NON-CODING RNA

The RNA World: molecular cooperation at the origins of life

Paul G. Higgs[1] and Niles Lehman[2]

Abstract | The RNA World concept posits that there was a period of time in primitive Earth's history — about 4 billion years ago — when the primary living substance was RNA or something chemically similar. In the past 50 years, this idea has gone from speculation to a prevailing idea. In this Review, we summarize the key logic behind the RNA World and describe some of the most important recent advances that have been made to support and expand this logic. We also discuss the ways in which molecular cooperation involving RNAs would facilitate the emergence and early evolution of life. The immediate future of RNA World research should be a very dynamic one.

RNA riboswitches
RNA molecules that respond to environmental conditions by changing secondary structure — and, in some cases, by modulating catalytic function — thereby affecting gene expression.

Ligases
Enzymes that covalently join polymers using ATP-derived energy.

Cooperation
The phenomenon whereby two or more entities interact to provide benefits for themselves that are greater than those possible by the operations of the entities in isolation.

[1]*Origins Institute and Department of Physics and Astronomy, McMaster University, Hamilton, Ontario L8S 4M1, Canada.*
[2]*Department of Chemistry, Portland State University, PO Box 751, Portland, Oregon 97207, USA.*
Correspondence to N.L.
e-mail: niles@pdx.edu
doi:10.1038/nrg3841
Published online
11 November 2014

The RNA World is the conceptual idea that there was a period in the early history of life on Earth when RNA, or something chemically very similar, carried out most of the information processing and metabolic transformations needed for biology to emerge from chemistry. This scenario, if it indeed existed, took place some 4 billion years ago. By contrast, the realization that RNA is a good candidate for the emergence of life is an idea that is only ~50 years old. It was recognized early on by Crick[1], Orgel[2] and others that RNA has both a genotype and a phenotype, and that a system based on RNA would be a plausible precursor to the much more complex system of DNA–RNA–proteins on which current life is based. It was also realized that the ribonucleotide coenzymes now used by many proteins may be molecular 'fossils' from an RNA-based metabolism[3]. Discoveries of naturally occurring ribozyme catalysts, such as self-splicing introns[4] and the ribonuclease P catalyst[5], were made in the 1980s and, with the demonstration that ribosomal RNA catalyses peptide bond formation in the ribosome[6], the credentials of RNA as a catalyst became firmly established.

Spiegelman's classic experiments with the bacteriophage Qβ showed how viral RNA could evolve over time in response to selection[7]. This study gave rise to the field of *in vitro* evolution[8] by demonstrating that RNA could behave in a Darwinian manner in the absence of cells. Once this realization has been made, pioneers such as Orgel[9], Eigen[10], Joyce[11,12], Gold[13] and Szostak[14] fully demonstrated the evolutionary capabilities of RNA and made it difficult to ignore the possibility that life started with RNA. However, additional pieces of data from both new and old angles then became available. The catalytic repertoire of RNA was shown to be diverse[15], RNA riboswitches were detected in bacteria and shown to be widespread in biology[16], and an autocatalytic cycle based on RNA ligases was found[17]. Furthermore, polymerase ribozymes that can use another sequence as a template were selected and improved[18–20] (see below).

Different research questions are gradually being brought together to assemble a complete picture of the emergence of life via the RNA World scenario (FIG. 1). New routes of chemical synthesis of ribonucleotides that could operate prebiotically are being studied, and there are also further experiments to isolate ribozymes from random RNA sequences. These areas are reviewed below. Theoretical models of RNA action are being developed to describe the origin and evolution of replicating systems. A point that emerges both from theoretical work and from laboratory experiments is that cooperation at the molecular level is essential for the survival of replicating sequences. A key aim of this Review is to describe the different senses in which cooperation is relevant in the RNA World. We argue that RNA replication must also fit into a broader thermodynamic and biological context if this were to form the basis of life (FIG. 1). What was the energy source that drove the synthesis of large macromolecules on early Earth? What were the environmental conditions at the location where these molecules were forming? How did RNA replication become associated with growth and division of protocells? Current ideas on some of these questions are considered later in this Review.

How did prebiotic synthesis of RNA occur?

The first, and in some ways the most important, problem facing the RNA World is the difficulty of prebiotic

Protocells
Membrane-enclosed compartments that may not contain all the components of present-day cells but that were presumably capable of some rudimentary means of growth and division. They can also refer to artificial cell-like structures created in the laboratory.

Activated
In this context, pertaining to nucleotides that are primed with a high-energy bond to facilitate their condensation with other nucleotides.

synthesis of RNA. This point has been made forcefully by Shapiro[21,22] and has remained a focal point of the efforts of prebiotic chemists for decades[23]. The 'traditional' thinking was that if one could assemble a ribose sugar, a nucleobase and a phosphate, then a nucleotide could arise through the creation of a glycosidic bond and a phosphoester bond. If nucleotides were then chemically activated in some form, then they could polymerize into an RNA chain. Each of these synthetic events poses tremendous hurdles for the prebiotic Earth, not to mention the often-invoked critique of the inherent instability of RNA in an aqueous solution[24–26]. Thus, the issue arises of whether there could have been a single environment in which all these steps took place. Benner has eloquently noted that single-pot reactions of sufficient complexity lead to 'asphaltization' (basically, the production of intractable 'goo') and argued for a discontinuous synthesis model[27]. In this model, different environments could generate distinct materials, and their compilation into RNA would have occurred sequentially as, for example, a stream percolates down a mountain into a pond. Along the way, various key reactions could have benefited from local environments. For example, the appearance of ribose as the main sugar for nucleic acids is enhanced in the presence of stabilizing borates[28] or silicates[29].

Nevertheless, efforts to simulate something close to a one-pot route to nucleotide synthesis continue and have proved to be successful. In 2009, Sutherland and colleagues realized the possibility that there are many alternative routes to nucleotide synthesis besides the sugar–nucleobases–phosphate scheme[30] by showing that a few prebiotically plausible molecules could be coaxed into an activated nucleotide under fairly mild conditions[31]. A key feature of this synthesis is the multiple roles that a single compound — in this case, inorganic phosphate — has in the overall reaction scheme. Phosphate acts as a catalyst, a pH buffer and a reactant in the route towards β-ribocytidine-2′,3′-cyclic phosphate. This is effectively a 'three-fer' in that three-benefits-in-one can be extracted from phosphate. As the prebiotic synthesis of RNA is indeed a recalcitrant problem, we should be looking for more situations in which multiple benefits can come from single facets. Moreover, although the prebiotic availability of inorganic phosphate was once a crucial missing link in itself, recent work by Pasek has shown that various phosphorus species, including inorganic phosphate and pyrophosphate, can result from the interaction of iron-rich meteorites with water[32]. Thus, some of the problems are now being turned into solutions. Along these lines, efforts of the Sutherland group and those of his former students are showing that other particulars of the abiotic Earth — such as reduced metal ions and high ultraviolet light flux[33], and/or the likelihood of thioacetates resulting from iron sulphides[34] (as previously suggested by Wächtershäuser[35]) — can actually be wielded to the advantage of making activated nucleotides.

