

# Self-assembly processes in the prebiotic environment

David Deamer<sup>1,\*</sup>, Sara Singaram<sup>1</sup>, Sudha Rajamani<sup>1</sup>,  
Vladimir Kompanichenko<sup>1</sup> and Stephen Guggenheim<sup>2</sup>

<sup>1</sup>*Department of Chemistry and Biochemistry, University of California, Santa Cruz, CA 95064, USA*

<sup>2</sup>*Department of Earth and Environmental Sciences, University of Illinois at Chicago, Chicago, IL 60607-7059, USA*

An important question guiding research on the origin of life concerns the environmental conditions where molecular systems with the properties of life first appeared on the early Earth. An appropriate site would require liquid water, a source of organic compounds, a source of energy to drive polymerization reactions and a process by which the compounds were sufficiently concentrated to undergo physical and chemical interactions. One such site is a geothermal setting, in which organic compounds interact with mineral surfaces to promote self-assembly and polymerization reactions. Here, we report an initial study of two geothermal sites where mixtures of representative organic solutes (amino acids, nucleobases, a fatty acid and glycerol) and phosphate were mixed with high-temperature water in clay-lined pools. Most of the added organics and phosphate were removed from solution with half-times measured in minutes to a few hours. Analysis of the clay, primarily smectite and kaolin, showed that the organics were adsorbed to the mineral surfaces at the acidic pH of the pools, but could subsequently be released in basic solutions. These results help to constrain the range of possible environments for the origin of life. A site conducive to self-assembly of organic solutes would be an aqueous environment relatively low in ionic solutes, at an intermediate temperature range and neutral pH ranges, in which cyclic concentration of the solutes can occur by transient dry intervals.

**Keywords:** self assembly; clay minerals; prebiotic environment

## 1. INTRODUCTION

Several scenarios guide research on the beginning of life. The first scenario concerns the nature of environmental conditions conducive to the origin of life on the early Earth. All such scenarios are constrained by the requirement for liquid water and a source of organic compounds. A possible site must necessarily be far from equilibrium with a source of energy that must have been present to drive synthetic reactions. There must also have been a process by which organic compounds were sufficiently concentrated to undergo physical and chemical interactions.

The second scenario is related to chemical and physical properties of the organic compounds available to participate in the origin of life. One set of compounds would be soluble in water with chemical properties of monomers so that they can undergo polymerization into linear polymers. A second set of compounds would be amphiphilic, a property which permits self-assembly into membrane-bounded compartments that encapsulate subsets of the mixed components. The permeability properties of the

membranes were such that smaller water-soluble monomers could enter the compartment. The energy-dependent internal polymerization reactions then produced larger molecules that were unable to escape from the compartment. Thus, the compartment served as a microscopic reactor that isolated large molecules from their surroundings and enhanced the probability of interacting with one another.

Cellular life began when a yet unknown subset of self-assembled membrane-bounded polymers had the ability not only to polymerize, but also to replicate their linear sequence of monomers. The growth and replication processes were catalysed by a second set of polymers having a linear sequence of monomers that were somehow determined by the replicating subset. The membranous boundary of the system was also able to grow by the addition of amphiphilic components, and a process developed that allowed compartments to divide following a certain amount of growth. Evolution began when compartmented cellular systems began to compete for energy and nutrients to support their growth and division. Those systems that were more efficient were selected to undergo further growth and passed on their specific set of components to their progeny.

An immense gap in our knowledge concerns the processes occurring during the several hundred million years of chemical evolution that led to the first living cells. Cellular life did not begin with nucleic acids,

\* Author for correspondence (deamer@soe.ucsc.edu).

One contribution of 19 to a Discussion Meeting Issue 'Conditions for the emergence of life on the early Earth'.

The electronic supplementary material is available at <http://dx.doi.org/10.1098/rstb.2006.1905> or via <http://www.journals.royalsoc.ac.uk>.

proteins, ribosomes, a genetic code and a boundary structure. Instead, some sort of scaffold structure self-assembled on the prebiotic Earth from existing components, using organic compounds, sources of chemical energy and perhaps mineral interfaces. Here, we will describe some recent experiments aimed at understanding the environmental conditions that provided life's original scaffold.

## 2. PRIMITIVE SELF-ASSEMBLY PROCESSES

The physical properties of certain kinds of molecules lead to their aggregation into complex structures with emergent properties. For instance, if we add soap molecules to water, at first there is nothing except a clear solution of individual molecules dissolved in the water, but as the concentration increases, additional molecules no longer dissolve, and instead begin to associate into small aggregates called micelles. As the concentration increases further, the micelles grow into membranous layers that cause the originally clear solution to become milky. Finally, if air is blown through the solution, much larger structures appear at the surface, the soap bubbles that are familiar to everyone.

Such emergent phenomena are referred to as self-assembly processes, or sometimes self-organization. Self-assembly is the basic process by which contemporary cellular life produces membranes, duplex DNA and folded proteins, and the first cells must have incorporated many of the same intermolecular forces and self-assembled structures. The origin of life must also have involved reactions that result from the energy flowing through molecular systems, thereby driving them towards increasingly complex structures having emergent properties. For instance, membranous boundaries and the compartments they produce have the potential to make energy available in the form of ion gradients and can also mediate selective inward transport of nutrients. Furthermore, membranous compartments, in principle, are capable of containing unique systems of macromolecules. If a yet unknown macromolecular replicating system of polymers could be encapsulated within a membrane-bounded compartment, the components of the system would share the same microenvironment, and the result would be a major step towards cellularity, speciation and true cellular function (Cavalier-Smith 1987; Morowitz 1992; Dyson 1999; Deamer *et al.* 2002).

The most striking examples of self-assembling molecules are amphiphiles, because they have both a hydrophilic and a hydrophobic group on the same molecule. Amphiphilic molecules are among the simplest of life's molecular components and are readily synthesized by non-biological processes. Virtually any hydrocarbon having 10 or more carbons in its chain takes on amphiphilic properties if one end of the molecule incorporates a polar or an ionic group. Therefore, the simplest common amphiphiles are molecules such as soaps, more technically referred to as monocarboxylic acids.

