Prebiotic Chemistry

International Edition: DOI: 10.1002/anie.201704048
German Edition: DOI: 10.1002/ange.201704048

The Narrow Road to the Deep Past: in Search of the Chemistry of the Origin of Life

*Jack W. Szostak**

origin of life · prebiotic chemistry · primer extension · protocells · RNA replication

The sequence of events that gave rise to the first life on our planet took place in the Earth's deep past, seemingly forever beyond our reach. Perhaps for that very reason the idea of reconstructing our ancient story is tantalizing, almost irresistible. Understanding the processes that led to synthesis of the chemical building blocks of biology and the ways in which these molecules self-assembled into cells that could grow, divide and evolve, nurtured by a rich and complex environment, seems at times insurmountably difficult. And yet, to my own surprise, simple experiments have revealed robust processes that could have driven the growth and division of primitive cell membranes. The nonenzymatic replication of RNA is more complicated and less well understood, but here too significant progress has come from surprising developments. Even our efforts to combine replicating compartments and genetic materials into a full protocell model have moved forward in unexpected ways. Fortunately, many challenges remain before we will be close to a full understanding of the origin of life, so the future of research in this field is brighter than ever!

1. Introduction

In the late 1600s, the great poet Bashō set off on his epic journey into the Deep North of Edo period Japan. In his celebrated chronicle of that journey,[1] Bashō explains that he wanted to see for himself those beautiful sites that he had only heard of from others. This Minireview is a brief and personal account of a different and certainly less perilous journey, but one that has also followed a narrow, twisting and difficult to follow path, leading from time to time to surprising and beautiful vistas. Like Bashō, I have not traveled alone, and it is the many brilliant students, postdocs and collaborators who have made this journey possible and who have kept it so interesting and exciting. My decision to commit fully to this exploration of the deep past followed a lengthy series of debates in the 1980s and 90s, between those of us who emphasized the role of nucleic acids and inheritance in the origin of life,[2,3] and others who emphasized the role of compartmentalization.[4,5] Although it was obvious to many that both compartments and informational polymers are crucial and universal aspects of cellular life,[6,7] it took some time for me to fully appreciate that the important question was how such a combined system could arise.[8] In other words, to understand the origin of cellular life we need to understand the transition from a collection of biological building blocks to the assembly of a protocell capable of growth, division and Darwinian evolution. Since that realization I have focused the efforts of my laboratory towards the synthesis of simple living systems, in the hope that this approach might give us clues as to how life began spontaneously on the early Earth. I will not address the great advances that have been made in the past decade in understanding the prebiotic synthesis of the nucleotides, amino acids and lipids needed to build biology;[9] instead my focus will be on the self-assembly processes that resulted in the formation and replication of the first cells.

[*] Prof. J. W. Szostak
Howard Hughes Medical Institute, Department of Molecular Biology and Center for Computational and Integrative Biology, Massachusetts General Hospital
Boston, MA 02114 (USA)
E-mail: szostak@molbio.mgh.harvard.edu

2. The Protocell Concept

Our work is based on the idea that the first living cells emerged from a large population of vesicles containing different and more or less random-sequence oligonucleotides; I say more or less because the sequences of these primordial oligonucleotides would of course have been biased by the

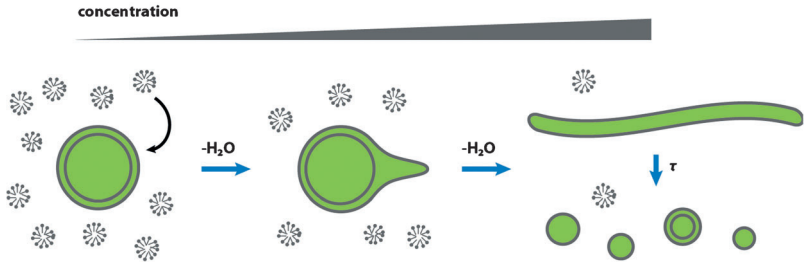


Figure 1. Concentration driven growth of protocell membranes, followed by turbulence induced division. Adapted with permission from Ref. [16]. Copyright 2012 American Chemical Society.

chemical and physical influences on their synthesis, degradation and replication.[10] If the vesicles could grow and divide and their genetic contents could replicate, and if some, probably very rare, nucleic acid sequences could form a catalyst or a structure that provided some benefit to their host cell, the stage would be set for the emergence of Darwinian evolution and thus life itself. We now understand that a variety of simple processes can drive vesicle growth and division, under prebiotically plausible early Earth conditions.[11] On the other hand, our understanding of chemical (i.e. nonenzymatic) replication of the primitive genetic material is less advanced, and is the current subject of intense research in a number of laboratories.[12] Continuing cycles of nucleic acid replication are likely to require the assistance of short peptides, and defining potential roles of peptides is an exciting new aspect of protocell research. The growth and division of protocells would clearly require a chemically rich and physically complex environment that could provide the materials and sources of energy necessary to drive cell reproduction. We are attempting to define such environments, and to show through laboratory experiments how protocell replication might be accomplished.