Importantly, creative new ways to assemble monomers into polymers are coming to light. In the past, one of the most successful means of doing this was to use

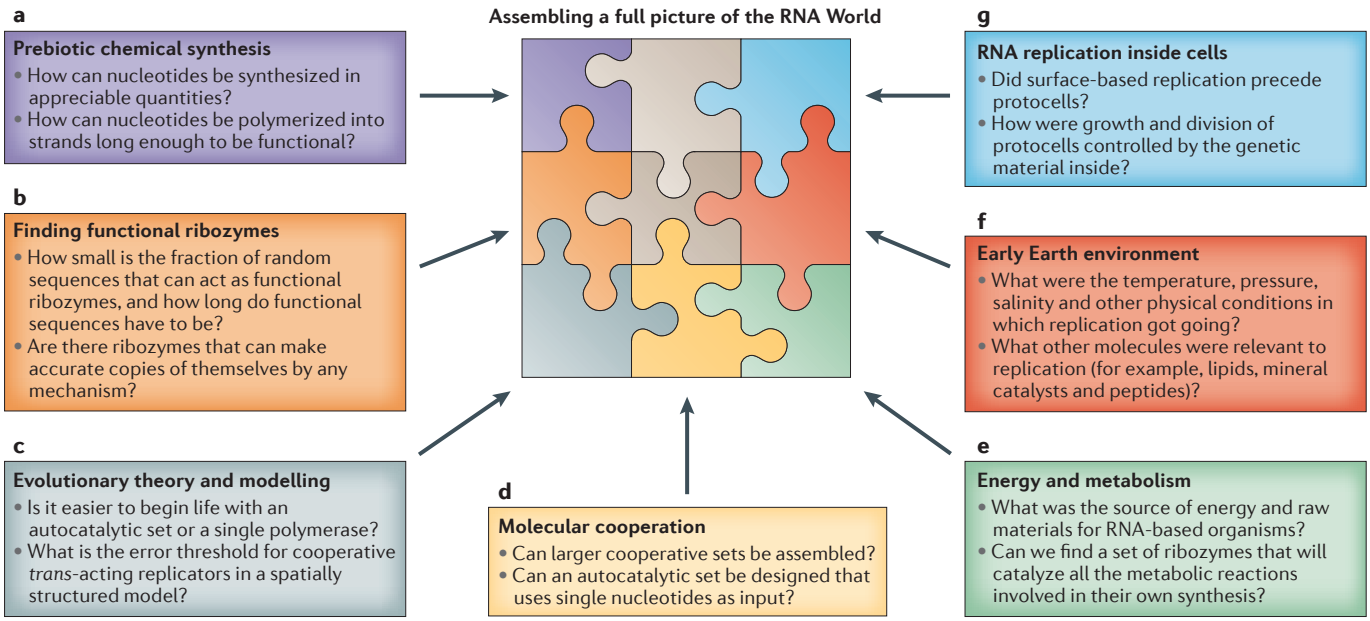


Figure 1 | **Research in different fields is coming together to assemble a more complete picture of the way the RNA World began and operated.** **a** | Progress in organic chemistry helps to show how nucleotides and RNA oligomers could have been synthesized before life. **b** | *In vitro* evolution studies discover functional ribozymes in the very large RNA sequence space. **c** | Theoretical models emphasize the importance of stochasticity and spatial structure for the evolution of replicators. **d** | Both experiment and theory point to the different ways in which molecular cooperation is essential. **e** | The way that a free energy source can drive RNA synthesis and replication is still poorly understood. **f** | Synthesis of ribozymes that operate under different conditions gives some clues as to what the early Earth environment might have been like. **g** | Experiments on linking RNA replication with vesicle growth and division are moving towards the creation of artificial cells.

Box 1 | **What alternatives to RNA could have existed?**

The idea that alternative nucleic acid chemistries could have formed more easily than RNA has a long history[26,36]. Eschenmoser provided a systematic survey of the possible alternatives for ribofuranosyl nucleic acids[101] and concluded that far more research on the synthetic routes available to RNA and RNA-like structures is needed before a consensus of the antiquity of RNA per se could be reached.

This avenue of investigation continues robustly today. One approach is to query the chemistry and evolutionary potential of alternative, or additional, nucleobases on a standard RNA backbone. The biotic world is replete with modified nucleosides, many of which display non-canonical base-pairing properties. Consequently, it is of interest to see whether any novel (that is, abiotic) nucleobases can provide information about the earliest chemistry of life. Various 'third base pairs' have been demonstrated[102,103]. In the past few years, attention has turned to the role of metal ions in mediating still more intriguing possibilities. An interesting example is a copper(II)-mediated base pair, purine-2,6-dicarboxylate·pyridine, which is not only metal-dependent but also recognizable by many contemporary DNA polymerases[104]. This result is exciting because copper ions could in fact turn out to be a 'two-fer', having a role not only in nucleotide synthesis but also in nucleobase pairing.

An analogous and parallel approach is to investigate whether an alternative polymeric backbone can exhibit desirable chemical and evolutionary qualities. If simpler alternatives to RNA, such as threose nucleic acid (TNA) or glycerol nucleic acid (GNA), are able to form a stable (but not too stable) duplex with RNA, then it could be envisioned that RNA could take over genetically from its predecessor as a result of selection for chemical versatility. It has been shown that a duplex can occur between TNA and TNA[105], between GNA and RNA, and between TNA and RNA[106], but not between GNA and TNA.

The Chaput and Holliger groups have recently shown that not only can TNA fold into functional three-dimensional structures, but iterative selection pressure can also differentiate some sequences from others. By creating libraries of TNA using a mutant RNA polymerase that accepts TNA monomers, it became possible to select for TNAs that were good binders to the human thrombin protein[107], to the HIV *trans*-acting response element and to lysozyme[108]. Another example is DNA itself, and there has been considerable progress in developing catalytic DNA-zymes[109,110] by *in vitro* selection. However, DNA is not usually thought of as a candidate for the first replicating polymer. It is generally believed that DNA arose after both RNAs and proteins because the usual biochemical synthesis of deoxyribose is from ribose, and because several non-homologous forms of DNA polymerases are found across the domains of life, whereas the ribosome itself is homologous in all domains[111,112].