We and others have found that such molecules readily form membranous vesicles, as shown in figure 1 (Hargreaves & Deamer 1978; Walde *et al.* 1994; Apel

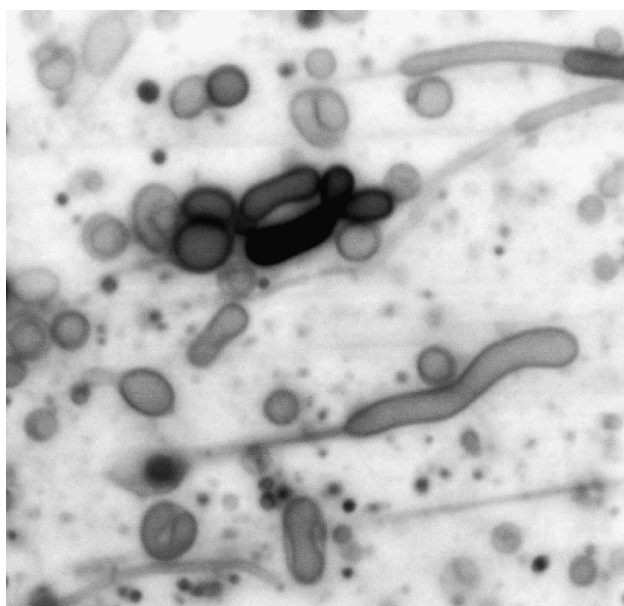


Figure 1. Membranous vesicles are produced when a mixture of decanoic acid–decanol is dispersed in aqueous salt solutions at neutral pH ranges. For the purposes of illustration, the original colour micrograph of fluorescent vesicles stained with rhodamine 6G was inverted to produce a greyscale image. Scale bar, 10  $\mu\text{m}$ . See Monnard *et al.* (2002).

*et al.* 2002; Hanczyc *et al.* 2003; Hanczyc & Szostak 2004). Similar molecules occur in the mixture of organic compounds present in carbonaceous meteorites and have been shown to self-assemble into membranous structures (Deamer & Pashley 1989). It is significant that amphiphilic molecules have been synthesized from simple gases under simulated geothermal conditions that commonly occur in volcanic environments (McCollom *et al.* 1999; Rushdi & Simoneit 2001). This will become important later when we discuss environmental factors related to the emergence of cellular life.

### (a) *Synthesis of linear polymers*

Two species of polymers dominate all life today: proteins and nucleic acids. The remarkable relationship between these two polymers virtually defines contemporary life. That is, protein enzymes catalyse the synthesis of nucleic acids and the sequence of bases in nucleic acids carries the information required to synthesize proteins. A primary objective of origin-of-life research is to understand the evolutionary origin of this relationship. However, to date, there is no clear understanding of plausible mechanisms by which linear polymers could be synthesized in the prebiotic environment.

The origin of peptides might have been a relatively simple process. In contrast to nucleic acids, only a single monomer species—amino acids—is involved. If we use the amino acid content of carbonaceous meteorites as a guide, at least 80 species of amino acids would be available in the prebiotic environment. The question, then, is how peptide bond formation could have occurred. One possibility is that the synthesis of covalent bonds between amino acids was driven by conditions in which water becomes a leaving group. Heating and drying cycles, such as those occurring at the

mineral–water interfaces in geothermal sites exposed to the atmosphere, are able to drive condensation reactions that produce polymers of simple organic molecules. The chemical potential is provided by the anhydrous condition, so that condensation reactions become favourable as water leaves the reactants. Early work by Fox & Harada (1958) first demonstrated that simple heating of amino acid mixtures to temperature of approximately 160°C produced polymeric substances they termed as ‘proteinoids’. Rohlfsing (1976) showed that, given sufficient time, similar polymers can form at lower temperatures as well.

Although there was considerable initial excitement about the ease of polymerizing amino acids, it became apparent that peptide bonds were a minor component and most of the bonds produced by anhydrous heating involved other kinds of covalent linkages. Thus, pathways that permit peptide bond synthesis in aqueous solution without the requirement for thermal cycles are desirable. An elegant synthetic mechanism was recently reported by Leman *et al.* (2004), in which carbonyl sulphide (COS) promotes the formation of peptide bonds between a variety of amino acids in solution, producing di-, tri- and tetrapeptides. Because COS is a component of volcanic gases, this reaction may have produced simple peptides in geothermal environments.

#### (b) Possible roles of mineral surfaces as organizing agents

Mineral surfaces have several important properties related to the origin of life (Orgel 1997; Hazen *et al.* 2001). First, mineral surfaces offer an interface that may adsorb thin films of organic solutes from aqueous solutions and thereby concentrate them. Minerals composed of iron and sulphur have the potential to act as oxidants and reducing agents, providing chemical free energy for reactions that could evolve into primitive metabolic pathways. Assuming that volcanic land masses emerged from the early ocean, mineral surfaces could have served as sites for synthetic reactions that produce polymers. Ferris and co-workers (see Ferris 2002) have explored the ability of montmorillonite, a type of smectite clay, to organize and perhaps even catalyse the synthesis of RNA from chemically activated mononucleotides, as a specific example of a mineral-dependent reaction pathway. RNA polymers up to 50 nucleotides in length can be produced by this reaction.

Huber & Wächtershäuser (1997) described another experimental model of mineral catalysis, in which a slurry of nickel and iron sulphides were found to promote the formation of acetic acid from carbon monoxide and methyl mercaptan (CH<sub>3</sub>SH). Peptide bond formation has also been demonstrated (Huber & Wächtershäuser 1998). These conditions were considered to simulate a primordial geothermal system in which metal sulphides at high temperatures (approx. 100°C) provide a reaction pathway for the initial steps of an autotrophic metabolism.

Martin & Russell (2003) have taken this concept a step further. They note that certain iron sulphide minerals contain microscopic pores in the size range of cells (approx. 10–100 µm) and proposed that such cavities could provide a mineral version of a

membranous boundary structure. The authors suggest that the cavities may be able to concentrate nutrient organic solutes that could serve as reactants in primitive metabolic pathways. They also propose that the iron sulphide membranes could provide a source of chemical energy, perhaps even chemiosmotic energy to drive early metabolism.

### 3. THE PHYSICAL ENVIRONMENT AND THE ORIGIN OF LIFE

How are self-assembly processes affected by physical parameters of the prebiotic environment at the time that cellular life emerged approximately 4 billion years ago? Can geological sites exist on the contemporary Earth that provide model systems of the early Earth environment? To address these questions, one possible site was chosen for initial investigations. Here, we report results from preliminary experiments to investigate the behaviour of organic compounds in a natural simulation of the early Earth environment.

#### (a) Darwin's pond revisited

In 1871, Charles Darwin wrote a letter to his friend Joseph Hooker in which the following idea was expressed:

It is often said that all the conditions for the first production of a living organism are now present, which could ever have been present. But if (and oh! What a big if) we could conceive in some warm little pond, with all sorts of ammonia and phosphoric salts, light, heat, electricity, &c., present, that a protein compound was chemically formed ready to undergo still more complex changes, at the present day such matter would be instantly devoured or absorbed, which would not have been the case before living creatures were formed.