3. Growth and Division: The Environment in Control

The first cells had by definition no evolved machinery that could control their growth and division, leaving the success or failure of these critical processes to the vagaries of the environment. What reasonable environments could drive the growth of protocell membrane compartments? Perhaps the simplest idea is an intermittent source of additional membrane forming molecules that could, in effect, feed the vesicles. This is easy to imagine with fatty acids, because of their pH-dependent phase transition from small micelles at high pH to bilayer membranes at neutral to mildly alkaline pH.[13] When an alkaline solution of fatty acid micelles is added to preformed vesicles in a buffered solution, the micelle phase becomes thermodynamically unstable and some of the added fatty acid molecules insert into the preformed vesicle bilayers, leading to growth of the membrane.[14] Remarkably this leads to the formation of long filamentous vesicles, which can then be induced to divide by gentle shear forces, or by chemical changes in the solution.[15] A geochemical version of this process would require an intermittent flow of an alkaline stream carrying fatty acids into a pond or lake at lower pH, in which the protocells lived. While not perhaps impossible, the complexity of this scenario did stimulate us to search for simpler possibilities. For example, simple evaporation can lead to an increase in the concentration of free fatty acids, which drives them into the bilayer phase, again causing filamentous growth, which could be followed by turbulence-induced division (Figure 1).[16] Rain could then return the overall concentration to the original level, probably dissolving some fraction of the vesicles in the process, and setting the stage for the next cycle of evaporation-induced growth and division.

An attractive alternative to such scenarios is growth by competition for limiting membrane components. Our first hint that competitive growth was possible came from studies of osmotically swollen fatty acid vesicles, which can grow at the expense of relaxed vesicles, which shrink.[17] However, growth of a swollen vesicle leads to a swollen sphere, which is difficult to divide. A few years later we took up this trail again, when our studies of the evolutionary transition from primitive membranes composed of single-chain amphiphiles to modern two-chain amphiphiles led to the surprising discovery that a small fraction of two-chain phospholipids in the membrane of a fatty acid vesicle allowed such vesicles to grow at the expense of neighboring vesicles that lacked phospholipids.[18] Two mechanisms drive this remarkable effect. The first is simply the entropically favored dilution of the phospholipid during growth. The second effect is that the two chain phospholipids induce order in the membrane, slowing the

Jack W. Szostak is an Investigator of the Howard Hughes Medical Institute, Professor of Genetics at Harvard Medical School, Professor of Chemistry and Chemical Biology at Harvard University, and the Alex Rich Distinguished Investigator in the Dept. of Molecular Biology and the Center for Computational and Integrative Biology at Massachusetts General Hospital. His research on telomeres was recognized by the 2006 Albert Lasker Basic Medical Research Award and the 2009 Nobel Prize in Physiology or Medicine. In the 1990s he developed in vitro selection as a tool for the isolation of functional RNA, DNA and protein molecules from large pools of random sequences. His current research interests are in the laboratory synthesis of self-replicating systems and the origin of life.

dissociation of fatty acids from the membrane, but leaving the association rate approximately unchanged. Thus, vesicles whose membranes contain some phospholipid tend to grow by absorbing fatty acids from vesicles that contain no or less phospholipid. Again, growth into long filaments was observed, as was shear induced division. Thus any heritable catalyst of phospholipid synthesis would drive vesicle growth, leading to a large competitive advantage.