Studies on various forms of nucleic acid-like molecules (known generally as XNA) do not unambiguously demonstrate that an alternative polymer preceded RNA, and they leave open the most parsimonious alternative that replication began with RNA itself. It remains to be seen whether the abiotic synthesis of XNA solves some of the problems of RNA and whether a probable route from XNA to RNA can be delineated. The essential aspect of nucleic acids is the complementary pairing between strands that enables heredity and evolution. The fact that this is far less possible with proteins is one of the reasons that points towards an RNA World scenario for the emergence of life. It also seems important to us that, after at least 3.5 billion years of evolution of proteins since the origin of translation and the genetic code, no organism has managed to rid itself of its nucleic acid ancestry and evolve a protein-only replication system. This suggests that the lack of a mechanism for protein replication equivalent to complementary pairing is a fundamental limitation of proteins.

some sort of template — either a pre-existing polymer strand[36] or a solid substrate such as clay[37] — to orient and guide monomers into position next to one another in order to lower the activation energy barrier needed to form inter-monomer (for example, phosphodiester) bonds. Alternatively (or additionally), the eutectic phase at the water–ice interface can help to drive otherwise unfavourable condensation reactions[38,39] (see below). However, Cafferty *et al.*[40] realized recently that the hydrophobic effect that drives base stacking is an often under-appreciated force that stabilizes nucleic acids and holds two strands together, and showed that this can drive the non-covalent polymerization of planar structures into very long polymers. This general strategy has since been extended to nucleotide pairs per se, where this effect augments the covalent bond formation among adjacent members in a stack[41].

Eutectic phase
A chemical mixture that has a lower freezing point than a composition of pure ingredients.

Although the evidence that RNA preceded proteins and DNA is relatively strong, it is less certain that RNA was the original replicating polymer, and it is possible that RNA is a step in a longer chain leading from the origin of life to modern biochemistry[42]. In BOX 1, we discuss progress with study of alternative nucleic acids that might have existed before RNA.

How could RNA have replicated itself?

The central principle of the RNA World is that RNA could have made a copy of itself, thereby 'setting the stage' for evolution and hence biology. This is supported by studies that have investigated RNA replication in the laboratory using three different approaches. The first is non-enzymatic template-directed polymerization, in which an RNA strand promotes the formation of its complementary strand without the presence of either a ribozyme or a protein enzyme. The second is RNA polymerization by a general RNA polymerase ribozyme (or replicase) that catalyses the copying of a template strand to make a complementary sequence of the template, and the third is the development of networks of molecules that catalyse the formation of other specific sequences in the network such that the network as a whole is mutually autocatalytic.

Non-enzymatic template-directed polymerization. This approach was initiated by Orgel's group[43,44], who showed that a pre-existing nucleic acid strand could promote the synthesis of a complementary strand using activated mononucleotides. Over the years, improvements have been made to this process, but problems due to product inhibition and nucleotide bias have persisted. In 2011, Deck *et al.*[45] used four strategies concurrently to show that oligonucleotides could be polymerized with high efficiency in a few days. One of the tools used in this study was cold temperatures (0°C). This turned out to be a 'two-fer', as the cold served not only to aid the binding of successive nucleotides to the template but also to slow the rate of hydrolysis of the growing RNA strand[45,46].

It is important to note that cold scenarios are gaining ground in the RNA community. Although the idea of a cold start to life has been around for some time[47,48], it has conventionally been thought that prebiotic chemistry had to have taken place in a rather thermally active environment. This may be because, in general, reactions (especially those with poor yields) tend to be augmented when more energy is present. Moreover, the idea that there was a hot start to life also traces its origins to the notion that the Hadean Earth was in fact very hot, and a mild example of this is the 'warm little pond' idea proposed by Darwin. The discovery of deep-sea hydrothermal vents and the extremophilic microorganisms that inhabit these localities also certainly provided further support to this notion. However, the utility of cold,

Class I ligase ribozyme
A ribozyme selected from a random pool of RNAs that can catalyse the ligation of an exogenous fragment of RNA to its own 5′ end.

Autocatalytic set
A collection of molecules that mutually cooperate in the sense that none of them can replicate without all the others, such that the reactions that form the components of the set are catalysed by other components of the set.

even sub-zero, temperatures in promoting various key reactions is coming to light. Regardless of the average temperature of the prebiotic Earth, there certainly would have been cold microenvironments, such as those at the poles or at depth in bodies of water.

RNA polymerization by a general RNA polymerase ribozyme. In this approach, a template is replicated as a result of the catalytic action of an RNA polymerase ribozyme. For many, such a polymerase stands as the 'holy grail' of the RNA World because, if it was able to act on a template of the same length as itself, then it could sustain an autocatalytic cycle by alternating copying of itself and its complementary sequence. An example of a polymerase of this type is R18, which is derived from the class I ligase ribozyme[18] (FIG. 2a). This approach has also benefited from the consideration of cold scenarios. The Holliger group had achieved notable success in this regard by engineering and selecting mutations in polymerase ribozymes that could catalyse the template-directed polymerization of an RNA chain of roughly half of its own length[49]. However, when they selected for further variants that are operable in the eutectic phase of a water–ice solution at −7 °C, they were able to achieve a substantial improvement in activity and showed, for the first time, that RNA could replicate strands of their own length (206 nucleotides in this case) or above[50]. This stands as the longest strand replicated by RNA-directed nucleotide-by-nucleotide polymerization, and the cold temperature again had an important role in that it helps to keep the ribozyme on task by promoting tighter binding to its substrates. Although the sequence that was replicated was not that of the ribozyme itself and this is therefore not yet a demonstration of self-replication, this study achieved a crucial proof-of-concept milestone[51]. It should also be noted that these template lengths are very much longer than those that result from non-enzymatic synthesis (see above), which have yet to generate products longer than 30–50 nucleotides[37,44].