Darwin could not have known that he was suggesting a scenario for a prebiotic environment which would still be under consideration nearly 140 years later. Other sites for the origin of life have since been proposed, ranging from transient melt zones in a global ice pack (Bada *et al.* 1994) to hydrothermal vents (Corliss *et al.* 1981; Baross & Hoffman 1983) and to subterranean lithic zones (Pace 1991). Many such sites contain microbial populations that are extremophiles, and it has been suggested that the last universal common ancestor lived in a deep hydrothermal site secure from sterilizing giant impacts (Sleep *et al.* 1989). However, a surface site at an intermediate temperature range, exposed to the atmosphere, with the ability to produce concentrated reactants by evaporation remains a viable alternative for the site of the origin of life. After life began, it could then evolve into a variety of niches, including the extremophile environments.

Darwin might have imagined his pond to resemble the tropical salt-water tide pools he had observed in the Galapagos, or a fresh-water pond similar to those that dot the English landscape. However, he did specify a form of fixed nitrogen (ammonium salt) and phosphate, presumably because he knew that nitrogen and phosphorus are among the six major elements of life (C, H, O, N, P and S). He did not include a carbon source explicitly or mention any organic compounds



other than proteins, because, in 1871, it was not yet understood that amino acids and nucleobases were the major monomers of life's polymeric macromolecules.

The plausibility of Darwin's pond has not been tested in the field. At a minimum, such a site is expected to have access to liquid water and organic compounds, either synthesized *in situ* or transported to the site. Possible sources of organics include geochemical synthesis or delivery by extraterrestrial sources during late accretion. From the earlier results described previously, it is possible that self-assembly of amphiphiles might occur under these conditions. It is even conceivable that synthesis of peptide and ester bonds would be possible. However, the monomers must remain in solution or adsorbed to mineral surfaces in a configuration that permits reactions to occur. If conditions are such that the monomers are tightly adsorbed to mineral surfaces or produce insoluble salts, polymerization chemistry would not be possible.

Volcanic aquifers are potential models of a prebiotic environment that may be conducive to synthetic organic reactions (Washington 2000; Furuuchi *et al.* 2005). To test this idea, we have performed preliminary experiments in geothermal sites in Kamchatka, Russia and Mt Lassen, California, in order to determine how a set of pertinent organic compounds behave in such settings. The Kamchatka Peninsula is situated in the far northeastern part of Russia and spans a region of approximately 400 by 1200 km. This region contains nearly 10% of all active volcanoes of the world, together with 150 exposed hydrothermal sites. We chose a 3 hectare geothermal exposure on the flank of Mt Mutnowski as a test site. In California, Mt Lassen has a similar geothermal area called Bumpass Hell, with boiling clay-lined ponds similar to those of Kamchatka. Both sites represent modern analogues for hydrothermal environments typical of volcanic sites active throughout Earth's history.

Instead of Darwin's 'warm little pond', small boiling pools were studied at both the sites. Our experimental system used four amino acids and nucleobases, phosphate and a fatty acid amphiphile. Glycerol was also added to test whether ester bond formation could occur, for instance, between the -OH groups of the glycerol and the carboxylate of the fatty acid. Three primary questions were addressed, which are as follows.

- (a) What is the fate of a mixture of organic compounds in a natural setting? Assuming that there was a source of organic compounds on the prebiotic Earth, at what rate would such compounds disappear from solution, either by adsorption to mineral surfaces or chemical degradation?
- (b) Can self-assembly of membranous vesicles occur in hot springs typical of volcanic regions?
- (c) Are any synthetic reactions or self-assembly processes possible under such conditions?

(i) *The geothermal environment*

The Kamchatka site is an exposed geothermal area on the flank of Mt Mutnowski, an active volcano approximately 50 km south of Petropavlovsk. There are numerous boiling springs with water quality

ranging from clear to moderately turbid. Most of the springs have significant flows of water from a neighbouring snowpack above the site and are therefore unsuitable because added organic compounds would be rapidly flushed away. One site had only a very slow influx seepage, which was diverted by digging a shallow bypass channel (figure 2a).

The pool has a heat source in the centre which maintained a constant boiling temperature of 97°C. The pool was lined with a layer of grey clay mineral several centimetres deep, which was disturbed by boiling to produce constant agitation and turbidity. The equivalent test site in northern California was a geothermal exposure called Bumpass Hell on the flank of Mt Lassen, dominated by clay-lined hot springs. Two small clay-lined boiling-water basins were chosen that resembled the Kamchatka site, but neither of them had a boiling action as active as the Kamchatka pool. One of the Lassen sites was agitated by occasional bubbles rising through approximately 5 cm water every few seconds (figure 2b), whereas the second, although obviously near boiling temperatures (93°C at the elevation of the site), was relatively undisturbed.

The obvious source of water for the Kamchatka pool was seepage from snowmelt 10–50 m above the geothermal area. The seepage was exposed to plant growth in the form of grass and small bushes, and some input of organic materials from biological sources is expected. However, the gram quantities of material added to the pool were far greater than any influx from the local flora; therefore, it could be distinguished from the traces of biological material.

The inorganic solutes of the Kamchatka pool are shown in table 1. The data are presented as milligram per litre from the original analysis, then converted to approximate millimolar concentrations. The total cationic charge present amounts to 4.2 mM, which includes 1.0 mM hydrogen ion. (The measured pH of the pool was 3.1.) The total anionic charge present as sulphate is 4.4 mM, in reasonable balance with the total cationic charge. We assume that the silicate in solution at pH 3.1 is present in its neutral protonated form and does not contribute to the anionic charge. The concentrations of cations are significant and will be discussed in the following sections.

(ii) *Solute mixture*

At both the Kamchatka and the Mt Lassen sites, we added a defined set of primary molecular constituents of life to the hot springs. This included four amino acids (glycine, L-alanine, L-aspartic acid and L-valine, 1 g each), four nucleobases (adenine, cytosine, guanine and uracil, 1 g each) sodium phosphate (3 g), glycerol (2 g) and myristic acid (1.5 g). We expect that this mixture is a useful guide for monitoring degradation reactions, in addition to serving as a potential source of reactants for any synthetic reactions that might occur. The powdered mixture was first dissolved in approximately 1.0 l hot water from the site and the resulting milky solution was poured into the boiling centre of the pond to start the experiment. Upon adding the mixture, a white precipitate appeared on the surface of the pool and