Phospholipid induced growth is an attractive process for several reasons. First, no periodic or even sporadic input of nutrients is needed. Second, the process is competitive, thus setting the stage for a Darwinian evolutionary process with a strong selective pressure for the synthesis of higher levels of phospholipids. The catalyst of phospholipid synthesis would have to be heritable, which leads us to ask whether an RNA encoded ribozyme might do the trick, or perhaps a ribozyme that synthesized a peptide catalyst. We showed experimentally that protocells with higher levels of phospholipids in their membranes could grow at the expense of surrounding cells with lower levels of phospholipids. As a result this simple physical effect would lead to an evolutionary arms race, with pressure to increase phospholipid levels until the biophysical properties of the cell membrane began to change and introduce new limitations, for example, slower import of essential nutrients such as nucleotides synthesized in the external environment. This would lead to a cascade of new selective pressures favoring the evolution of primitive transport machinery. Some of this was probably peptide based; for example, primitive ion transporters may have been similar to cyclic depsipeptides such as valinomycin that catalyze the exchange of potassium ions across the cell membrane.[19] The growing importance of transport and catalytic peptides would then lead to the evolutionary optimization of peptide synthesis, first in a non-coded fashion by an ancestral version of the 50S ribosomal peptidyl synthase center, and later on through the addition of the small subunit decoding center.

The evolutionary transition to less permeable membranes would also make internal metabolic activities more beneficial to the primitive cell, because useful metabolic products would no longer leak out and feed neighboring cells. We therefore hypothesized that the changing biophysical properties of primitive cell membranes may have triggered the evolution of genetically encoded internal metabolic pathways.

The above scenario implies a chemically rich and physically active environment. For example, the substrates for phospholipid synthesis would have to have been available through prebiotic chemical processes. Given that the environment would also have had to support the synthesis and replication of nucleic acids, this does not seem like a large additional hurdle. In modern organisms, two-chain phospholipids are synthesized by two similar pathways.[20] In both, one substrate is a lysophospholipid, that is, an acyl chain esterified to glycerol phosphate. The second substrate is a fatty acid in which the carboxylate is activated as either a thioester or a carboxyphosphate anhydride. If the ambient chemical environment could generate such substrates from the fatty acids in protocell membranes, then two-chain phospholipid synthesis requires only one acyl transfer step, which in principle could be catalyzed by either a ribozyme or a peptide. Thus, the chemistry required to drive a transition in membrane composition does not seem terribly complex, and similar ribozyme-catalyzed acyl-transfer reactions have been described.[21] The reconstitution of this scenario in the laboratory is thus an exciting goal for those of us studying potential pathways for the emergence of life.

4. Surprises in RNA Replication: Covalent Nucleophilic Catalysis

In order for any evolutionary process to take place, a mechanism for the inheritance of useful functions must exist. RNA-based systems are attractive because inheritance and function can be embodied within the same class of molecules. Proposals for RNA based functional genomes were made in the late 1960s,[22] as it became apparent that RNA could generate complex folded structures, which it was supposed might then be capable of catalysis. However, these ideas languished until the early 1980s, when direct observations of RNA enzymes in biology revolutionized thinking about the origin of life.[23] All attention became focused on the possibility of RNA-catalyzed RNA replication, and the RNA World hypothesis was born.[24] Over the subsequent decades, the remarkable catalytic abilities of ribozymes have been explored, including the directed evolution of ribozymes that are increasingly effective RNA polymerase enzymes,[25] although these are still not quite good enough to catalyze their own replication.

What would be required in order for the first ribozymes to emerge? An initial stage of chemical (i.e. nonenzymatic) RNA replication would seem to be necessary to set the stage. The prebiotic chemistry leading to nucleotide synthesis is an active area of study in several laboratories, and great progress has been made.[9] There are still gaps in our understanding of how activated 5′-phosphorylated nucleotides could have been generated, but it has long been clear that the availability of such activated monomers makes the synthesis of RNA chains relatively straightforward. For example, activated monomers can polymerize on mineral surfaces,[26] or simply by freezing,[27] which results in very high concentrations of monomers in between the water ice crystals so that proximity induced polymerization follows. However, the interesting and difficult challenge is to show how RNA strands, once generated, could be replicated prior to the emergence of the first enzymes.[12,28] Once chemical RNA replication inside protocells could occur, some sequence or set of short oligonucleotides that could assemble into a useful ribozyme would eventually emerge. This could be any ribozyme, or even a structural RNA, that provided some advantage in survival or reproduction to its host protocell; one potential example out of many possibilities would be the phospholipid synthase discussed above. At that point, there would have been strong selective pressure to replicate that beneficial sequence more efficiently, rapidly and accurately, possibly leading to a rapid increase in the complexity of the RNA replication machinery, which would in turn allow for the evolution and maintenance of more and more ribozymes. Thus I suggest that the emergence of the first ribozyme was the trigger that led inevitably to the evolution

of increasingly complex cells and the development of metabolism, coded translation, archival storage of large amounts of information in DNA, and so on.

