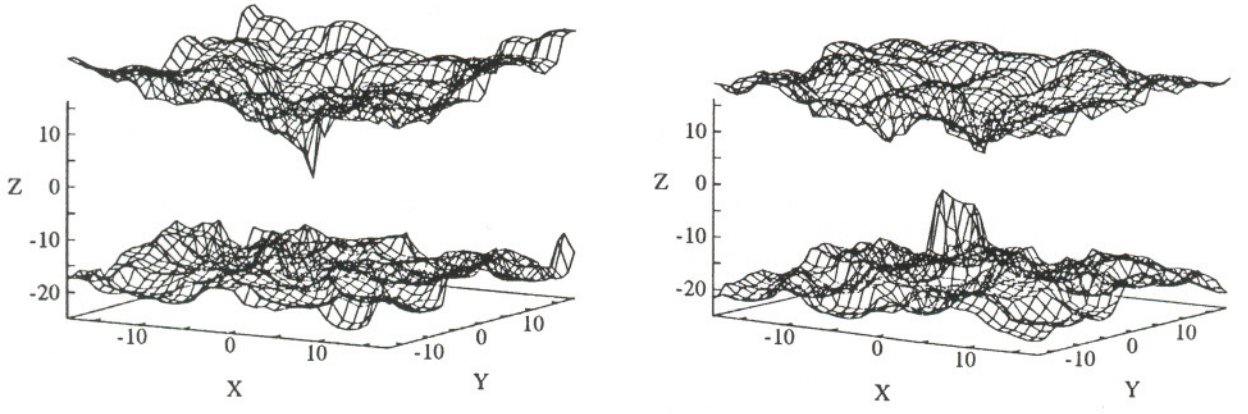


Figure 1: Instantaneous position of the water surface as Na^+ crosses the membrane (left) as the ion enters the membrane from above, creating a downward bulge in the water surface and (right) as the ion leaves the membrane from below where an upward bulge has formed in the lower surface.

The ion inside the membrane looses some of its hydration shell but this loss is compensated by near–neighbor interactions with oxygen atoms from GMO head groups. As a result, the total solvation number, defined as the average number of oxygen atoms from water molecules and GMO head groups around the ion, remains constant throughout the whole transport process. Both the formation of local, asymmetric defects in the bilayer and partially solvation of the ion reduce the activation barrier to charge transfer. These features are not captured in the simple dielectric model. For Na^+, the calculated decreases to ΔA to 26 kcal/mol yielding the permeability of approximately 10^{-13} cm sec^{-1}. This agrees well with the experimental value [29]. For Cl^-, the calculated permeability is 1-2 orders of magnitude higher than for Na^+, also in agreement with experiment [29].

The mechanism of unassisted ion transport focuses attention on an important property of lipid bilayers, namely their ability to deform, even in the absence of ions, from a rigid planar structure, such that the local width of the membrane fluctuates in time and space. This property was studied in detail

in computer simulations of the pure water–GMO system [16]. It was shown that the probability of forming thinning defects decreases exponentially with their depth. This implies that the permeability of thin membranes to ions should be considerably more sensitive to the membrane width than predicted from Eq. 2. This conclusion is consistent with recent measurements of the ionic permeabilities of phosphatidylcholines with varying hydrocarbon chain lengths [36] The observed sensitivity might lead to establishing limits on the width of protocellular walls; membranes that were too thin would not provide an effective barrier to ions while membranes that were too wide would be practically impervious to charges.

5 Directional Proton Transport

Contemporary cells utilize a variety of complex mechanisms for energy acquisition and transduction. A common motif, however, is the conversion of acquired energy into a proton gradient that is then used to do useful work [37], such as the synthesis of "high energy" compounds. The universality of this mechanism suggests that it must have emerged at an early stage of protobiological evolution.

The creation and maintenence of a transmembrane proton gradient requires a system that irreversibly transports protons across the protocellular boundary. There are several possible early sources of protons, including chemical reactions and light. Probably the simplest system which is capable of photo–generating a transmembrane proton gradient consists of polycyclic aromatic hydrocarbons incorporated into liposome membranes [38]. Upon the absorption of light, the chromophore releases protons either to the exterior or the interior of the liposome. Protons in the environment dissipate whereas those inside the liposome accumulate, thereby creating a proton gradient. This system, however, is not directional; it lacks a "gate–keeper" mechanism which ensures that protons transferred to the protocell interior are not used to regenerate the protonated chromophore.