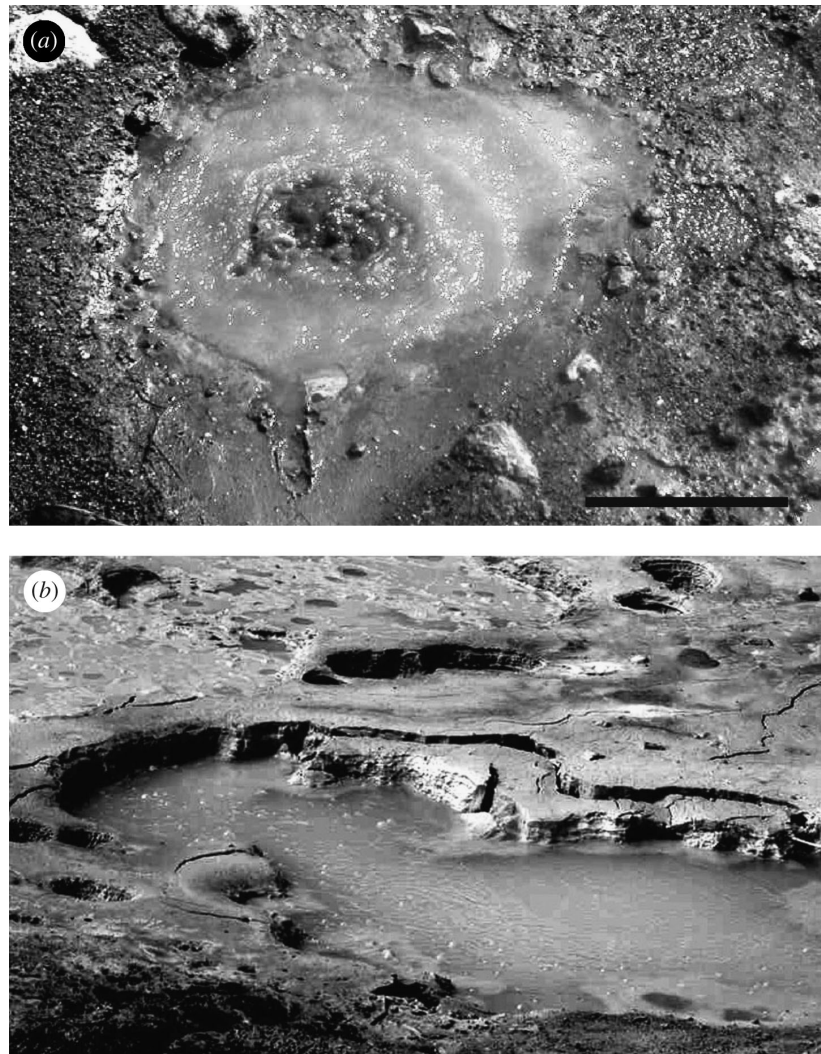


Figure 2. Boiling pools investigated in (a) Kamchatka and (b) Mt Lassen sites. Note the depression at the lower edge of the Kamchatka pool that indicates one of the clay sampling sites. Scale bar, 25 cm.

Table 1. Ionic composition and concentrations in the Kamchatka boiling pool.

cations	$\text{NH}_4^+$	$\text{Na}^+$	$\text{Ca}^{2+}$	$\text{Fe}^{3+}$	$\text{K}^+$	$\text{Mg}^{2+}$	$\text{Al}^{3+}$	$\text{H}^+$
$\text{mg l}^{-1}$	15	14	12.4	6	4.8	3.1	3.1	—
mM	0.9	0.6	0.31	0.11	0.12	0.13	0.11	1.0
anions	$\text{SO}_4^{2-}$	$\text{Cl}^-$	$\text{SiO}_4$					
$\text{mg l}^{-1}$	211	1.4	288					
mM	2.2	0.04	3.1					

accumulated around the border, then slowly dispersed over a period of an hour. Water samples (50 ml) were taken before the addition of the powder, then at 1, 5, 30, 60, 120 min, 24 h and 9 days later. We also sampled approximately 50 g of clay at the edge of the pool before and 2 h after the addition of the organic solutes. The samples were taken by scraping off a layer of clay approximately 1 cm thick (figure 2). All the samples were returned to our laboratory where they were stored at 4°C during the three months required for analysis. At the end of analysis, the samples were examined by phase-contrast and fluorescence microscopies for possible bacterial growth, but none was found.

### (iii) Fate of added solutes

Details of analytical procedures are provided in the electronic supplementary material. Briefly, the dissolved solutes in the water samples were analysed by high performance liquid chromatography (HPLC) either directly or after reacting with agents that produced a spectral absorption at UV wavelengths (amino acids and a fatty acid) or visible wavelengths (phosphate). The strong UV absorbance of the nucleobases permitted their direct measurement at 260 nm as aliquots of the water samples were passed through a reverse phase HPLC column.

In the Kamchatka site, as at one site at Mt Lassen, most of the added solutes rapidly disappeared from