All of this brings us back to our original question of how chemistry could drive RNA replication. This is not a new question by any means, and experimental work began in the late 1960s,[29] and was intensively pursued by Leslie Orgel and his students and colleagues from the 1970s through the 1990s.[7] Great progress was made during this early work, culminating in the discovery that nucleotides activated as 2-methyl-phosphor-imidazolides were excellent substrates for template-directed nonenzymatic primer extension.[30] Indeed short oligo C templates could be rapidly copied by activated G monomers. However, the copying of mixed sequence templates was inefficient, and templates containing all four nucleotides could not be copied at all, except under extreme conditions that are incompatible with a protocell environment. By the late 1990s, faith in the pure RNA World idea was waning, and the search for simpler progenitors of RNA began, in the hope that some plausible genetic molecule would be found that would be easier to synthesize and to replicate than RNA, and which could eventually be replaced by RNA at some later stage in the evolution of life. While this did indeed lead to a flowering of discovery and the synthesis of many beautiful nucleic acids that are alternative Watson–Crick base pairing systems,[31] so far none seem easier to synthesize or replicate than RNA. Meanwhile, as noted above, prebiotic chemical routes to the ribonucleotides have been emerging. Taken together, these developments led me to reexamine the problems that have long stalled efforts to replicate RNA in a purely chemical system.[28]

Here I will focus on the key issue of how to copy mixed sequence RNA templates in an efficient and general manner. While this is far from a solved problem, considerable progress has been made, and I suspect that the nonenzymatic replication of RNA oligonucleotides long enough to fold into functional structures will be possible within a few years. What developments have led to our current ability to copy mixed sequence templates? The solution emerged from mechanistic studies that have both changed our understanding of the fundamental chemistry of nonenzymatic primer extension, and have allowed us to develop new strategies that enable more efficient template copying. For many years, the reaction of a primer with an incoming activated monomer (e.g. a 2-methylimidazolide) was thought to be a classical S_N2 in-line nucleophilic substitution reaction, where the 3′-hydroxyl of the primer would attack the phosphorus of the incoming nucleotide, displacing its methylimidazole leaving group.[32] However, even very early experiments hinted that this was not the whole story. For example, on oligo-C templates, internal nucleotides were copied rapidly, but the last nucleotide was incorporated much more slowly.[33] In 1992, Wu and Orgel showed that the nucleotide downstream of the one adjacent to the primer played an important catalytic role, and they even suggested that this catalytic role was due to some interaction between the leaving groups of adjacent template bound monomers.[33] Unfortunately these seminal observations were never followed up, and the underlying basis of these observations remained unstudied for the following 25 years.

Our encounter with this puzzle came about, ironically, through a desire to explore alternatives to primer extension with monomers, which is a very biologically inspired model for replication.[34] I thought that the assembly of monomers into short oligonucleotides, followed by the template directed ligation of these oligos into longer and longer oligos, might lead more effectively to template copying, in part by avoiding the problems caused by the weak base-pairing of A and U, and in part by providing a more rigid framework for the chemical reaction step (Figure 2). To test this idea, we decided to compare the rate of ligation and monomer addition reactions, in an experiment designed to make the chemical step identical in each case.[35] To do this, we used a primer ending in G, and monitored the template-directed reaction of the primer with either incoming activated G monomers or with an oligonucleotide with an identically activated 5′-G residue. We expected the ligation reaction to be faster, since the primer and downstream oligonucleotide would be more pre-organized in the required A-type helical geometry. We were extremely surprised when the monomer addition reaction turned out to be about 100-fold faster than the ligation reaction. Follow up experiments showed that the critical difference was the presence of a downstream activated nucleotide in the monomer addition reaction. We thus rediscovered, after a 25 year lag, Orgel's earlier insight! At this point, we did go slightly further, in showing that while a downstream activated monomer was a good catalyst, a short downstream activated oligonucleotide was much better, with

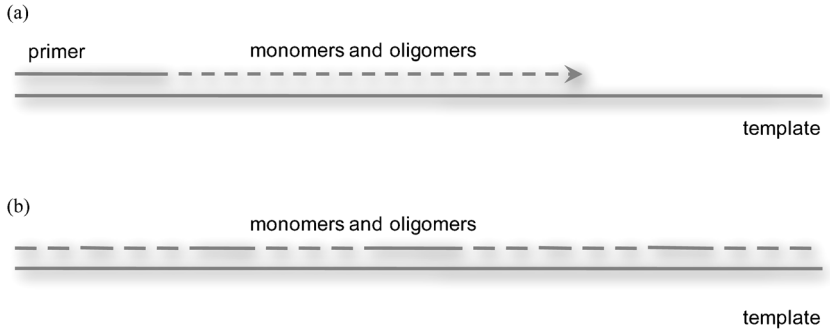


Figure 2. RNA replication models. a) Biologically inspired primer extension with monomers. b) Hierarchical assembly by ligation of oligonucleotides. From Ref. [34].

activated trimers providing up to a 1000-fold rate enhancement. We were able to take advantage of this observation to copy mixed sequence templates, by iterating primer extension one step at a time using pairs of activated monomers and corresponding downstream activated trimers. While more complex than simple monomer addition, this setup allowed us for the first time to copy templates containing all four nucleotides in a one pot reaction under mild conditions.