A schematic of a simple directional proton transport system is shown in Fig. 2. In this figure, PS refers to a proton source located near the center of the membrane and A_1 and A_2 are a pair of proton acceptors that are part of the gate–keeper complex. The PS initiates the proton transport across the membrane. It could be comprised of a chemical reaction, a chromophore and an ionizable species [39] or, more simply, of a polycyclic aromatic hydrocarbon [38]. The only constraint placed on the PS is that, when protonated, it transfers its proton to A_1. This might not only require that A_1 be in close proximity to PS, but also that the pK_a of A_1 be coupled to the state of PS. The secondary

proton acceptor, A_2, is located sufficiently close to the protocell interior that any proton it accepts is quickly released to the aqueous solution. After PS transfers its proton to A_1, the reverse reaction, $A_1H + PS^- \rightarrow A_1 + PS$, might become highly probable. However, all that is needed to create a proton gradient is a nonvanishing probability of irreversible proton transfer from A_1 to A_2. Protons transferred to A_2 would then be injected into the protocellular interior.

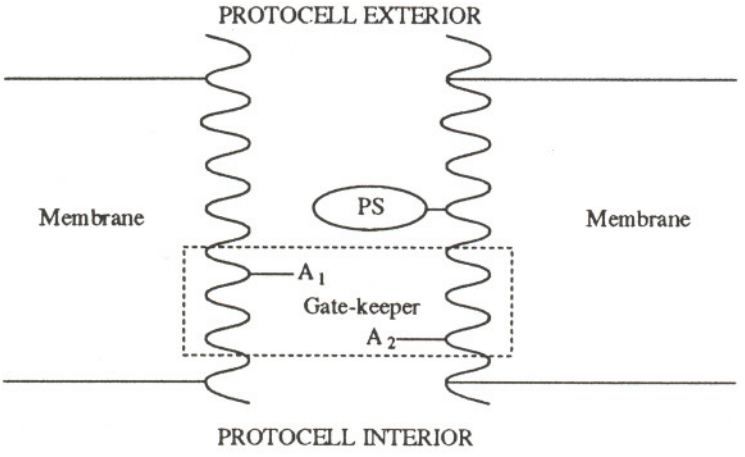


Figure 2: "Gate–keeper" scheme for directional proton transfer. PS is an activated proton source, A_1 and A_2 are proton acceptors that are part of the gate–keeper complex.

After the proton has been tranferred to the protocellular interior, PS must be reconstituted by a proton from the exterior if the proton gradient is to be maintained. Since the proposed proton pump is assymetrical, this can be accomplished if reconstitution from the exterior is faster than from the interior. After deprotonation, the charged PS alters the environment within the gate–keeper complex, slowing down the "uphill" proton transfer from the interior. Reconstitution of PS from the exterior would continue until the interior pH of the protocell drops below the pK_a of A_2, at which point the formation of protonated A_2 would begin to drive the back reaction.

Three different mechanisms have been considered for increasing the irreversibility of the transfer of a proton from A_1 to A_2. One mechanism involves a transient water bridge. Transient chains of hydrogen–bonded water molecules have been postulated to account for the anomalously high proton permeability of membranes [40] and for the proton conductivity of the gramicidin channel [41]. Furthermore, *ab initio* quantum mechanical calculations on the formic acid–water–formate system (see Fig. 3) have revealed that the barrier for proton

transfer from a formic acid to a formate ion across a water bridge is only 0.7 kcal/mole[42]. With a pK_a of 3.7, formic acid is a good model for the acidic amino acids glutamate and aspartate which have pK_as in aqueous solution of 3.95 and 4.4, respectively. These two amino acids are good candidates for A_1 and A_2. A transient water bridge between them would then provide an efficient mechanism for irreversibly transferring a proton. The back transfer of the proton is impeded by the free energy required to move the proton uphill towards PS, as well as by the disruption of the water bridge resulting from the hydration of the negatively charged A_1^- moiety. The efficiency of this mechanism will decrease if the pH inside the protocell becomes low enough to cause the spontaneous formation of protonated A_2, driving the back reaction.