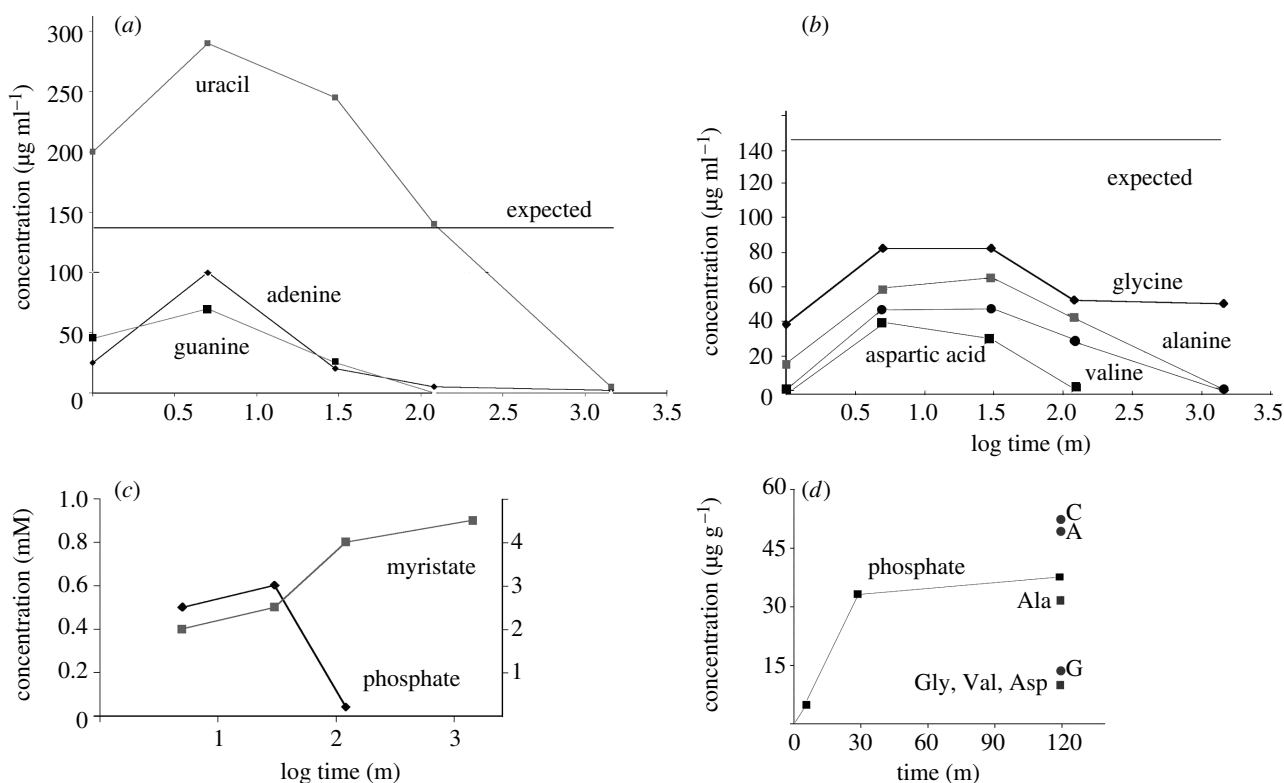


Figure 3. Time course of organic solute and phosphate concentration in the Kamchatka site. (a) Purine bases were lost from solution with half-times of 30 min, while uracil, a pyrimidine, remained in solution for longer time periods. Although four nucleobases were added, only three, adenine, guanine and uracil, could be detected in solution. Cytosine may have undergone deamination to produce uracil, which would account for the apparent doubling of the uracil concentration. The line at  $100 \mu\text{g ml}^{-1}$  shows the expected concentration of nucleobases, assuming that the pool contained 10 l volume that was estimated from the known amount of phosphate added. (b) Amino acids disappeared from solution with half-times comparable to those of the nucleobases, in the range of 30 min. Aspartic acid was lost most rapidly, while glycine could still be detected after 24 h. The line at  $100 \mu\text{g ml}^{-1}$  shows the expected concentration of amino acids, assuming that the pool contained 10 l volume. (c) Phosphate rapidly went into solution, but then became undetectable after 2 h. In contrast, myristic acid remained throughout a 24 h period and could still be detected 9 days after it was added. (d) Treatment with 0.1 M NaOH, but not 0.1 M HCl, released phosphate, nucleobases and amino acids from clay samples taken from the edge of the pool (see figure 1).

solution, some within minutes and others over a period of several hours or days. From the known amount of adenine added and its final concentration, we determined that the pool contained approximately 10 l of water, indicating that the gram quantities of organic solutes and phosphate should be easily detectable. Figure 3a shows the behaviour of the four nucleobases at the Kamchatka site. The solid line indicates the expected concentration for the addition of 1.0 g nucleobase to 10 l water. Note that the two purines—guanine and adenine—approached the expected line but were not detectable after 2 h. Cytosine could not be detected as a solute, but uracil was present at twice the expected concentration and became undetectable over a somewhat longer time-scale. The most probable explanation is that much of the cytosine was rapidly deaminated to produce uracil, although a small amount of cytosine could be detected bound to the clay lining the hot spring, as described in figure 3.

The phosphate and added amino acids were below detectable limits in minutes to hours, where aspartic acid was lost most rapidly and glycine more slowly (figure 3b,c). Of the solutes, only myristic acid could be detected after 24 h. (The added glycerol was not analysed.)

This experiment was repeated in the two pools at the Mt Lassen site. In the first pool, which was agitated by occasional vapour bubbles rising through the water, the results with nucleobases were similar to those at the Kamchatka site. None was detectable 2 h after addition. However, in the second undisturbed pool, the added nucleobases were present after 2 h. We assume that agitation is required to expose the added organics to suspended clay, which is then deposited in the bottom layer of the pool. In the absence of agitation the organics will remain in solution for longer time periods.

The observation that organic compounds were below detection limits so rapidly was surprising. We first considered whether some of the compounds might be precipitating after reacting with ionic solutes in the pool. The ionic composition of the pool is shown in table 1 and cation concentrations become significant when possible interactions with added solutes are considered, particularly phosphate and myristic acid. The concentration of phosphate (added as  $3.0 \text{ g NaH}_2(\text{PO}_4) \cdot \text{H}_2\text{O}$ ) would be equivalent to approximately 2 mM. At pH 3.1, phosphate would not be expected to form insoluble precipitates with calcium or magnesium ions, but could form insoluble iron or aluminium salts. However, even if this occurs, there is only enough iron and aluminium in



solution to bind a small fraction of the added phosphate as an insoluble salt. In contrast, 1.5 g myristic acid added to 10 l volume would be 0.6 mM, and iron and aluminium in solution would be sufficient to precipitate at least some of the added myristate as iron and aluminium soaps. This explains our observation that immediately after adding the organic mixture, a white precipitate appeared on the surface of the pool and accumulated around the border. Myristic acid is a liquid at the temperature of the pool, so the precipitate could only be a fatty acid soap with a much higher melting point.

The possibility that the organic compounds might be degraded by the conditions of the pool was considered, but control experiments in the laboratory showed that none of the solutes were lost after 2 h boiling in dilute sulphuric acid, pH 3.1 (the pH and acid content of the Kamchatka water). This suggested that, with the exception of the fatty acid, the compounds were adsorbed to the suspended clay particles and deposited in the several centimetres thick clay layer lining the pool. For this reason, the clay was analysed for adsorbed organics and phosphate.

Samples of clay (2 g wet weight) were extracted in both acidic and basic conditions by the addition of 20 ml of 0.1 M HCl or NaOH. In clay samples that were obtained before addition of the solute mixture, no soluble organics or phosphate were extracted either by acid or base treatment. However, in the NaOH extracts (but not HCl), nucleobases, amino acids and phosphate were all released from the clay. The amount of phosphate released was related to the time following the addition of the solute mixture (figure 3d). Time points were not analysed for the amino acids and nucleobases; therefore, only the values of solutes adsorbed after 2 h are presented in the figure.

We also tested the ability of the Kamchatka clay sample to bind adenine under laboratory conditions. Total organic carbon bound to the clay was kindly measured by Robert Hazen at the Carnegie Institution of Washington, and found to be 0.2% of the dry mass. None of this adsorbed material was detected by any of our tests, but strongly absorbed light in the UV wavelengths, so presumably it was a mixture of degradation products derived from biological sources, such as the grass that grew within several metres of the sampling site. To reduce interference in the UV range, the adsorbed organic compounds were removed from the clay by three alkaline washes (10 ml of 0.1 M NaOH per gram of clay), with the clay separated from the supernatant by low-speed centrifugation (4000g, 5 min). The clay was washed in 0.1 M HCl, followed by three water rinses. At the end of this procedure, no UV-absorbing solutes could be detected in the supernatant above the final clay pellet.