What was the mechanism of this surprising catalytic effect? At first we assumed that a non-covalent interaction of the 2-methylimidazole leaving groups on adjacent monomers somehow aligned the upstream leaving group for in-line attack by the primer hydroxy. Indeed, MD simulations showed that the leaving groups of adjacent monomers could potentially interact in a number of ways, including π-stacking, cation–π and hydrogen bonding interactions [Li and Szostak, unpublished]. To our surprise, stable phosphonate analogs of the reactive imidazolide substrates failed to show significant catalytic activity, and crystallographic studies of monomers bound to templates failed to show any favored non-covalent interaction between adjacent leaving groups.[36] Finally, careful and detailed kinetic studies set us back on the correct path. Our initial attempts to carry out quantitative reaction kinetics were frustrated by irreproducible results, suggesting that some key variable was not being controlled. Considerable efforts to prepare highly pure monomers, under standardized conditions, failed to solve the problem. Ultimately it turned out that the hidden variable was the pH of the activated monomer solution, *before* it was added to the primer extension reaction (which of course was always carried out in highly buffered solution at a controlled pH).[37] The optimum pH for monomer pre-incubation was at the pK_a of the leaving group, consistent with an interaction between a protonated and an unprotonated monomer. To our surprise, preincubation of monomer at this optimal pH required roughly 20 minutes to attain optimal primer-extension activity, and this observation suggested that we should look for the accumulation of a covalent intermediate. NMR experiments then quickly led to the identification of an imidazolium-bridged dinucleotide (Figure 3) as a candidate for the intermediate. Partial purification of the intermediate showed that it was highly reactive in primer extension, and monitoring the concentration of the intermediate suggested that its formation is sufficient to account for all primer extension. The Richert laboratory has also noted the high reactivity of imidazolium-bridged dinucleotides.[38] Our current model for the nonenzymatic primer extension reaction is that monomers (or activated helper oligonucleotides) react first to form an imidazolium-bridged intermediate. In a second step, the primer 3′-hydroxy attacks the phosphate of the upstream nucleotide, displacing the entire downstream activated monomer (or oligonucleotide) as the true leaving group.

Recognition of the fact that upstream and downstream monomers play distinct roles in the primer extension reaction allowed us to perform structure–activity relationship (SAR) studies on the leaving group, looking specifically for improved downstream catalytic activity.[39] A small screen of substituted imidazoles revealed that the substituent at the 2-position had to be small, and that higher pK_a was beneficial. These

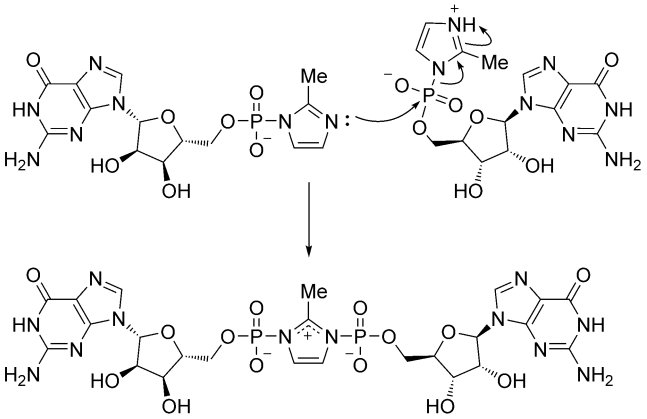


Figure 3. Formation of the imidazolium-bridged dinucleotide intermediate in primer extension by reaction of two activated mononucleotides. Adapted with permission from Ref. [37]. Copyright 2016 American Chemical Society.

constraints led us to examine 2-aminoimidazole as an activating group; indeed, 2-aminoimidazole turned out to be superior to 2-methylimidazole as an activating group, in both upstream and downstream positions. Using this new activating group in combination with the previous activated monomer plus activated trimer strategy has resulted in greatly improved template copying activity, such that mixed sequence templates can be fully extended for up to 7 nucleotides. We are continuing to search for additional improvements that will allow for the efficient copying of even longer templates.