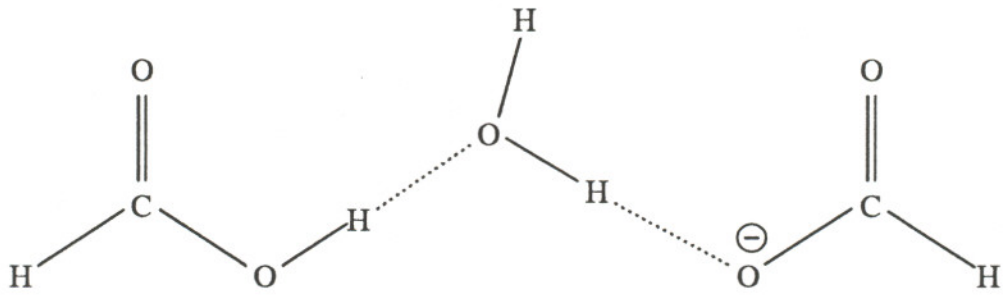
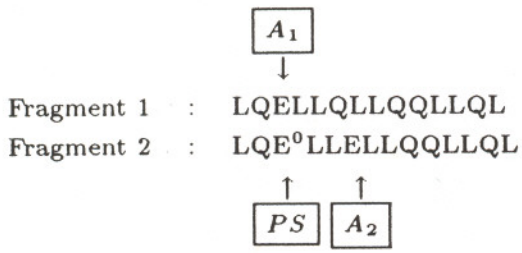


Figure 3: A sketch of the water–bridged formic acid–water–formate proton transfer system.

In contemporary cells, transmembrane helix bundles are the most common motif for proteins conducting proton transfer. Their simplest precursor might have been a system consisting of a proton donor and an acceptor affixed to a single transmembrane α–helix. However, geometrical considerations rule out this arrangement — all conformations of side chains that bring the proton donor into the proximity of the acceptor, with or without a water bridge, are sterically disallowed.

A somewhat more complex, two–helix system does not suffer from similar steric restrictions. The feasibility of the water–bridge mechanism was tested by constructing a pair of transmembrane helix fragments with sequences conducive to this type of proton transport. At pH 7, only three amino acids could serve as proton acceptors: aspartate (D), glutamate (E) and histidine (H). We used a glutamic acid (E^0) as the proton donor and glutamate for both A_1 and A_2. We further assumed that the pK_a of the glutamic acid was coupled to a chromophore so that the photo–excited chromophore induced the glutamic acid to ionize. The rest of the helical fragments was constructed from nonpolar leucine (L) and polar, nonionizable glutamine (Q). In the folded fragments,

the leucines were in contact with the nonpolar membrane while the glutamines formed a polar core within which the proton transfer could occur. The sequences of the two helix fragments studied are:



The helix fragments were arranged so that the glutamate on fragment 1 (A_1) and the glutamic acid on fragment 2 (proton donor of the PS) formed a bifurcated hydrogen bond. A_1, in turn, is separated from the A_2 glutamate by approximately 5Å. While this distance is too far for direct proton transfer between the two glutamates, it is nearly ideal for a water–bridged transfer.

Another mechanism to prevent back transfer is a conformational shift of A_1 after it is deprotonated. To examine the magnitude of the conformational shift accompanying deprotonation in a simple system, we considered two helix fragments differing only in the protonation state of the first glutamate:

Fragment 3 : LQE°LLEL
Fragment 4 : LQELLEL

Minimization of each protein resulted in different side chain conformations of the glutamate. The oxygen initially bearing the proton, $O_{\epsilon 2}$, is displaced by 4.8Å after deprotonation. The two relevant torsional angles of the glutamate side chains, $NC_\alpha C_\beta C_\gamma$ and $C_\alpha C_\beta C_\gamma C_\delta$, change upon deprotonation from 72 and -176 to 63 and -83, respectively, upon deprotonation. The displacement of the deprotonated carboxyl oxygen is large enough to disrupt a proton transfer chain. When coupled with a transient water bridge, it would ensure the irreversibility of the $A_1 \rightarrow A_2$ proton transfer. This coupling of a conformational shift with a hydrogen–bonded chain of proton acceptors to generate irreversibly a transmembrane proton gradient is thought to be used by bacteriorhodopsin [43].