The pellet was then suspended in 20 ml aqueous adenine solutions of increasing concentration, brought to boiling for 1 min, cooled and centrifuged as before. We found that 4 mg adenine per gram clay was completely adsorbed. However, only 30% of the adenine at 10 mg g<sup>-1</sup> clay was adsorbed and 20 mg g<sup>-1</sup> saturated the clay adsorptive surface. This result confirmed that clay was well able to adsorb adenine at the concentration that was added to the pool.

#### (iv) Clay-mineral components

It is significant that the clay mineral apparently bound most of the added solutes. The area of the clay layer lining the pool was estimated from the pool diameter to be approximately 2500 cm<sup>2</sup>. The clay layer was approximately 2 cm thick, so that the total clay was about 5 kg, presumably more than sufficient to adsorb a few grams of added organic compounds and phosphate, even though the deeper clay-mineral surfaces were not available for adsorption. Since the type of clay affects adsorption, we were interested in the mineral components of the clay. Samples were analysed according to the procedures outlined in the electronic supplementary material.

Three samples were analysed: one of the original clay which was taken directly from the edge of the acidic pool; and the second and third samples after being exposed to acid and base solutions to desorb bound organics. The original sample was tan grey in colour when wet, the acid-treated was tan brown and the base-treated was red brown.

The untreated clay sample produced a diffraction pattern with a broad (001) peak at  $d=14.63 \text{ \AA}$ , which was expanded by ethylene-glycol solvation to approximately 17.5 Å with a significant low-angle shoulder. The acid-treated sample showed a peak after solvation at 17.31 Å with no significant low-angle shoulder. The base-treated sample produced a solvation peak at 17.48 Å and a significant low-angle shoulder. A typical smectite clay will produce a (001)  $d$ -value of about 16.9 Å, so the tentative identification based on ethylene glycol treatment is smectite, but additional data are required (see below) to explain the greater-than-normal expansion. The higher-order smectite peaks after ethylene-glycol treatment are either low in intensity or non-existent, but the shape of the 001 peak is broad with a significant low-angle shoulder, the latter suggesting a small amount of random interstratification, probably with illite. After heating to 300°C, a small peak at 10 Å was produced, indicating collapse of the smectite. The base-treated sample showed almost no peak at 10 Å, whereas the acid-treated sample showed a more prominent peak at 10 Å after heating to 300°C. All samples showed a phase with  $d(001)$ , which proved to be kaolin, because heating at 550°C showed that no chlorite was present (i.e. no 14 Å peak was present after treatment at 550°C) and the peak at 7.285 Å disappeared, indicating that the kaolin became amorphous as expected at this high temperature. Elevated background at low angles of about 9.91 Å occurs even in patterns subjected to 550°C, possibly suggesting that some interlayer water remained in the smectite structure even at this very high temperature.

An Mg-exchanged sample followed by glycerol treatment was used to obtain more information about the nature of the smectite. The  $d(001)$  of the Mg-exchanged smectite in the glycerol-treated pattern is 18.1 Å. Christidis & Eberl (2003) showed that a peak at about 17.8 Å after Mg saturation and glycerol treatment is related to a high-charge smectite or vermiculite. Further analysis using K saturation and glycerol treatment indicates that the clay contains smectite at high charge and consists of a very fine

particle size. The unexchanged and exchanged smectite too produced a distinct diffraction band near 20–22.5°C, suggesting that the smectite shows severe (turbostratic) stacking disorder. This region is also affected by general broadening of the pattern caused by the quartz glass mounting plates with broadening features from 15 to 30°C and a maximum at 21°C.

Very sharp peaks at  $d=9.00$  and  $4.04$  Å are too sharp to be related to clay. These peaks are very weak and may be related to the presence of a zeolite, such as stilbite (e.g. 100 and 50% peaks at 9.13 and 4.04 Å, respectively). No other unaccounted peaks were observed.

#### 4. DISCUSSION AND CONCLUSIONS

This initial study of the behaviour of organic compounds in a natural geothermal environment has been instructive in guiding future experiments on self-assembly, polymer synthesis and origin of cellular life. Four primary physical and chemical parameters markedly affect the fate of organic compounds and phosphate under the following conditions: organic solute concentration; temperature; pH and cation composition and mineral adsorption. In the following discussion, we consider each of these parameters.

##### (a) *Organic solute concentration*

The solutes added are all plausible components of a possible prebiotic mixture. They are present at concentrations in the millimolar range, probably much higher than the concentration of organic compounds in the early ocean when added in gram quantities to the Kamchatka boiling pool. Only in the fluctuating environment at the edge of the pool would concentrations be sufficiently high to permit self-assembly or synthetic polymerization reactions. No evidence of self-assembled structures or covalent bond formation in the clay samples taken from the pool edge was noted, but other factors intervened, as discussed in the following sections.

##### (b) *Effect of temperature*

Although self-assembly of amphiphilic molecules promotes the formation of complex molecular systems, the physical and chemical properties of an aqueous phase can significantly inhibit such processes, possibly constraining the environments in which cellular life first appeared. One such constraint is that temperature strongly influences the stability of vesicle membranes. It has been proposed that the last common ancestor, and even the first forms of life, were hyperthermophiles which developed in geothermal regions, such as hydrothermal vents (Baross & Hoffman 1983) or deep subterranean hot aquifers (Pace 1991). Such environments have the advantage of providing chemical energy in the form of redox potentials and abundant mineral surfaces to act as potential catalysts and adsorbents. However, because the intermolecular forces that stabilize the self-assembled molecular systems are relatively weak, it is difficult to imagine how lipid bilayer membranes assembling from plausible prebiotic constituents would be stable under these conditions. All hyperthermophiles today have

highly specialized lipid components, and it seems probable that these are the result of more recent adaptation rather than a molecular fossil of the earliest forms of cellular life.

##### (c) *Effects of pH and cations on self-assembly*

Fatty acids readily form membranous vesicles at neutral pH ranges near the pK of the carboxylate head group. However, at low pH ranges where the head group is protonated, only a bulk-phase structure occurs in the form of liquid or solid, depending on the melting point of the fatty acid. For this reason, we did not expect to find any self-assembled membranous structures produced by the fatty acid itself. However, it is known that glycerol esters of fatty acids form stable lipid bilayers over a variety of pH ranges, because the glycerol head group does not interact with hydrogen ions. Although no significant yield of monoglycerides was found, they might have been synthesized but not detected in this preliminary study.