5. All Together Now: Compatibility of Vesicle and Genetic Systems

The spontaneous primer extension reaction described above requires metal catalysis to achieve significant rates. Typically, Mg^{2+} is used, but the binding of the divalent cation to the reaction center is weak, so that high concentrations on the order of 50–100 mM are commonly used. In addition to being implausible in an environment that presumable must contain significant free phosphate, such high concentrations of Mg^{2+} are problematic for other reasons. Most dramatically, high concentrations of divalent cations cause immediate disruption of fatty acid based membranes, and they also catalyze the degradation of RNA. How can we reconcile the need for metal ion catalysis for RNA replication, with the need for a low-metal environment for membrane stability and RNA integrity?

In approaching the difficult problems of the origin of life, it is often necessary to begin with a proof of principle experiment that provides a fresh view of the landscape. Later work may then lead to more prebiotically plausible solutions to the problem. At first, we simply side-stepped the problem by showing that a phosphoramidate nucleic acid, which does not require Mg^{2+} for polymerization, could be synthesized by template-directed primer extension within fatty acid vesicles.[40] This was a satisfying step, as it showed that there were

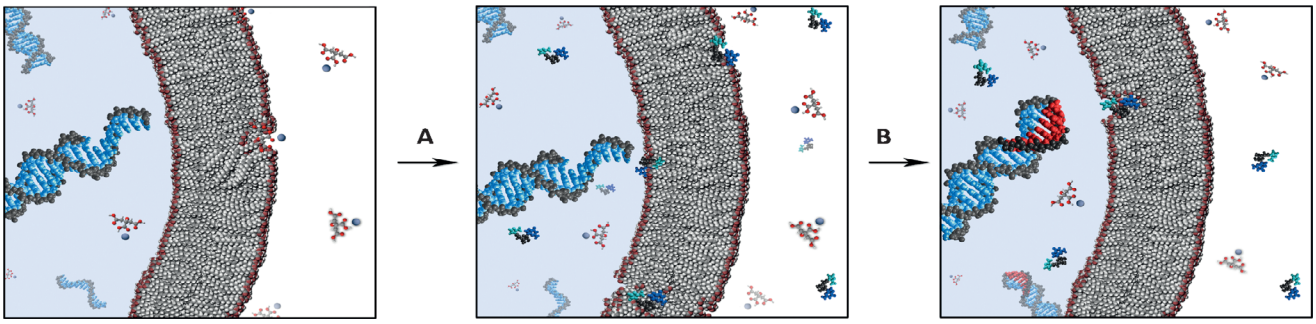


Figure 4. Copying of a template by primer extension within a fatty acid vesicle. A) activated nucleotides diffuse across the membrane to the vesicle interior, where they B) take part in template-directed primer extension. Modified from Ref. [41].

no other hidden problems in combining the membrane and genetic systems. Our second step was another proof of principle experiment, this time showing that RNA chemistry and vesicle integrity could be made to be compatible, albeit in a somewhat artificial way.[41] A small screen of di- and tricarboxylic acids for activity as chelators of Mg^{2+} ions revealed that citrate binds to Mg^{2+} in such a way that it protects fatty acid membranes from disruption by Mg^{2+}, and yet interference with catalysis of RNA copying is minimal. These observations allowed us to perform an experiment that we had wanted to do for many years, namely copy an RNA template inside a fatty acid vesicle (that we already knew could grow and divide indefinitely). To do this, we first encapsulated a primer–template complex inside oleate vesicles, and purified away all unencapsulated material. We then added Mg^{2+}, citrate, and activated monomer, all of which spontaneously crossed the membrane, after which primer extension commenced inside the vesicles.[41] This experiment is important because it shows that if the remaining problems in RNA replication can be solved, then RNA replication within vesicles should be possible (Figure 4).

As pleased as we were with the ability of citrate to enable RNA copying within vesicles, we had to recognize that the presence of sufficient citrate in early protocell environments is unlikely. What are the alternatives? Here biology may be giving us a clue, for in the heart of one of the most important of all enzymes, cellular RNA polymerase, lies a small peptide loop that coordinates the catalytic metal ion via the carboxylates of three aspartate residues.[42] Could small asparate-rich peptides play the role of citrate? If there was a way to bring such a peptide close to the reaction center on the primer–template–monomer complex, it seems possible that the small peptide could act as a primitive enzyme by increasing the local concentration of Mg^{2+}; we are currently exploring this possibility. This approach has the added benefit that only low overall Mg^{2+} concentrations would be needed, on the order of the K_d of the peptide for the metal ion, suggesting that protocells might be able to originate in an environment with low Mg^{2+} concentrations. If we are able to demonstrate a role for simple peptides in RNA replication within vesicles, we could be seeing the earliest beginnings of the modern nucleic acid–peptide–membrane framework of biology.