Another key point is that the rate of ribozyme-catalysed polymerization should be much less sensitive to the sequence of the template than that of non-enzymatic synthesis. A polymerase ribozyme could also replicate different sequences that have other useful functions but that are not themselves polymerases. In this way, we envisage building up a system of many ribozymes that control a metabolic function, all of which are copied by the same polymerase. This would be comparable to the genetic system of modern organisms in which there is a single mechanism for DNA replication, RNA transcription or protein translation that operates on genes of different functions. A polymerase ribozyme is a cooperator in the sense that it copies another template sequence rather than directly copying itself (FIG. 2a,b). Replication of the polymerase thus requires a second copy of the polymerase. Cooperative polymerases are vulnerable to exploitation by parasites that act as templates but that do not contribute to the replication of the polymerase. Solutions to the problem of parasitic templates (FIG. 2c,d) are discussed below.

Mutually autocatalytic RNA networks. If a general RNA polymerase were operating in the RNA World, then it was presumably fairly long (for example, 200 nucleotides or more), and it is still not clear how likely it would be for the first such molecule to be created by prebiotic chemistry. For this reason, another approach to RNA replication focuses on the possible combined action of many shorter strands rather than on the search for a single general polymerase. Given that mononucleotides can in fact polymerize on various substrates such as clay[44,52] up to at least 20 nucleotides in length, one can envision scenarios in which these RNAs can interact and react to promote phosphoester bond breakage and formation. Kauffman[53] introduced the idea of an autocatalytic set (FIG. 3a) and proposed that all molecules in the set can be synthesized by reactions that are catalysed by other molecules in the set. The set as a whole is mutually autocatalytic, even though none of the molecules is individually autocatalytic. The chemical subtleties of

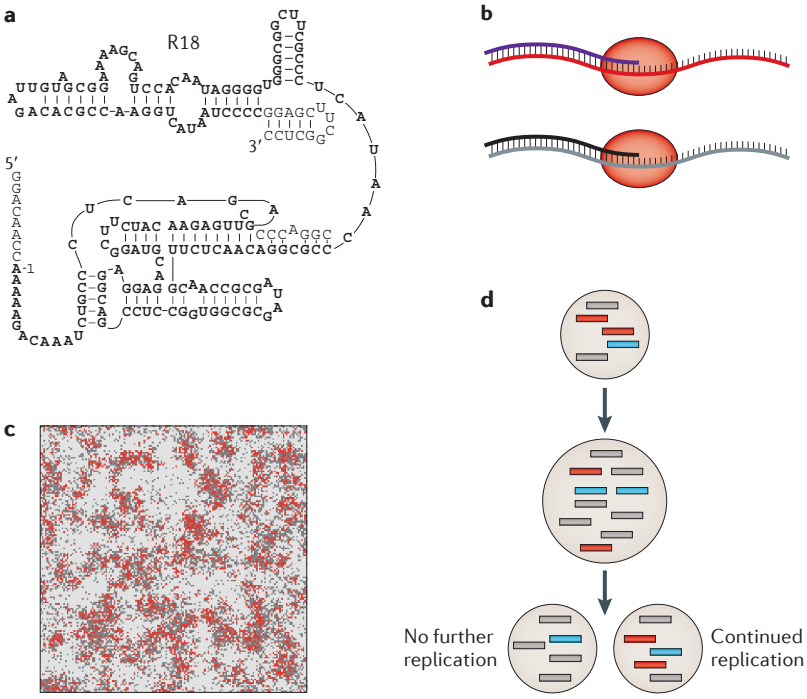


Figure 2 | **RNA polymerases as altruistic cooperators.** **a** | A *trans*-acting RNA polymerase (or replicase) such as the R18 polymerase[18] is a likely mechanism for supporting replication in the RNA World. **b** | Such a polymerase can use a template that is either another copy of itself (red) or an unrelated sequence (grey). Well-mixed systems of altruistic replicators are destroyed by parasites. There are two ways in which cooperators can resist parasites. **c** | First, survival of replicases is possible in two-dimensional models on surfaces. A simulation of altruistic replicators (red) diffusing on a surface in the presence of parasites (grey) demonstrates survival of the replicators as a result of spatial clustering[82]. **d** | Second, when small groups of molecules are packaged in compartments, group selection can occur. Functional RNAs (red and blue) are shown replicating in a protocell compartment in the presence of parasites (grey). Random segregation creates cells of varying composition. Cells with an excess of parasites are unviable, and this prevents parasites from over-running the system, even if the parasites multiply at a higher rate within a cell. Part **a** reproduced from REF. 50, Nature Publishing Group. Part **c** reprinted from *J. Theor. Biol.*, Vol. **364**, Shay, J. A., Huynh, C. & Higgs, P. G., The origin and spread of a cooperative replicase in a prebiotic chemical system, 249–259, Copyright (2014), with permission from Elsevier.