The final gate–keeper mechanism is a generalization of the water–bridged mechanism whereby the water molecule acts as an amphiprotic species — a species that can both accept and donate protons. A similar gate–keeper mechanism could be constructed from other amphiprotic species. For example, the

two ring nitrogen atoms of histidine can both accept protons, although at pH 7 only one of them is protonated. Proton transfer to the second nitrogen, occurs with a pK_a of slightly greater than 7. Proton transfer would proceed from histidine to a secondary acceptor and the back reaction would be prevented by tautomerization of the histidine ring. This mechanism would be very sensitive to the exact nature of the PS since the pK_a's of the two nitrogen atoms are very sensitive to the local environment.

Modern proton pumps are thought to utilize a complex chain of hydrogen–bonded residues as well as internal isomerizations for irreversible proton transport across cell membranes. We have demonstrated that these same mechanisms could have been used by potential precursors of these pumps to facilitate directional proton transport across the protocellular boundary. The structures needed to perform this function are simple and not highly specific and, therefore, are compatible with protocellular conditions.

6 Organization of Peptides at Membrane Interfaces

Even the simplest protocell must have had the capability to catalyze the chemical reactions needed for its survival and growth. One group of potential early catalysts were peptides — possible precursors of contemporary protein enzymes. In modern enzymes, catalytic activity almost invariable depends upon the structure into which the protein folds which, in turn, depends upon the specific sequence of the amino acid residues along the protein backbone. This poses two problems for peptides to act as protocellular catalysts: First, in the absence of information molecules, high sequence specificity of peptides could not have been required for their catalytic activity. Second, short peptides typically do not exhibit secondary structure in aqueous solution and, therefore, do not appear to be suitable candidates for protoenzymes. There is, however, a growing body of evidence that peptides, which are disordered in water, acquire secondary structure at water–air or water–membrane interfaces if they have a proper sequence of polar and nonpolar residues. Structures that are stable at the interface are amphiphilic; polar residues are immersed in water and nonpolar residues are exposed to air or the membrane interior. The specific identity of the residues is less important, a desirable property in the protocellular environment. All main elements of secondary structure — α–helix [45,46] β-sheet [45] and β–turn [49] — have been observed at aqueous interfaces, sometimes for peptides less than 10 residues long.

To examine the effect of sequence amphiphilicity on the secondary structure of simple peptides at aqueous interfaces, we studied two heptamers placed at the water–air interface. A similar behavior is expected at more complex

water–oil and water–membrane interfaces [53,54]. The peptides were composed of two residues, nonpolar leucine (L), and polar glutamine (Q). Their specific sequences were (LQQLLQL) and (LQLQLQL). These sequences were designed to maximize the amphiphilicity of an α–helix and a β–strand, respectively, by exposing their polar side chains to the aqueous phase and their nonpolar residues to the air (see Fig. 4).

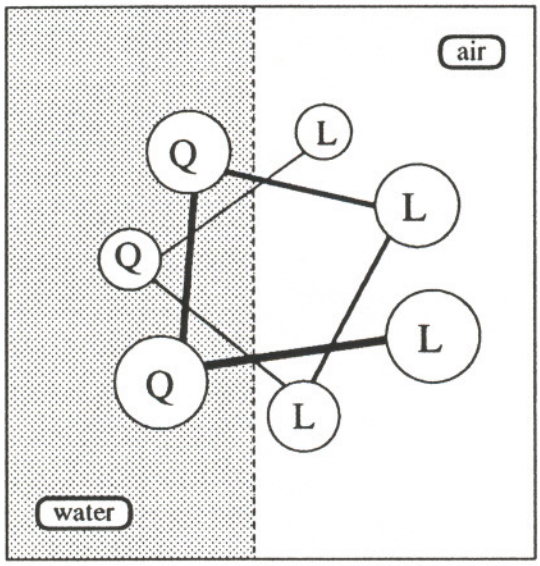


Figure 4: Axial projection of the (LQQLLQL) α–helix peptide at the water–air interface. Hydrophobic leucine residues are exposed toward the air, whereas hydrophilic glutamine residues are buried in the aqueous phase

In one set of molecular dynamics simulations, (LQQLLQL) and (LQLQL QL) were initially arranged at the water–air interface in amphiphilic secondary structures (α–helix and β–strand, respectively.) After approximately 3 ns. of molecular dynamics trajectory, both peptides remained at the interface. The α–helix remained stable, showing only small deviations from the initial set of angles (ϕ_i; ψ_i). The only exceptions were the tail–end leucine residues, for which an equilibrium between ($\psi_i \simeq -40°$) and ($\psi_i \simeq 150°$) was observed.