6. Evolution at Last: The Beginnings of Biology

The very first cells, because of their reliance on the environment for growth, division and genomic replication, could "live long and prosper" when conditions were favorable, but would have been "hungry, just barely alive" when conditions became adverse. Similarly, the spread of these primitive organisms from limited local environments with ideal conditions to the wide world beyond would have required a lengthy series of evolutionary adaptations, each providing some incremental (or perhaps enormous) advantage in ability to survive or propagate in increasingly challenging environments. To begin to see such changes in controlled laboratory experiments is the ultimate goal of all of the work I have described above. Eventually it may even be possible to recreate the long lost path from chemistry to the simplest beginnings of life and on to the complexity of modern life, replete with DNA, a translational apparatus, and enzyme catalyzed metabolic complexity.

Conflict of interest

The author declares no conflict of interest.

[1] M. Bashō, *The Narrow Road to the Deep North*, **1694**.
[2] D. P. Bartel, J. W. Szostak, *Science* **1993**, *261*, 1411.
[3] a) T. R. Cech, *Gene* **1993**, *135*, 33; b) G. F. Joyce, *Nature* **2002**, *418*, 214.
[4] a) R. Wick, P. Walde, P. L. Luigi, *J. Am. Chem. Soc.* **1995**, *117*, 1435; b) F. Mavelli, P. L. Luisi, *J. Phys. Chem.* **1996**, *100*, 16600.
[5] D. W. Deamer, *Microbiol. Mol. Biol. Rev.* **1997**, *61*, 239.
[6] T. Gánti, *BioSystems* **1975**, *7*, 15.
[7] L. E. Orgel, *Nature* **1992**, *358*, 203.
[8] J. W. Szostak, D. P. Bartel, P. L. Luisi, *Nature* **2001**, *409*, 387.
[9] a) M. W. Powner, B. Gerland, J. D. Sutherland, *Nature* **2009**, *459*, 239; b) B. J. Cafferty, N. V. Hud, *Curr. Opin. Chem. Biol.* **2014**, *22*, 146; c) J. D. Sutherland, *Angew. Chem. Int. Ed.* **2016**, *55*, 104; *Angew. Chem.* **2016**, *128*, 108.
[10] I. A. Chen, M. A. Nowak, *Acc. Chem. Res.* **2012**, *45*, 2088.
[11] I. Budin, J. W. Szostak, *Annu. Rev. Biophys.* **2010**, *39*, 245.
[12] a) F. R. Bowler, C. K. Chan, C. D. Duffy, B. Gerland, S. Islam, M. W. Powner, J. D. Sutherland, J. Xu, *Nat. Chem.* **2013**, *5*, 383; b) J. C. Blain, J. W. Szostak, *Annu. Rev. Biochem.* **2014**, *83*, 615;