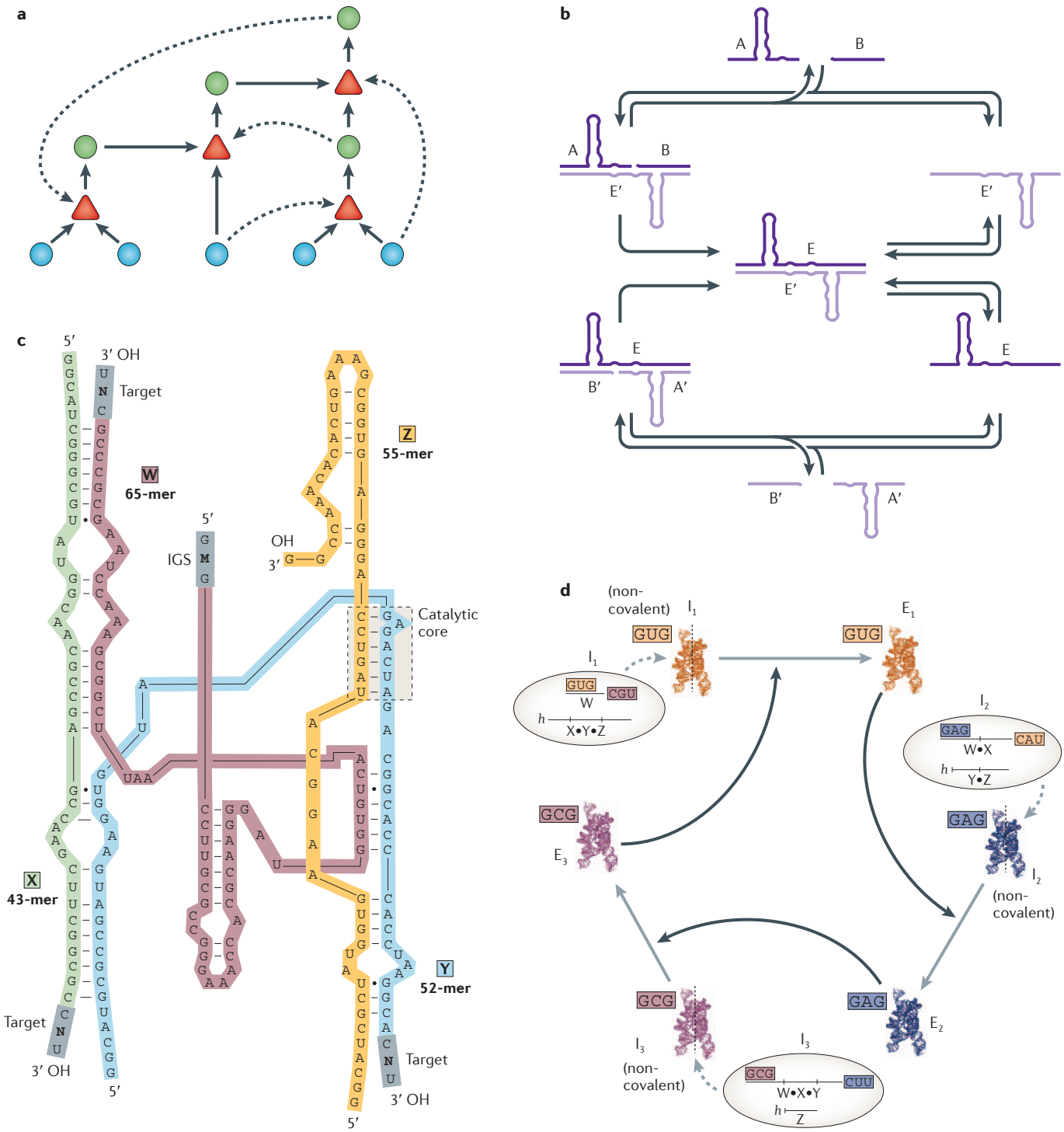


Figure 3 | **Different senses of molecular cooperation.** **a** | Cooperation in networks of mutually dependent strands with different functions is shown. An autocatalytic set involves precursor molecules present in the environment (blue circles) and molecules synthesized by the cycle (green circles). Each reaction (red triangles) is catalysed by a molecule that is part of the set (dashed arrows). **b** | An autocatalytic set composed of two cross-catalytic ligases[17] is shown. RNA A and RNA B are ligated together by ribozyme E′ to create ribozyme E, which can reciprocate and ligate RNA A′ and RNA B′ to create ribozyme E′. **c** | Cooperation between multiple strands that assemble to perform a single function is shown. Ribozymes, such as the *Azoarcus* recombinase[57], can be made from several short strands that assemble as a result of RNA secondary structure formation and information contained in internal guide sequences (IGSs) and complementary targets (grey). **d** | While strands can cooperate to form ribozymes, these ribozymes can then potentially cooperate at an even higher level to construct an autocatalytic set, such as a three-membered cycle[60]. In this scheme, the *Azoarcus* ribozyme is fragmented into two pieces in three different ways (that is, at three different junctions). The IGS of the ribozyme in E_1 is adjusted in such a way that it can only covalently assemble the E_2 ribozyme and so on, such that for assembly of all the ribozymes an obligatory cooperative process must (and does) occur. Part **b** from Lincoln, T. A. & Joyce, G. F. Self-sustained replication of an RNA enzyme. *Science* **323**, 1229–1232 (2009). Reprinted with permission from AAAS. Parts **c** and **d** adapted from REF. 60, Nature Publishing Group.

Box 2 | **Types of replication reactions and their relationship to cooperation**

We distinguish between four types of reaction systems that are relevant to RNA replication. In each reaction, catalyst molecules are written above the reaction arrow.

RNA virus replication

$$X \xrightarrow{E} X + X'; \quad X' \xrightarrow{E} X' + X. \qquad (1)$$

X is the virus genome, *X′* is its complementary sequence, and *E* is a protein enzyme that catalyses RNA replication. Experiments with the Qβ replicase[7,9,10] work in this way. In laboratory experiments, *E* is provided, whereas for naturally occurring viruses the RNA has to encode it. Given the presence of *E*, each different *X* sequence is an independent replicator; hence, this system does not require cooperation.

A general RNA polymerase ribozyme in the RNA World

$$X \xrightarrow{X} X + X'; \quad X' \xrightarrow{X} X' + X. \qquad (2)$$

X is the polymerase, *X′* is its complementary sequence (which is not a catalyst), and both can act as templates. Current laboratory studies of polymerases aim to create ribozymes of this type[49,50]. If *X* is a general polymerase, then it can also use any other RNA as a template (such as a non-functional parasitic RNA or a ribozyme with a different function). As *X* is an altruistic cooperator, this reaction system requires either spatial structure or compartmentalization to stabilize cooperation.

The hypercycle

$$X_1 \xrightarrow{X_2} 2X_1; \quad X_2 \xrightarrow{X_1} 2X_2. \qquad (3)$$

X_1 and X_2 are replicators that produce another copy of themselves with the aid of catalysis provided by the other sequence. In hypercycles with more than two components, each component catalyses the next one in a circular arrangement. The components of hypercycles are also altruistic cooperators, and either spatial structure or compartmentalization is therefore required to stabilize cooperation.

The autocatalytic set

$$A_1 + B_1 \xrightarrow{X_2} X_1; \quad A_2 + B_2 \xrightarrow{X_1} X_2. \qquad (4)$$

X_1 and X_2 are catalysts that catalyse one another's formation from precursors (*A*s and *B*s). This two-component example can be generalized to a larger number of components to form a network of the type shown in FIG. 3d. Ligase[17] and recombinase[60,61] systems that fit this reaction pattern have been studied in the laboratory. The components of an autocatalytic set are not altruistic cooperators, but they mutually cooperate in the sense that none of them can replicate without all of the others.

this process have been worked out in theory[54] and shown to be applicable to RNA systems[55], and this idea should hold even if the set components are not polymers[56].

The system of two mutually catalytic ligases[17] (FIG. 3b) is an elegant example of a rudimentary autocatalytic set constructed from RNA sequences. A second example is recombinase systems, in which RNA oligomers of ~50 nucleotides can assemble to form catalysts that can promote the synthesis of more of themselves[57,58] (FIG. 3c). These RNA sequences recombine to produce new sequences that can be improved and shortened through selection[59]. Moreover, higher-level cooperation among sets of RNAs (as opposed to single fragments) (FIG. 3d) could out-compete selfishly replicating systems when pitted together for common resources[60]. Furthermore, RNA fragments have the capacity, when they cooperate, to recycle themselves in and out of larger RNAs in a manner that is subject to natural selection[61].