In contrast, fluctuations within the backbone of the β–strand were much larger, with several excursions of the ψ_i angles from *ca.* 150° to −40°. Examples of these fluctuations are shown in Fig. 5. This clearly reveals the instability of the β–strand for a single peptide molecule at the water-air interface. However, at higher concentrations of the peptide this conformation might still be favored due to the association of several molecules into β–sheets stabilized by a network on intermolecular hydrogen bonds. Since our system

contained only one peptide molecule this possibility was not explored.

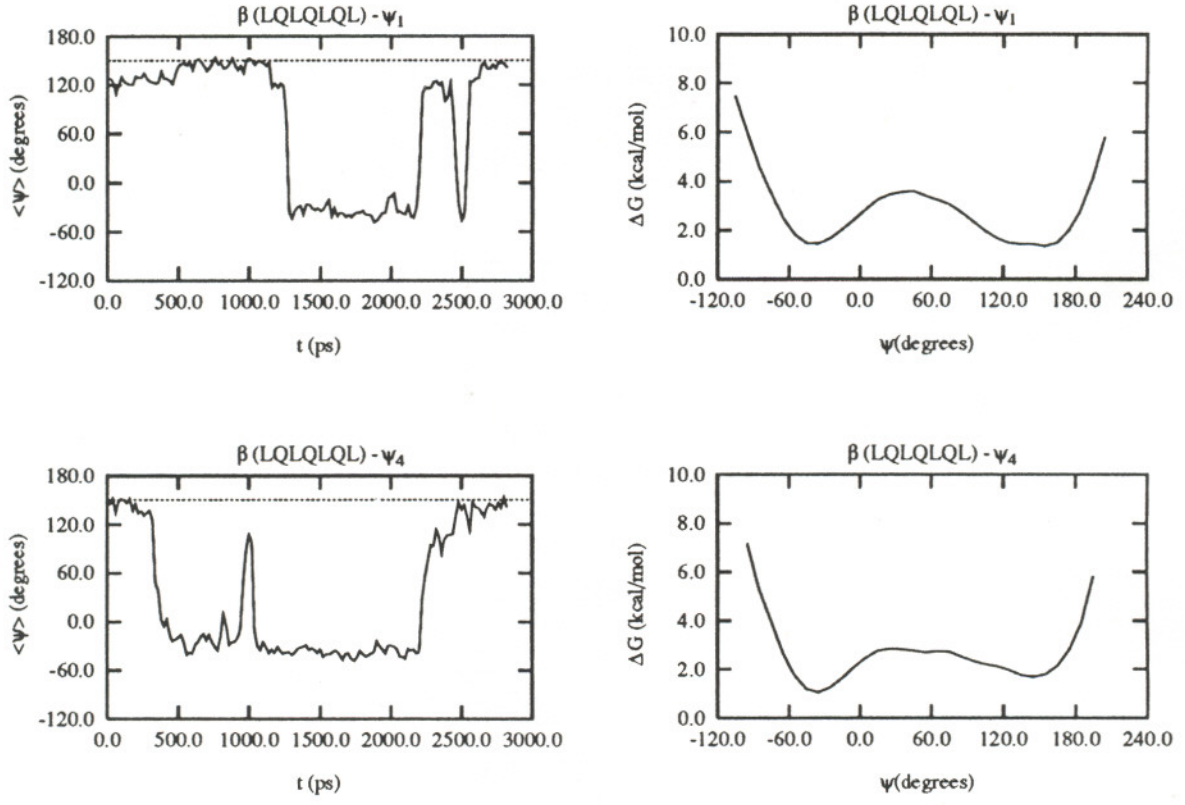


Figure 5: Time–history of ψ angles (left), and the corresponding free energy profiles (right), for residues 1 and 4 of the (LQLQLQL) β–strand

To investigate further the relationship between the sequence of peptides and their secondary structure at the interface, a second set of molecular dynamics simulations was performed. In these simulations, (LQQLLQL) was placed at the interface in the β–strand conformation and (LQLQLQL) was arranged as the α–helix. Neither of these initial structures was amphiphilic. A trajectory 15 ns. long was generated for this system.