A second effect on self-assembly is cation composition of the geothermal environment. Di- and trivalent cations strongly interact with anionic lipid head groups. A white scum appeared in the Kamchatka pool within minutes of adding the organic mixture. The precipitate is probably a mixed iron and aluminium soap, which would remove the fatty acid as a potential reactant. A similar fate might govern amphiphile compositions in a marine environment. The concentrations of divalent cations, in particular  $Mg^{2+}$  and  $Ca^{2+}$ , were likely to exceed 10 mM in the early oceans. In the absence of oxygen,  $Fe^{2+}$  would also be present at millimolar concentrations. All such divalent cations have a strong tendency to bind to the anionic head groups of amphiphilic molecules, strongly inhibiting their ability to form stable membranes. We emphasize again that glycerol esters of fatty acids are virtually unaffected by pH and divalent cations. Significant yields of monoglyceride are known to occur under laboratory conditions that simulate the temperature regime of geothermal pools (Apel & Deamer 2005). This suggests that the first step in chemical evolution towards self-assembling amphiphiles might have been the esterification of naturally occurring monocarboxylic acids. The resulting monoglycerides could form relatively stable membranes that would be resistant to pH and divalent cations, perhaps resolving the concern that acidic and high salt environments would not be conducive to self-assembly processes. It is interesting to note that most contemporary membrane-forming lipids use fatty acid ester or ether bonds linked to a glycerol backbone.

##### (d) *Adsorption to mineral surfaces*

Adsorption of organic compounds to clay minerals has been the subject of a number of earlier publications (Kell 1994; Hedges & Kell 1995; Mayer & Xing 2001). There are both favourable and unfavourable consequences of adsorption to mineral surfaces. Since virtually all the added solutes, including phosphate, adsorbed to the clay surfaces, such solutes can be concentrated from a dilute solution onto a two-dimensional surface. Thus, polymerization reactions would be promoted if they are thermodynamically



favoured. However, the diffusional mobility of the solutes is reduced by adsorption, which would inhibit the ability of potential reactants to undergo collisions required for a synthetic reaction.

Perhaps more important is that even if the reactants are concentrated by adsorption to a mineral surface, they are likely to be isolated from one another. A physical property of clays in this regard is the immense surface area of the clay mineral. Clay particle dimensions are by definition less than 4  $\mu\text{m}$ , and smectite clay has a surface area ranging from 80 to 120  $\text{m}^2 \text{g}^{-1}$  (Mitchell 1993). From our measurements of adsorbed organics (figure 3d), a total of approximately 170  $\mu\text{g}$  of amino acids and nucleobases were adsorbed per gram of clay, assuming complete stirring of the 2 cm layer that was sampled. However, it is probable that the solutes adsorbed only to the surface layer; therefore, for the purposes of calculation, we will assume that all the solutes adsorbed to a 1 mm thick clay layer. The molecular weights of the solutes range from 75 (glycine) to 151 (guanine). If we assume an average molecular weight of 100, and 100  $\text{m}^2$  surface area per gram clay, one molecule of reactant would be bound per 5  $\text{nm}^2$ , which is too far apart for reactions to take place between molecules. By way of contrast, when clay is used to promote RNA synthesis from activated nucleotides (Ferris 2002), the amount of added nucleotides is equivalent to 2 nucleotides per  $\text{nm}^2$ , clearly more conducive to the intermolecular contacts required for phosphodiester bond formation between nucleotides.

To summarize, the organic compounds required for the origin of life in a natural setting would have had a variety of possible fates other than those observed in a laboratory setting, where pure compounds react in glass containers. Our results show that self-assembly of boundary structures cannot occur under hot acidic conditions, at least with a simple amphiphile such as myristic acid. Furthermore, if di- and trivalent cations are present, as in the Kamchatka pool described here, the fatty acids will precipitate as insoluble soaps, which inhibits potential membrane formation. This limitation occurs not only for the boiling hot spring described here, but also for seawater, which has a relatively high content of calcium and magnesium ions (Monnard *et al.* 2002). We are presently exploring other possible amphiphiles that may be able to form stable membranes over a range of pH values and in the presence of divalent cations. As mentioned earlier, monoglycerides are possible candidates.

Interactions of organic solutes with mineral surfaces would tend to concentrate organic solutes, and it is possible that such processes would enhance condensation reactions leading to polymers. However, adsorption of organic solutes and phosphate to clay-mineral surfaces also has the potential to isolate reactant molecules and thereby inhibit polymerization.

From a biophysical perspective, these considerations suggest that the most plausible planetary environment for the origin of life would be an aqueous phase at moderate temperature ranges (less than 60°C) and low ionic strength, having a pH value near neutrality with divalent cations at submillimolar concentrations. This suggestion is in marked contrast to the view that life

most probably began in a geothermal or marine environment, perhaps even the extreme environment of a hydrothermal vent. One argument favouring a marine site for life's beginning is that fresh water would be rare on the early Earth. Even with today's extensive continental crust, fresh water represents only approximately 1% of the contemporary Earth's reservoir of liquid water. Another concern about a fresh water origin of life is that the lifetime of fresh water bodies tends to be short on a geological time-scale. On the other hand, if the temperature and ionic composition of a geothermal spring or seawater markedly inhibits self-assembly processes, we may need to reconsider the assumption that life inevitably began in such environments. A more plausible site for the origin of cellular life may be fresh water pools maintained by the rain falling on the earliest volcanic land masses and small continents. After the first form of cellular life was able to establish itself in a relatively benign environment, it would rapidly begin to adapt through Darwinian selection to more rigorous environments, including extreme temperatures, salt concentrations and pH ranges that we now associate with the limits of life on the Earth.

Partial support was obtained from the Astrobiology Program at the Carnegie Institution of Washington, and the National Science Foundation under grant no. EAR-0308588.