It should be noted that the above three approaches to replication need not have been mutually exclusive. As noted previously[62], the evolution of replication probably proceeded in stages.

Altruism
When one individual provides a benefit to another while gaining no benefit itself (or even while suffering a detriment).

Hypercycles
Cooperative replicative sets of molecules in which hyperbolic growth is possible.

Error threshold
The theoretical maximum mutation rate that can sustain information genetic polymers of a particular length.

Cooperation in the RNA World

As discussed above, there are three senses in which cooperation is relevant in the RNA World: cooperation between polymerase and template strands (FIG. 2); cooperation between components in an autocatalytic set (FIG. 3a,b,d); and cooperation between RNA strands that assemble to form a catalyst (FIG. 3c). The evolution of cooperation has fascinated biologists because altruism should be selected against. Nevertheless, examples of cooperation are seen in biology at many levels, and several reasons have been found that cooperation and altruism can be selected over purely selfish behaviour[63–65].

We compare below the different kinds of cooperation that are relevant to the RNA World. To complement this discussion, BOX 2 emphasizes the kinds of chemical reaction systems that are associated with these cases. To begin, we note that if each sequence is independently able to be replicated, as in the case of RNA virus replication, then no cooperation is required for the survival of the replicators. Cooperation becomes relevant when an RNA polymerase ribozyme acts in *trans* (that is, without a covalent linkage) on a template sequence, as with the empirical RNA polymerases developed so far[18,19,49,50] and as envisaged for a general polymerase in the RNA World. A *trans*-acting polymerase is an altruistic cooperator because it replicates other sequences but is only replicated itself when another polymerase uses it as a template.

From the standpoint of evolutionary theory, there are many known examples of cooperators that are vulnerable to parasites in well-mixed systems (that is, when molecules react with one another at random in proportion to their concentration) but that can survive more easily in two-dimensional spatial models in which interactions occur between neighbouring molecules on a surface. This has been shown in many different types of model, including 'hypercycles' (REFS 66–68) — a simple model of a *trans*-acting RNA polymerase[69] — and models of metabolic replicators with several types of cooperating RNAs[70,71]. For the same reason, if a faster polymerase evolves in a well-mixed system, then it is only marginally evolutionarily stable because it benefits other competing sequences as much as its benefits itself[72]. In a spatially distributed system, improved polymerases can be positively selected because they benefit other copies of the same sequence that are close neighbours. A 'successful' polymerase would have to act sometimes as a catalyst and sometimes as a template, and these functions may be in opposition (for example, a well-folded tertiary structure is presumably necessary to be a catalyst, whereas an unfolded single strand might be a better template). Simulations also show that both template and polymerase abilities can co-evolve in spatially distributed models[73].

A general RNA polymerase ribozyme replicates by means of a two-component cycle that involves itself and its complement. This is different from a two-component hypercycle, in which each component is a specific catalyst for the replication of the other component (BOX 2). The hypercycle was originally proposed as a means of overcoming the problems of the error threshold[74,75]. Although

hypercycles have been widely studied theoretically[66–68,76], we are not aware of any experimental RNA system that corresponds to a hypercycle. There have also been theoretical studies of *trans*-acting polymerases[69,77–79], although most of these have ignored the alternation between complementary strands and assumed that a strand is a template for another copy of itself. Other models that consider strand alternation are more complex in that they include a folding algorithm for each sequence and a fitness that depends on the secondary structure[80], or they involve replication via a more complex strand displacement reaction[81]. We have recently considered the origin and spread of a cooperative polymerase that emerges as a single copy in a prebiotic system which supports the synthesis of random RNAs and non-enzymatic template-directed synthesis[82]. The newly evolved polymerase has to contend with parasitic templates that are already present owing to prebiotic chemistry, as well as with new parasites that are made by inaccurate replication of the polymerase. A simulation of this case (FIG. 2c) emphasizes that spatial clustering occurs in two-dimensional models, such that cooperators tend to be close to other cooperators that support them.

Another factor that permits altruistic replicators to survive is compartmentalization, as seen in a protocell in which several different sequences that replicate are present (FIG. 2d). Cell division and random segregation of sequences between daughter cells then occur. The daughter cell survives if it inherits all the relevant functional sequences but dies if it lacks functional sequences or has too many parasites. In a well-mixed system without compartments, rapidly multiplying parasites would destroy the system but, when compartments are present, group selection can operate and can sometimes overcome individual selection on single molecules. This was originally shown in the stochastic corrector model[83,84]. More recent models have permitted the comparison of compartmentalization and spatial clustering[76,85], showing that both mechanisms allow the survival of RNA replicators. It is still unclear how RNA replicators became associated with compartments and the extent to which spatially distributed replicating systems could evolve and become more complex in the absence of compartments. It is also possible that the first replicators were already inside protocell compartments (see below).

The second sense in which RNA sequences could have cooperated in the RNA World is by the formation of an autocatalytic set of mutually dependent strands[53–61] (see above). The components of an autocatalytic set are cooperators in the sense that they can only function when all the others are present but, unlike the general RNA polymerase and the components of a hypercycle, they are not displaying altruism. The essential difference is that the reactions in the autocatalytic set form a sequence from precursors (for example, by ligation or recombination reactions) and not by replicating an existing sequence (BOX 2). Autocatalytic sets can be stable in well-mixed reaction systems, and spatial structure is not necessary to prevent takeover by parasites, unlike for general RNA polymerase ribozymes. The key aspect of this second type of cooperation is that each of the molecules has a different function that contributes to the replication of the system as a whole. This also applies to the metabolic replicator model[86], in which several types of catalysts cooperate to synthesize the monomers required for RNA synthesis.