During the course of simulations both peptides remained interfacially active. For (LQQLLQL), the first two ψ angles of the β–strand rapidly shifted to $-40°$. The resulting structure, although different from an α–helix, was rigorously amphiphilic. The initial, non-amphiphilic α–helix of (LQLQLQL) was also found to be unstable at the interface and never refolded to the β–strand.

It is clear that the current molecular dynamics simulations are not sufficient to describe *all* possible folding pathways of the two peptides. To do

so, the free energy as a function of backbone angles $(\phi_i; \psi_i)$ has to be fully explored using the approach outlined at the end of the method section (see Eq. (1)). This work is currently in progress.

The results of our simulations illuminate three important properties of small peptides at aqueous interfaces. First, peptides that contain both polar and nonpolar amino acids tend to accumulate at the interface. Second, amphiphilicity provides a strong force driving the peptides towards specific, organized structures. This force is absent in bulk media, such as water or the membrane interior. This tendency to organize at the interface, driven by the amphiphilicity of the structure rather than a specific sequence, is consistent with the concept of an active interface and might have been conducive to primitive catalysis under protobiological conditions. Finally, the degree of structural organization of the peptide backbone changes with the position in the sequence. The backbone is considerably more disordered at the ends of the peptide than in the middle.

The existence of secondary structure in membrane–bound peptides does not necessarily imply their catalytic activity, however. Only a few examples of such activity are known to to date. Peptides containing an alternating Leu-Lys sequence, which folded at the interface to the β-sheet geometry, have been shown to hydrolyze polyribonucleotides by the general acid–base mechanism[55]. A decapeptide exhibited almost the same activity as the polypeptide. In another example, a 14–residue peptide, which formed an amphiphilic α–helix, catalyzed decarboxylation of oxaloacetate[56]. In both cases, a binding pocket was not necessary to achieve catalytic activity. A simple active center was created by placing a small ligand (*e.g.* iron) between a bundle of four amphiphillic α–helices[57]. However, in general, the link between the interfacial structure of peptides and their catalytic activity remains largely unexplored.

7 Conclusions

The most direct and, perhaps, the only path to understanding the origins and earliest evolution of cellular life requires gaining advanced knowledge of contemporary cells and physico-chemical principles determining cellular organization and functions. This paper has been devoted to translating this idea into specific examples of protocellular functions through the application of molecular–level computer simulation methods. The ability to generalize the results for properly chosen models to other systems makes these methods particularly useful.

We have focused on the role of membranes as the main structures that distinguish a cell from other microenvironments. In particular, we stressed the

role of membranes as barriers to charged species and explored it in two different contexts. We considered possible activated mechanisms for the formation of a proton gradient across protocellular walls than could be further utilized as an energy source. We paid special attention to the directionality of this process which was essential to sustain the gradient and showed that a simple gate–keeping mechanism could be created by placing a proton donor and two proton acceptors at proper locations in the membrane. This mechanism might also involve a water molecule. One example of such a system consists of two transmembrane, helical peptides. We also investigated how simple ions could permeate cell walls and showed that this process is facilitated by highly flexible membranes. Ion transport is accompanied by the formation of thinning defects in the membrane which are needed to reduce the high activation barrier to this process.

Despite considerable experimental [14] and theoretical [3] progress in establishing a molecular basis of protocellular life, most aspects of this problem remain poorly understood. In particular, catalytic systems and mechanisms leading to synthesis of essential cellular components (*e.g* amphiphiles, short peptides) are not known. Also, a specific, directional mechanism of capturing and utilizing energy has not yet been demonstrated. Solving these two problems are among the most important and challenging tasks in studies of protocellular life. Based on rapid increase in computational capabilities and fast progress in understanding membrane–based functions, we anticipate that computer simulations will be very useful to this end.