c) M. Jauker, H. Griesser, C. Richert, *Angew. Chem. Int. Ed.* **2015**, *54*, 14559; *Angew. Chem.* **2015**, *127*, 14767.
[13] T. H. Haines, *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 160.
[14] M. M. Hanczyc, S. M. Fujikawa, J. W. Szostak, *Science* **2003**, *302*, 618.
[15] a) T. F. Zhu, J. W. Szostak, *J. Am. Chem. Soc.* **2009**, *131*, 5705; b) T. F. Zhu, K. Adamala, N. Zhang, J. W. Szostak, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 9828.
[16] I. Budin, A. Debnath, J. W. Szostak, *J. Am. Chem. Soc.* **2012**, *134*, 20812.
[17] I. A. Chen, R. W. Roberts, J. W. Szostak, *Science* **2004**, *305*, 1474.
[18] I. Budin, J. W. Szostak, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 5249.
[19] H. Brockmann, H. Geeren, *Liebigs Ann. Chem.* **1957**, *603*, 216.
[20] Y. M. Zhang, C. O. Rock, *Nat. Rev. Microbiol.* **2008**, *6*, 222.
[21] a) H. Suga, P. Lohse, J. W. Szostak, *J. Am. Chem. Soc.* **1998**, *120*, 1151; b) N. Li, F. Huang, *Biochemistry* **2005**, *44*, 4582.
[22] a) C. R. Woese, in *The Genetic Code: The Molecular Basis for Genetic Expression*, *Vol. 8*, Harper & Row, New York, **1967**, pp. 179–195; b) F. H. C. Crick, *J. Mol. Biol.* **1968**, *38*, 367; c) L. E. Orgel, *J. Mol. Biol.* **1968**, *38*, 381.
[23] a) K. Kruger, P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling, T. R. Cech, *Cell* **1982**, *31*, 147; b) C. Guerrier-Takada, K. Gardiner, T. Marsh, N. Pace, S. Altman, *Cell* **1983**, *35*, 849.
[24] W. Gilbert, *Nature* **1986**, *319*, 618.
[25] a) W. K. Johnston, P. J. Unrau, M. S. Lawrence, M. E. Glasner, D. P. Bartel, *Science* **2001**, *292*, 1319; b) A. Wochner, J. Attwater, A. Coulson, P. Holliger, *Science* **2011**, *332*, 209; c) D. P. Horning, G. F. Joyce, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 9786.
[26] J. P. Ferris, A. R. Hill Jr, R. Liu, L. E. Orgel, *Nature* **1996**, *381*, 59.
[27] P. A. Monnard, A. Kanavarioti, D. W. Deamer, *J. Am. Chem. Soc.* **2003**, *125*, 13734.
[28] J. W. Szostak, *J. Syst. Chem.* **2012**, *3*, 2.
[29] B. J. Weimann, R. Lohrmann, L. E. Orgel, H. Schneider-Bernloehr, J. E. Sulston, *Science* **1968**, *161*, 387.
[30] T. Inoue, L. E. Orgel, *J. Mol. Biol.* **1982**, *162*, 201.
[31] K. Schoning, P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy, A. Eschenmoser, *Science* **2000**, *290*, 1347.
[32] L. E. Orgel, *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 99.
[33] T. Wu, L. E. Orgel, *J. Am. Chem. Soc.* **1992**, *114*, 5496.
[34] J. W. Szostak, *Philos. Trans. R. Soc. B* **2011**, *366*, 2894.
[35] N. Prywes, J. C. Blain, F. Del Frate, J. W. Szostak, *eLife* **2016**, *5*, e17756.
[36] W. Zhang, C. P. Tam, J. Wang, J. W. Szostak, *ACS Cent. Sci.* **2016**, *2*, 916.
[37] T. Walton, J. W. Szostak, *J. Am. Chem. Soc.* **2016**, *138*, 11996.
[38] E. Kervio, M. Sosson, C. Richert, *Nucleic Acids Res.* **2016**, *44*, 5504.
[39] L. Li, N. Prywes, C. P. Tam, D. K. O'Flaherty, V. S. Lelyveld, E. C. Izgu, A. Pal, J. W. Szostak, *J. Am. Chem. Soc.* **2017**, *139*, 1810.
[40] S. S. Mansy, J. P. Schrum, M. Krishnamurthy, S. Tobé, D. A. Treco, J. W. Szostak, *Nature* **2008**, *454*, 122.
[41] K. Adamala, J. W. Szostak, *Science* **2013**, *342*, 1098.
[42] a) G. Zhang, E. A. Campbell, L. Minakhin, C. Richter, K. Severinov, S. A. Darst, *Cell* **1999**, *98*, 811; b) P. Cramer, D. A. Bushnell, R. D. Kornberg, *Science* **2001**, *292*, 1863.

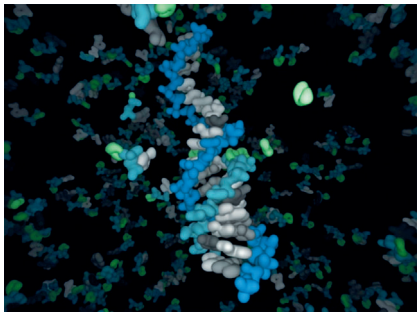
Manuscript received: April 19, 2017
Accepted manuscript online: May 17, 2017
Version of record online: ■■ ■■, ■■■■

Minireviews

Prebiotic Chemistry

J. W. Szostak* _____ ■■■■–■■■■

The Narrow Road to the Deep Past: in Search of the Chemistry of the Origin of Life



Life from chemistry: How did the first living cells reproduce before the evolution of enzymes? Primitive cell membranes can grow and divide in response to a variety of environmental fluctuations or competitive situations. Recent progress brings us closer to understanding the nonenzymatic replication of the primordial genetic material, but many puzzles remain.