The third sense of cooperation is between strands that assemble into a ribozyme which has a single function[57,58,87] (FIG. 3c). The shorter strands are cooperators because they can only function when all of them come together. As each short strand on its own has no function, the strands have no option but to cooperate. There is thus no altruism involved. This kind of cooperation is potentially important for the emergence of life and may have temporally preceded the other types of cooperation discussed above. There are several reasons for this. First, the abundance of strands formed by random chemistry typically decays exponentially as the sequence length increases[45,88,89], and the formation of even just a few long functional sequences may be extremely unlikely. Thus, any mechanism that allows shorter sequences to be functional will make the emergence of life more likely. Second, polymerases are limited by the error rate. The standard error threshold theory[75] states that the maximum rate of error per nucleotide at which a replicator can survive is inversely proportional to the length of the sequence. It is therefore plausible that a ribozyme composed of shorter components could survive at a much higher per-base error rate than if it were a single long strand. Third, an important limitation of polymerase ribozymes so far developed in laboratories is that they are not completely processive along the template strand. Even the most processive developed so far[50] has a maximum template length of ~200 nucleotides. It is possible that a non-processive polymerase composed of several shorter components might be able to replicate each of the component strands, whereas it would be unable to replicate a single strand of the same total length. In this case it would be conceivable for a crude ribozyme function that solely consists of ligation or recombination chemistry to evolve before the processivity function, which would greatly simplify the problem of the origin of the polymerase.

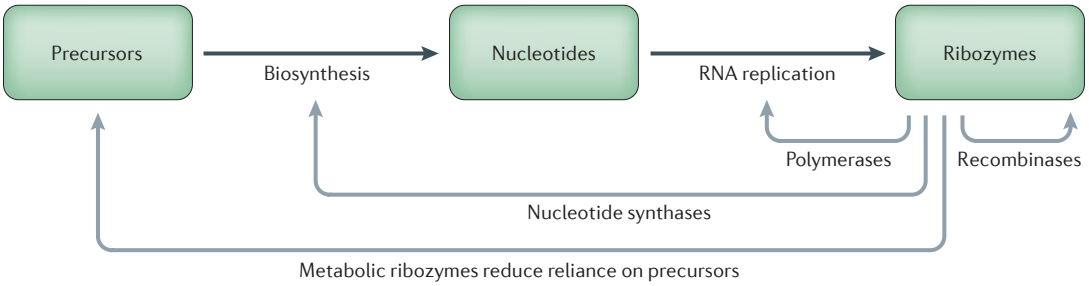
At present, it is difficult to know whether life is more likely to originate via a single polymerase or via an autocatalytic set. Which of these seems simpler depends on what one sees as being simple or complex. A network is complex in that it has many components, but a single polymerase ribozyme might be considered complex because its function is difficult to find in a mixture of random RNAs. If the reactions occurring in the network are relatively simple ligation and recombination reactions, then replication of a network of interacting molecules might actually be easier to achieve than the chemistry needed for general replication by a single polymerase, in which case replication might be a later biotic invention[62]. Although the ligase and recombinase laboratory systems have established that RNA sequences can maintain autocatalytic sets, it is less clear whether such a system could emerge from scratch from prebiotic chemistry. In the experimental cases, the precursor strands are too long to be created in appreciable concentrations by random polymerization. For an autocatalytic set to be viable at the time of the origin of life, it would have

Group selection
Selection that acts on a group of entities as a whole (such as animals living in a social group or molecules inside a protocell) and that favours survival of the whole group, in contrast to selection acting on individual members of a group that leads to competition between the individuals.

Box 3 | **The transition to the living state**

a Chemistry
The prebiotic world: a dead state

b Biology
The RNA World: an autocatalytic living state



Recent theoretical work[78,79,82,113–117] shows the difference between non-living chemistry (also known as 'pre-life') and a living state, such as the RNA World. In the prebiotic world, chemistry can generate nucleotides and random RNA sequences to some extent. Very rarely, a functional ribozyme might be generated among these random sequences, but the concentration would be very low. This state is 'dead' because it is controlled by random chemistry and has no heredity or evolution (see the figure). By contrast, in the RNA World, longer RNAs, including functional ribozymes, are present at substantial concentrations. This state is living because reactions are catalysed by ribozymes that are part of the system, because there is heredity of specific functional sequences and because evolution is possible by mutations or by adding sequences with new functions (see the figure).

We have studied several theoretical models for RNA polymerization[118,119] and simpler replicator models[78,79,82], and shown that the chemical reaction equations have both non-living and living solutions. For life to emerge, the system must be able to progress from the non-living state to the living state. In an infinite, well-mixed chemical system, the non-living state is stable forever, and no transition occurs. For finite numbers of molecules in a finite volume, the transition to life is possible because of concentration fluctuations[116]. However, the transition to life occurs much more easily in two-dimensional spatial models because concentration fluctuations are local. This allows the living state to originate in a small patch that can then spread across the surface[78,82]. The same conditions of local two-dimensional interactions and limited diffusion rate are important both for the stochastic transition that creates life and for the survival of cooperating replicators in the presence of parasites.

The key to maintenance of autocatalysis in the living state is a positive feedback mechanism such that the higher the concentration of ribozymes, the higher the rate of synthesis of new ribozymes. The figure shows that several different types of ribozymes can create this feedback. It is clear that polymerases that speed up RNA synthesis can cause feedback. It is also intuitive that nucleotide synthase ribozymes increase the concentration of nucleotides, which increases the rate of RNA synthesis and hence promote feedback. It is less intuitive that recombinase ribozymes can cause feedback; however, we have shown that if there is a mechanism to bias recombination in one direction over the other, then recombinases can also cause feedback that stabilizes the living state[117]. Ma *et al.*[118–120] have also shown in simulations that several different kinds of ribozymes could sustain a replicating system.

to use as input only monomers or short oligomers that we may assume to have been present in reasonable concentrations before life. An important advantage of the polymerase point of view is that it operates with only monomers as input.

In this Review, we adopt the view that if a replicating RNA system were present in the RNA World, then this would count as 'life', as it would satisfy the frequently used definition that life is a self-sustaining chemical system that is capable of evolution. In BOX 3, we review recent progress in understanding how a living, replicating system could have emerged from a non-living prebiotic chemical system.

When did cells arise?

We now return to the question of whether cell-like compartments were the driver of life and existed from the beginning, or whether they were a product of life that enhanced its metabolic and evolutionary potential. There are obvious reasons that life would benefit from cells, including (but not limited to) the advantage of concentrating reagents, the energetic potential of forming a concentration gradient between the inside and the outside of a cell, and the chance to link the phenotype of an entity spatially with the genotype that encoded it[90,91]. The existence of compartments also permits group selection to operate, which is one of the

important mechanisms for stabilizing the evolution of cooperation (see above).

