PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B

Phil. Trans. R. Soc. B (2011) **366**, 2853–2856
doi:10.1098/rstb.2011.0133

Introduction

The chemical origins of life and its early evolution: an introduction

David M. J. Lilley[1,*] **and John Sutherland**[2]

[1]*Cancer Research UK Nucleic Acid Structure Research Group, MSI/WTB Complex, The University of Dundee, Dow Street, Dundee DD1 5EH, UK*
[2]*MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, UK*

Can we look at contemporary biology and couple this with chemical insight to propose some plausible mechanisms for the origin of life on the planet? In what follows, we examine some promising chemical reactions by which the building blocks for nucleic acids might have been created about a billion years after the Earth formed. This could have led to self-assembling systems that were based on an all-RNA metabolism, where RNA is both catalytic and informational. We consider the breadth of RNA enzymes presently existing in biology, and to what extent these might have covered a wider range of chemistry in the RNA world. Ultimately, the RNA world would probably have given way to protein-based life quite quickly, and the origins of peptidyl transferase activity are discussed below.

Keywords: prebiotic chemistry; self-assembling systems; RNA world; ribozymes; ribosome

1. INTRODUCTION

A discussion meeting on the Origin of Life was held at the Royal Society in February 2011. The meeting had been conceived about 18 months earlier, with the intention of bringing together people from a number of different backgrounds to discuss plausible chemical scenarios for how life might have originated on this planet. We decided to take a distinctly RNA-centric view of this problem, and for this purpose we were rather selective about which topics to focus on—perhaps a little more focused than our title might have suggested. Our plan for the meeting was to discuss the possible chemistries that could have generated the building blocks for nucleic acids, the likely components of a primitive RNA-based life and then how to make the transition to living cells that use proteins. Put another way, our central theme was the RNA world, how it came to be, and how life emerged from it.

We are not experts on the origins of life. But who is? There is no record left of that time, no fossils. All we can do is to look at contemporary life and try to work backwards using imagination guided by chemical insight. In this, we may be helped by the incremental nature of evolution that has shaped the subsequent development of life. Evolution is what Crick called the 'great tinkerer'. It cannot start with a clean sheet, it neither designs nor is intelligent. It blindly selects modifications of what has come before. Thus, for example, contemporary peptidyl transferase may not be the way it would be designed from the ground up, but rather reflects a unidirectional evolution from structures in the RNA world. So a careful examination of the current ribosome may provide insight into these origins.

So with that perspective in mind, we have put together a juxtaposition of topics that we hope will provide some insight into the process by which life might have arisen. What follows is intended to be a kind of roadmap through this hypothetical process.

The solar system formed about 5.5 Ga, and took a billion or so years to cool. From that point on the environment was probably set to allow the chemistry to get going. But where exactly might this have occurred—can we identify candidates for the role of Darwin's 'warm little pond'? Sleep *et al.* [1] (Stanford) argue that a suitable environment might have been provided at submarine hydrothermal vents above serpentinite rocks, where chemical gradients of dissolved hydrogen gas and hydrogen ions could have provided a source of chemical free energy. Primitive life might have been established in pores contained within vent chimneys.

What kind of chemistries might have generated the building blocks on which the earliest life forms might have been based? Everyone knows the early experiments of Stanley Miller showing that subjecting a gaseous chemical mixture that might resemble the early Earth's atmosphere to electrical discharge could produce relatively complex compounds such as amino acids. But if we make the assumption for now that nucleic acids were required at an early stage, can we construct a plausible scenario that could have led to prebiotic synthesis of nucleosides, the basic building blocks of RNA? While it was previously assumed that

* Author for correspondence (d.m.j.lilley@dundee.ac.uk).

One contribution of 17 to a Discussion Meeting Issue 'The chemical origins of life and its early evolution'.

nucleosides would be made by joining pre-assembled ribose and nucleobase entities, Powner & Sutherland [2] have shown that a complete pyrimidine nucleoside can be built in a single entity beginning with a reaction between cyanamide and glycolaldehyde. Of course, an important aspect of all biological macromolecules is that they are constructed from chiral monomers, and the possible origins of enantiomerically pure amino acids and ribonucleotides are discussed by Blackmond [3] (Scripps).

We next consider how self-assembling systems could have arisen at the early stages of life. This would require some kind of compartmentalization to form protocells, possibly involving lipid membranes. Hanczyc [4] (Odense) shows how self-assembled oil droplets can move around, powered by the hydrolysis of oleic anhydride, exhibiting a kind of chemotaxis sensitive to pH gradients. Szostak [5] (Harvard Medical School) discusses membrane self-assembly, and the advantages of mixtures over single compounds and systems chemistry in general.