Acknowledgments

This work was supported by a grant from the NASA Exobiology Program. C.C. was supported by a National Research Council Associateship. Computer resources for this work were provided in part by the Numerical Aerodynamics Simulator (NAS) Program and by National Cancer Institute Supercomputer Center. We are grateful to the UCSF Computer Graphics laboratory (NIH RR-1081, Thomas Ferrin, P.I.) for graphics support. and the Computer Graphics Facility at the College of Chemistry, UC Berkeley (NIH S10 RR-5651A) for graphics support.

References

1. Monod, J., *Chance and Necessity*, Knopf: New York, 1971.
2. Morowitz, H. J., *Beginnings of Cellular Life*, Yale University Press: New Haven, 1992.

3. Pohorille, A.; Wilson, M. A., *Origins of Life and Evolution of Biosphere* **1995**, *25*, 21.
4. Schopf, J.W. in *Early Life on Earth*, Bengtson, S., Ed., Nobel Symposium No. 84. Columbia University Press, New York, 1994.
5. Chyba, C. F.; Thomas, P. J.; Brookshaw, L.; Sagan, C., *Science* **1990**, *249*, 366.
6. Lowe, D.R. in *Early Life on Earth*, Bengtson, S., Ed., Nobel Symposium No. 84. Columbia University Press, New York, 1994.
7. Chang, S. in *Early Life on Earth*, Bengtson, S., Ed., Nobel Symposium No. 84. Columbia University Press, New York, 1994.
8. Lasic, D.D., Elsevier: Amsterdam, 1993.
9. Deamer, D. W.; Pashley, R. M., *Origins of Life and Evolution of Biosphere* **1989**, *19*, 21.
10. Morowitz, H. J.; Heinz, B.; Deamer, D. W., *Origins of Life and Evolution of Biosphere* **1988**, *18*, 281.
11. Gilbert, W., *Nature* **1984**, *319*, 7618.
12. Wächtershäuser, G., *Microbiological Reviews* **1988**, 452.
13. Allen, M.P.; Tildesley, D.J., *Computer Simulation of Liquids*, Clarendon Press: Oxford, 1987.
14. Deamer, D.W.; Harang Mahon, E.; Bosco, G. in *Early Life on Earth*, Bengtson, S., Ed., Nobel Symposium No. 84. Columbia University Press, New York, 1994.
15. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L., *J. Chem. Phys.* **1983**, *79*, 926.
16. Wilson, M.; Pohorille, A., *J. Am. Chem. Soc.* **1994**, *116*, 1490.
17. Cornell, W.D.; Cieplak, P.; Bayly, C.I.; Gould, I.R.; Merz Jr., K.M.; Ferguson, D.M.; Spellmeyer, D.C.; Fox, T.; Caldwell, J.C.; Kollman, P.A., *J. Am. Chem. Soc.* **1995**, *117*, 5179.
18. Pohorille, A.; Wilson, M. A., *J. Chem. Phys.* (to be submitted).
19. Chipot, C.; Wilson, M.A.; Pohorille, A., *J. Phys. Chem.* **1995**. (to be submitted).
20. Diamond, J.M.; Katz, Y., *J. Membrane Biol.* **1974**, *17*, 121.
21. Marrink, S.J.; Berendsen, H.J.C., *J. Phys. Chem.* **1994**, *98*, 4155.
22. Pohorille, A.; Wilson, M.A., *Chem. Phys.* **1995**. (in press).
23. Pohorille, A.; Pratt, L.R., *J. Am. Chem. Soc.* **1990**, *112*, 5066.
24. Pratt, L.R.; Pohorille, A., *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 2999.
25. Walter, A.; Gutknecht, J., *J. Membr. Biol.* **1986**, *90*, 207.
26. Parsegian, V.A., *Nature* **1969**, *221*, 844.
27. Ohki, S., *J. Theor. Biol.* **1968**, *19*, 97.