## REFERENCES

- Apel, C. & Deamer, D. W. 2005 The formation of glycerol monodecanoate by a dehydration condensation reaction. *Orig. Life Evol. Biosph.* **35**, 323–332.
- Apel, C. L., Deamer, D. W. & Mautner, M. 2002 Self-assembled vesicles of monocarboxylic acids and alcohols: conditions for stability and for the encapsulation of biopolymers. *Biochim. Biophys. Acta* **1559**, 1–10. (doi:10.1016/S0005-2736(01)00400-X)
- Bada, J. L., Bigham, C. & Miller, S. L. 1994 Impact melting of frozen oceans on the early Earth—implications for the origin of life. *Proc. Natl Acad. Sci. USA* **91**, 1248–1250. (doi:10.1073/pnas.91.4.1248)
- Baross, J. A. & Hoffman, S. E. 1983 Submarine hydrothermal vents and associated gradient environments as sites for the origin and evolution of life. *Orig. Life* **15**, 327–345.
- Cavalier-Smith, T. 1987 The origin of cells: a symbiosis between genes, catalysts and membranes. *Cold. Spring Harb. Symp. Quant. Biol.* **LII**, 805–824.
- Christidis, G. E. & Eberl, D. D. 2003 Determination of layer-charge characteristics of smectites. *Clay. Clay Miner.* **51**, 644–655. (doi:10.1346/CCMN.2003.0510607)
- Corliss, J. B., Baross, J. A. & Hoffman, S. E. 1981 An hypothesis concerning the relationship between submarine hot springs and the origin of life on Earth. *Oceanol. Acta. Proc. 26th Int. Geol. Congress, Geology of the Oceans Symp., Paris*, pp. 59–69.
- Deamer, D. W. & Pashley, R. M. 1989 Amphiphilic components of carbonaceous meteorites. *Orig. Life Evol. Biosph.* **19**, 21–33. (doi:10.1007/BF01808285)
- Deamer, D. W., Dworkin, J. P., Sandford, S. A., Bernstein, M. P. & Allamandola, L. J. 2002 The first cell membranes. *Astrobiology* **2**, 371–382. (doi:10.1089/153110702762470482)
- Dyson, F. 1999 *The origins of life*. Princeton, NJ: Princeton University Press.

- Ferris, J. 2002 Montmorillonite catalysis of 30–50 mer oligonucleotides: laboratory demonstration of potential steps in the origin of the RNA world. *Orig. Life Evol. Biosph.* **32**, 311–332. (doi:10.1023/A:1020543312109)
- Fox, S. W. & Harada, K. 1958 Thermal copolymerization of amino acids to a product resembling protein. *Science* **128**, 1214.
- Furuuchi, R., Imai, E.-I., Honda, H., Hatori, K. & Matsuno, K. 2005 Evolving lipid vesicles in prebiotic hydrothermal environments. *Orig. Life Evol. Biosph.* **35**, 333–343. (doi:10.1007/s11084-005-2039-7)
- Hanczyc, M. M. & Szostak, J. W. 2004 Replicating vesicles as models of primitive cell growth and division. *Curr. Opin. Chem. Biol.* **28**, 660–664. (doi:10.1016/j.cbpa.2004.10.002)
- Hanczyc, M. M., Fujikawa, S. M. & Szostak, J. W. 2003 Experimental models of primitive cellular compartments: encapsulation, growth, and division. *Science* **302**, 618–622. (doi:10.1126/science.1089904)
- Hargreaves, W. R. & Deamer, D. W. 1978 Liposomes from ionic, single-chain amphiphiles. *Biochemistry* **17**, 3759–3768. (doi:10.1021/bi00611a014)
- Hazen, R. M., Filley, T. R. & Goodfriend, G. A. 2001 Selective adsorption of L- and D-amino acids on calcite: implications for biochemical homochirality. *Proc. Natl Acad. Sci. USA* **98**, 5487–5490. (doi:10.1073/pnas.101085998)
- Hedges, J. T. & Kell, R. G. 1995 Sedimentary organic matter preservation: an assessment and speculative synthesis. *Mar. Chem.* **49**, 81–115. (doi:10.1016/0304-4203(95)00008-F)
- Huber, C. & Wächtershäuser, G. 1997 Activated acetic acid by carbon fixation on (Fe,Ni)S under primordial conditions. *Science* **276**, 245–247. (doi:10.1126/science.276.5310.245)
- Huber, C. & Wächtershäuser, G. 1998 Peptides by activation of amino acids with CO on (Ni,Fe)S surfaces: implications for the origin of life. *Science* **281**, 670–672. (doi:10.1126/science.281.5377.670)
- Kell, R. G. 1994 Sorptive preservation of labile organic matter in marine sediments. *Nature* **370**, 549–552. (doi:10.1038/370549a0)
- Leman, L., Orgel, L. & Ghadiri, M. R. 2004 Carbonyl sulfide-mediated prebiotic formation of peptides. *Science* **306**, 283–286. (doi:10.1126/science.1102722)
- McCollom, T. M., Ritter, G. & Simoneit, B. R. T. 1999 Lipid synthesis under hydrothermal conditions by Fischer–Tropsch-type reactions. *Orig. Life Evol. Biosph.* **29**, 153–166. (doi:10.1023/A:1006592502746)
- Martin, W. & Russell, M. J. 2003 On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. *Phil. Trans. R. Soc. B* **358**, 59–85. (doi:10.1098/rstb.2002.1183)
- Mayer, L. M. & Xing, B. 2001 Organic carbon–surface area–clay relationships in acid soils. *Soil Sci. Soc. Am. J.* **65**, 250–258.
- Mitchell, J. K. 1993 *Fundamentals of soil behavior*. New York, NY: Wiley.
- Monnard, P.-A., Apel, C. L., Kanavarioti, A. & Deamer, D. W. 2002 Influence of ionic solutes on self-assembly and polymerization processes related to early forms of life: Implications for a prebiotic aqueous medium. *Astrobiology* **2**, 139–152. (doi:10.1089/15311070260192237)
- Morowitz, H. J. 1992 *Beginnings of cellular life*. New Haven, CT: Yale University Press.
- Orgel, L. 1997 Polymerization on the rocks: theoretical introduction. *Orig. Life Evol. Biosph.* **28**, 227–234. (doi:10.1023/A:1006595411403)
- Pace, N. R. 1991 Origin of life—facing up to the physical setting. *Cell* **65**, 531–533. (doi:10.1016/0092-8674(91)90082-A)
- Rohlfing, D. L. 1976 Thermal polyamino acids: synthesis at less than 100°C. *Science* **193**, 68–70.
- Rushdi, A. I. & Simoneit, B. 2001 Lipid formation by aqueous Fischer–Tropsch type synthesis over a temperature range of 100 to 400°C. *Orig. Life Evol. Biosph.* **31**, 103–118. (doi:10.1023/A:1006702503954)
- Sleep, N. H., Zahnle, K., Kasting, J. F. & Morowitz, H. J. 1989 Annihilation of ecosystems by large asteroid impacts on the early Earth. *Nature* **342**, 139–142. (doi:10.1038/342139a0)
- Walde, P., Goto, A., Monnard, P.-A., Wessicken, M. & Luisi, P. L. 1994 Oparin's reactions revisited: enzymatic synthesis of poly(adenylic acid) in micelles and self-reproducing vesicles. *J. Am. Chem. Soc.* **116**, 7541–7547. (doi:10.1021/ja00096a010)
- Washington, J. 2000 The possible role of volcanic aquifers in prebiological genesis of organic compounds and RNA. *Orig. Life Evol. Biosph.* **30**, 53–79. (doi:10.1023/A:1006692606492)