At this point, we must address one of the main themes of the discussion meeting, the proposed central role of RNA. In all living cells in the present biological world, there is a marked division of labour between the macromolecules used to store and transmit genetic information (nucleic acids) and those that perform the processing (proteins). Proteins are the chemical workhorse of the cell that keeps its metabolism running along. They are well suited to the role of chemical catalysis, being able to bring a range of side chains with different chemical structures to bear on this task. The nucleic acids seem indispensable, because without their capacity to store information there can be no evolution, and without that there can be no refinement of primitive organisms. But a massive chicken and an egg problem emerges in thinking about how all this got started. In contemporary biology, nucleic acids require enzymes to make their precursors and for replication, and proteins require nucleic acids to provide the information for their synthesis. It is hard to imagine how this complicated interwoven system began. On the other hand, if one type of molecule carried out all the required functions, informational and catalytic, then this would simplify the process greatly. So it was proposed many years ago that perhaps an early form of self-replicating life was based solely on RNA, with no proteins involved. RNA contains information in the order of its nucleotide bases, and could specify its own replication. But could it also act as a primitive enzyme? For a long time, there was very little precedent for the idea. However, the discovery of the group I intron [6] and ribonuclease P [7] ribozymes, and perhaps the lead ion-induced cleavage reaction of tRNA [8] provided that precedent, increasing the possibility of RNA-based life existing approximately 3.6 Ga. This has been called the RNA World hypothesis, as elegantly set out recently by Mike Yarus (Boulder) [9].

RNA occupies a small fraction of the chemical space that proteins cover, consisting of ribose-phosphate and four rather similar heterocyclic bases. Unlike proteins, RNA is a poly-anion, and will be associated with metal ions to neutralize charge. A few ribozymes can be found in contemporary cells, including the above-mentioned RNase P and self-splicing introns—in addition, mRNA splicing in eukaryotes is probably catalysed at least in part by the U2 : U6 small nuclear RNA. What is arguably the most important reaction in the cell, the condensation of amino acids to form polypeptides, is also RNA catalysed. Thus, the ribosome is a ribozyme, and ancient. RNase P is present in all domains of life, and Altman [10] (Yale) argues persuasively that this is also an ancient RNA.

Peptidyl transferase aside, all the remaining ribozymes we find catalyse phosphoryl transfer reactions, either hydrolysis or transesterification. Precedent from the protein world suggests that there are two ways to achieve this, using either metal ions or general acid–base catalysis. Lilley [11] argues that all the nucleolytic ribozymes use variations on the theme of general acid–base catalysis, predominantly using nucleobases. Catalytic efficiency is inherently limited by the unfavourable pK_a values of these, but if this is factored out then these ribozymes are potentially almost as good as a protein like ribonuclease A. As we move to the larger ribozymes, such as RNase P and the self-splicing introns, we seem to switch mechanisms to metal ion-based catalysis.

GlmS provides a fascinating glimpse into how the limitations of chemical functionality in the nucleobases could be expanded by recruitment of a small-molecule cofactor. Here, the role of the general acid in cleavage is taken by glucosamine-6-phosphate, as shown in the crystal structures of Hiller & Strobel [12] (Yale) and Ferré-D'Amaré [13] (National Institutes of Health). The role of guanosine in the group I intron ribozyme is another pointer, although this is not catalytic. This extension of ribozyme activity might also suggest that there could be other ribozymes using cofactors in modern cells. These might be using quite different chemistry, and so have escaped detection to this point.

However, for now, the range of chemistry employed by the known ribozymes is really quite limited, and a workable RNA-based metabolism would have required a much wider range of ribozyme activity. One way to address this is to use *in vitro* selection to isolate activities that might have been essential to the viability of such life. Demonstration of the feasibility of such ribozymes offers a kind of proof of principle of an RNA world. RNA catalysts that accelerate radically different reactions have been isolated, including Diels-Alder pericyclic reactions [14] and Michael addition [15]. Clearly, a fundamental requirement would be a ribozyme that can replicate RNA, and the related ligase, and these have been isolated in the Bartel laboratory [16,17]. Piccirilli & Koldobskaya [18] (Chicago) present the structure of the Bartel ligase RNA, arguing for a two-metal ion mechanism. Would such ribozymes have been efficient enough to support self-sustaining RNA-based organisms? One essential requirement would have been a polymerase activity that could replicate at least a few hundred nucleotides with reasonable accuracy, and this has not yet been achieved by *in vitro* selection, although progress continues to be made. On the other hand, chemists have only been doing this for a few years, while nature had millions of years to experiment, so perhaps we must keep an open mind.