28. Dilger, J.P.; Fisher, L.R.; Haydon, D.A., *Chem. Phys. Lipids* **1982**, *30*, 159.
29. Hauser, H.; Oldani, D.; Philips, M.C., *Biochemistry* **1973**, *12*, 12.
30. Nozaki, Y.; Tanford, C., *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 4324.
31. Deamer, D.W.; Bramhall, J., *Chem. Phys. Lipids* **1986**, *40*, 40.
32. Nagle, J.F.; Scott, H.L., *Biochim. Biophys. Acta* **1978**, *513*, 236.
33. Hamilton, R.T.; Kaler, E.W., *J. Phys. Chem.* **1990**, *94*, 2560.
34. Cruzeiro-Hansson, L.; Mauritsen, O.G., *Biochim. Biophys. Acta* **1988**, *944*, 63.
35. Wilson, M.A.; Pohorille, A., *J. Am. Chem. Soc.* **1995**. (to be submitted).
36. Deamer, D.W.; Volkov, A.G., in *Permeability and Stability of Lipid Bilayers*, Disalvo, E.; Simon, S., Eds., CRC Press: Boca Raton, 1995.
37. Skulachev, V.P. Chemiosmotic systems and the basic principles of cell energetics. in *Bioenergetics*, Ernster, L., Ed., vol. 23, *New Comprehensive Biochemistry*. Elsevier Science Publishers B.V., Amsterdam, 1984.
38. Deamer, D.W., *Advances in Space Research* **1992**, *12*, 183.
39. Oesterhelt, D.; Braüchle, C.; Hampp, N., *Quart. Rev. Biophys.* **1991**, *24*, 425.
40. Deamer, D.W.; Nichols, J.W., *J. Membrane Biol.* **1989**, *107*, 91.
41. Akeson, M.; Deamer, D.W., *Biochim. Biophys. Acta* **1990**, *60*, 101.
42. Abashkin, Y.; Burt, S.K. Private communication, 1995.
43. Lanyi, J., *Biochim. Biophys. Acta* **1993**, *183*, 241.
44. Creighton, T.E., in *Proteins, Structures and Molecular Properties*, Freeman and Company: New York, 1993.
45. DeGrado, W.F.; Lear, J.D., *J. Am. Chem. Soc.* **1985**, *107*, 7684.
46. Parente, R.A.; Nadasdi, L.; Subbarao, N.K.; Szoka, F.C., *Biochemistry* **1990**, *29*, 8713.
47. Kato, T.; Lee, S.; Ono, S.; Agawa, Y.; Aoyagi, H.; Ohno, M.; Nishino, N., *Biochim. Biophys. Acta* **1991**, *1063*, 191.
48. Ono, S.; Lee, S.; Mihara, H.; Aoyagi, H.; Kato, T.; Yamasaki, N., *Biochim. Biophys. Acta* **1990**, *1022*, 237.
49. Macquaire, F.; Baleux, F.; Giaccobi, E.; Huynh-Dinh, T.; Neumann, J.M.; Sanson, A., *Biochemistry* **1991**, *31*, 2576.
50. Ravishanker, G.; Mezei, M.; Beveridge, D.L., *J. Comput. Chem.* **1986**, *7*, 345.
51. Anderson, A.; Hermans, J., *Proteins* **1988**, *3*, 262.
52. Tobias, D.J.; Brooks, C.L., *J. Phys. Chem.* **1992**, *96*, 3864.
53. Pohorille, A.; Wilson, M.A. in *Reaction Dynamics in Clusters and Condensed Phases – The Jerusalem Symposia on Quantum Chemistry and*

Biochemistry, Jortner, J.; Levine, R. and ; Pullman, B., Eds., vol. 26. Kluwer, Dordrecht, 1993, p. 207.

54. Pohorille, A.; Wilson, M.A. Interaction of a model peptide with a water-bilayer system. in *Structure and Reactivity in Aqueous Solution: Characterization of Chemical and Biological Systems*, Cramer, C.; Truhlar, D., Eds., ACS Symposium Series No. 568. ACS, Washington D.C., 1994, ch. 11.
55. Brack, A.; Barbier, B., *Origins of Life and Evolution of Biosphere* **1990**, *20*, 139.
56. Johnsson, K.; Allemann, R.K.; Widmer, H.; Benner, S.A., *Nature* **1993**, *365*, 530.
57. Sasaki, T.; Kaiser, E.T., *J. Am. Chem. Soc.* **1989**, *111*, 380.
58. Tanford, C., *Science* **1980**, *200*, 1012.