But ultimately, the clear advantages offered by proteins as the cell's chemical engine had to win out, perhaps leaving only the remnants of RNA catalysis in a few remaining ribozymes and coenzymes like NAD. The RNA world might only have lasted a relatively short period of time, and may always have had amino acids and short non-coded peptides associated with it because of concurrent prebiotic chemical synthesis. GlmS may suggest how this transition might have started. Perhaps similar ribozymes recruited amino acids or short peptides to augment their function, and from that point it is a fairly short step conceptually at least to the point at which the protein tail wagged the RNA dog. But the essential requirement of a transition to protein-based life is the creation of some kind of translation system.

The first requirement of translation is to link together the RNA components and the protein building blocks. Yarus [19] discusses some very basic reactions carried out by short oligonucleotides with aminoacyl adenylate, a kind of primitive aminoacylation reaction, and using *in vitro* selection Suga *et al.* [20] (Tokyo) have isolated RNA species that carry out amino acid charging reactions on tRNAs. Convincing thermodynamic arguments made by Pascal & Boiteau [21] (Montpellier) point to a crucial early role for aminoacyl adenylates. These would be synthesized by chemical aminoacylation of ribonucleotides with high free energy amino acid derivatives such as *N*-carboxyanhydrides, rather than by reaction of amino acids with ribonucleoside triphosphates. With 20 different amino acids in contemporary biology, the fidelity of aminoacylation is a challenge, discussed by Schimmel [22] (Scripps).

Using aminoacylated RNA molecules, the final requirement in the transition to the protein world is a translation apparatus. In our cells, this is of course the ribosome. But how did this emerge from the RNA world? The ribosome must carry out two functions, decoding the genetic code of the RNA, and catalysing the condensation of the amino acids to form the polypeptide. These functions are the responsibility of the small and large subunits, respectively. The RNA world hypothesis would require that the components of the translation apparatus would have to be made of RNA initially at least, and remarkably this is still true. Twenty years ago, Noller *et al.* [23] showed that a ribosome severely depleted of protein could still carry out peptidyl transfer, suggesting that perhaps the catalytic component was RNA rather than a protein enzyme. This was confirmed when Steitz *et al.* [24] solved the structure of the large ribosomal subunit, showing that there was no protein within 20 Å of the peptidyl transferase centre. Krupkin *et al.* [25] (Rehovot) have now identified a symmetrical core element in the rRNA that might be a vestige of that primitive peptidyl tranferase centre. It is hard not to be impressed by the thought that this could be a 3.6 billion-year old molecular fossil from an earlier RNA world lying at the heart of our most basic systems for gene expression! Finally, Wohlgemuth *et al.* [26] (Göttingen) bring us to the present day, describing how the modern ribosome achieves high fidelity of translation.

The purpose of this discussion meeting was just that, to promote discussion and make connections between disciplines, and we hope these collections of papers will contribute to the synthesis of ideas in this fascinating field. As organizers, and editors of this volume, we want to thank Christine Phillips and Joanna Bolesworth of the Royal Society, our active session chairs Fritz Eckstein and Venki Ramakrishnan and the more-than 300 people who attended the meeting and contributed to fruitful discussions.

REFERENCES

1 Sleep, N. H., Bird, D. K. & Pope, E. C. 2011 Serpentinite and the dawn of life. *Phil. Trans. R. Soc. B* **366**, 2857–2869. (doi:10.1098/rstb.2011.0129)
2 Powner, M. W. & Sutherland, J. D. 2011 Prebiotic chemistry: a new *modus operandi*. *Phil. Trans. R. Soc. B* **366**, 2870–2877. (doi:10.1098/rstb.2011.0134)
3 Blackmond, D. G. 2011 The origin of biological homochirality. *Phil. Trans. R. Soc. B* **366**, 2878–2884. (doi:10.1098/rstb.2011.0130)
4 Hanczyc, M. M. 2011 Metabolism and motility in prebiotic structures. *Phil. Trans. R. Soc. B* **366**, 2885–2893. (doi:10.1098/rstb.2011.0141)
5 Szostak, J. W. 2011 An optimal degree of physical and chemical heterogeneity for the origin of life? *Phil. Trans. R. Soc. B* **366**, 2894–2901. (doi:10.1098/rstb.2011.0140)
6 Grabowski, P. J., Zaug, A. J. & Cech, T. R. 1981 The intervening sequence of the ribosomal RNA precursor is converted to a circular RNA in isolated nuclei of *Tetrahymena*. *Cell* **23**, 467–476. (doi:10.1016/0092-8674(81)90142-2)
7 Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. 1983 The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35**, 849–857. (doi:10.1016/0092-8674(83)90117-4)
8 Brown, R. S., Hingerty, B. E., Dewan, J. C. & Klug, A. 1983 Pb(II)-catalysed cleavage of the sugar–phosphate backbone of yeast tRNAPhe—implications for lead toxicity and self-splicing RNA. *Nature* **303**, 543–546. (doi:10.1038/303543a0)
9 Yarus, M. 2010 *Life from an RNA world*, pp. 1–198. Cambridge, MA: Harvard University Press.
10 Altman, S. 2011 Ribonuclease P. *Phil. Trans. R. Soc. B* **366**, 2936–2941. (doi:10.1098/rstb.2011.0142)
11 Lilley, D. M. J. 2011 Mechanisms of RNA catalysis. *Phil. Trans. R. Soc. B* **366**, 2910–2917. (doi:10.1098/rstb.2011.0132)
12 Hiller, D. A. & Strobel, S. A. 2011 The chemical versatility of RNA. *Phil. Trans. R. Soc. B* **366**, 2929–2935. (doi:10.1098/rstb.2011.0143)
13 Ferré-D'Amaré, A. R. 2011 Use of a coenzyme by the *glmS* ribozyme-riboswitch suggests primordial expansion of RNA chemistry by small molecules. *Phil. Trans. R. Soc. B* **366**, 2942–2948. (doi:10.1098/rstb.2011.0131)
14 Seelig, B. & Jäschke, A. 1999 A small catalytic RNA motif with Diels-Alderase activity. *Chem. Biol.* **6**, 167–176. (doi:10.1016/S1074-5521(99)89008-5)
15 Sengle, G., Eisenfuh, R. A., Arora, P. S., Nowick, J. S. & Famulok, M. 2001 Novel RNA catalysts for the Michael reaction. *Chem. Biol.* **8**, 459–473. (doi:10.1016/S1074-5521(01)00026-6)
16 Johnston, W. K., Unrau, P. J., Lawrence, M. S., Glasner, M. E. & Bartel, D. P. 2001 RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. *Science* **292**, 1319–1325. (doi:10.1126/science.1060786)
17 Bagby, S. C., Bergman, N. H., Shechner, D. M., Yen, C. & Bartel, D. P. 2009 A class I ligase ribozyme with reduced Mg^{2+} dependence: selection, sequence analysis, and

identification of functional tertiary interactions. *RNA* **15**, 2129–2146. (doi:10.1261/rna.1912509)
18 Piccirilli, J. A. & Koldobskaya, Y. 2011 Crystal structure of an RNA polymerase ribozyme in complex with an antibody fragment. *Phil. Trans. R. Soc. B* **366**, 2918–2928. (doi:10.1098/rstb.2011.0144)
19 Yarus, M. 2011 The meaning of a minuscule ribozyme. *Phil. Trans. R. Soc. B* **366**, 2902–2909. (doi:10.1098/rstb.2011.0139)
20 Suga, H., Hayashi, G. & Terasaka, N. 2011 The RNA origin of transfer RNA aminoacylation and beyond. *Phil. Trans. R. Soc. B* **366**, 2959–2964. (doi:10.1098/rstb.2011.0137)
21 Pascal, R. & Boiteau, L. 2011 Energy flows, metabolism and translation. *Phil. Trans. R. Soc. B* **366**, 2949–2958. (doi:10.1098/rstb.2011.0135)
22 Schimmel, P. 2011 Mistranslation and its control by tRNA synthetases. *Phil. Trans. R. Soc. B* **366**, 2965–2971. (doi:10.1098/rstb.2011.0158)
23 Noller, H. F., Hoffarth, V. & Zimniak, L. 1992 Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* **256**, 1416–1419. (doi:10.1126/science.1604315)
24 Nissen, P., Hansen, J., Ban, N., Moore, P. B. & Steitz, T. A. 2000 The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**, 920–930. (doi:10.1126/science.289.5481.920)
25 Krupkin, M., Matzov, D., Tang, H., Metz, M., Kalaora, R., Belousoff, M. J., Zimmerman, E., Bashan, A. & Yonath, A. 2011 A vestige of a prebiotic bonding machine is functioning within the contemporary ribosome. *Phil. Trans. R. Soc. B* **366**, 2972–2978. (doi:10.1098/rstb.2011.0146)
26 Wohlgemuth, I., Pohl, C., Mittelstaet, J., Konevega, A. L. & Rodnina, M. V. 2011 Evolutionary optimization of speed and accuracy of decoding on the ribosome. *Phil. Trans. R. Soc. B* **366**, 2979–2986. (doi:10.1098/rstb.2011.0138)