Many efforts have succeeded in creating various types of protocells, many of which were based on lipids, others on protein components, and still others on water-in-oil emulsions. These are too numerous to list here and are reviewed elsewhere[92]. Some recent work has augmented these studies while providing new insights that are relevant to cooperation and conflict in the RNA World, and it has already been shown that protocells can compete with each other in a manner that elicits selection[93,94]. First, RNA-directed catalysis can actually be improved upon compartmentalization[95]. The mechanism for this seems to be molecular crowding, such that the local concentration of RNA can be increased, leading to a catalytic enhancement of up to 70-fold. Thus, the mere process of putting RNA into a tiny cell-like structure can confer a selective advantage on a genotype. Second, protocells can locally affect the distribution of substances such as Mg^{2+} ions, which are required in high concentrations to allow either recombinase ribozymes[57] or polymerase ribozymes[50] to function efficiently. The Szostak group has now demonstrated that the co-encapsulation of citrate ions can allow such high concentrations of Mg^{2+} ions to accumulate in lipid-based protocells, which enables non-enzymatic template-directed RNA primer extension to take place inside[96]. Of note is that citrate ions also seems to protect the nascent RNA from spontaneous hydrolytic degradation and is thus at least a 'two-fer' in this context. Finally, some exciting work from the Yomo group shows how compartmentalization can influence the evolutionary trajectory of its components. Bansho *et al.* used water-in-oil droplets to encapsulate a cell-free transcription–translation system that replicated the Qβ RNA genome and found that the smaller the cell, the more protected the genome was from the encroachment of short parasitic RNA species that have a tendency to hijack *in vitro* evolution experiments[97]. This result fits well with hypotheses on the origins of life because anything simpler, including smaller, would have been more prebiotically accessible[98]; it also hints that the RNA World may have at least overlapped with a cellular world at some point. Later, this system was shown to demonstrate 'Darwinian-type' evolution by the accumulation of point mutations over time[99], further supporting the notion that RNA may have benefited from being inside cells. However, there are remaining issues as to whether RNA led a cell-free life before compartmentalization and/or whether RNA compartments that are alternative to cell-like structures (for example, rock fissures) preceded cells.

Conclusions and perspectives

Contemporary studies both in the laboratory and by simulation are beginning to reveal the cooperative nature of the RNA World, as well as how various types of cooperation and conflict probably guided the earliest evolutionary processes. The basics of the RNA World concept are well established, but the details continue to evolve. A powerful theme has begun to emerge from many of these new approaches and their results. One way to describe this is systems chemistry[100]. By focusing not on single reactions in isolation but on the collective set of processes that must all occur contemporaneously, systems chemistry can lead to the discovery of 'multifers' in which the system as a whole can establish itself through the shared use of common conditions, reactants or products. Another way to look at this process, one that we highlighted in this Review, is that of chemical cooperation (BOX 1). This newer view, embedded in the broader perspective of intermolecular cooperation and conflict, is pervading other aspects of study of the RNA World, as we explored above.

Key questions for the next decade of research include the following. How long were the first ribozymes? In other words, what is the extent of the error threshold problem? How specific were the first ribozymes? Was life sparked by a single polymerase or a random autocatalytic set? How far can the RNA World go without being encapsulated in a cell, or were there cells already at the earliest stage? What was the energy source for the RNA World? Thus, how was it possible to link a stable energy supply to a metabolic synthesis of RNA? The RNA World idea emphasizes replication, but thermodynamic driving force is still needed for synthesis. We anticipate that the answers to many of these questions will not only be within our conceptual reach in the next decade or two, but will also invoke the insights gained from a more systematic appreciation of how conflict and cooperation can influence molecular processes.

1. Crick, F. H. C. The origin of the genetic code. *J. Mol. Biol.* **38**, 367–379 (1968).
2. Orgel, L. E. Evolution of the genetic apparatus. *J. Mol. Biol.* **38**, 381–393 (1968).
3. White, H. B. Coenzymes as fossils of an earlier metabolic state. *J. Mol. Evol.* **7**, 101–104 (1976).
4. Kruger, K. *et al.* Self-splicing RNA: autoexcisionand autocyclization of the ribosomal RNA interveningsequence of Tetrahymena. *Cell* **31**, 147–157 (1982).
5. Guerrier-Takada, C. *et al.* The RNA moiety of ribonuclease P is thecatalytic subunit of the enzyme. *Cell* **35**, 849–857 (1983).
6. Nissan, P., Hansen, J., Ban, N., Moore, P. B. & Steitz, T. A. The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**, 920–930 (2000).
7. Mills, D. R., Peterson, R. L. & Spiegelman, S. An extracellular darwinian evolution experiment with a self-duplicating nucleic acid molecule. *Proc. Natl Acad. Sci. USA* **58**, 217–224 (1967).
8. Joyce, G. F. Forty years of *in vitro* evolution. *Angew. Chem. Int. Ed.* **46**, 6420–6436 (2007).
This review covers most of the test-tube evolutionary studies since the classic experiments by Spiegleman.
9. Biebricher, C. K. & Orgel, L. E. An RNA that multiplies indefinitely with DNA-dependent RNA polymerase: selection from a random copolymer. *Proc. Natl Acad. Sci. USA* **70**, 934–938 (1973).
10. Biebricher, C. K., Eigen, M, & Gardiner, W. C. Kinetics of RNA replication: competition and selection among self-replicating RNA species. *Biochemistry* **24**, 6550–6560 (1985).
11. Robertson, D. L. & Joyce, G. F. Selection *in vitro* of an RNA enzyme that specifically cleaves singlestranded DNA. *Nature* **344**, 467–468 (1990).
12. Beaudry, A. A. & Joyce, G. F. Directed evolution of an RNA enzyme. *Science* **257**, 635–641 (1992).
13. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505–510 (1990).
14. Ellington, A. D. & Szostak, J. W. *In vitro* selection of RNA molecules that bind specific ligands. *Nature* **346**, 818–822 (1990).
15. Chen, X., Li, N. & Ellington, A. D. Ribozyme catalysis of metabolism in the RNA world. *Chem. Biodivers.* **4**, 633–655 (2007).
16. Mandal, R. & Breaker, R. R. Gene regulation byriboswitches. *Nature Rev. Mol. Cell. Biol.* **5**, 451–463 (2004).
17. Lincoln, T. A. & Joyce, G. F. Self-sustained replication of an RNA enzyme. *Science* **323**, 1229–1232 (2009).
18. Johnston, W. K., Unrau, P. J., Lawrence, M. S., Glasner, M. E. & Bartel, D. P. RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. *Science* **292**, 1319–1325 (2001).
This study is the first laboratory demonstration of RNA-catalysed RNA polymerization on a template (that is, the first hint that an RNA replicase might be possible).

19. Zaher, H. S. & Unrau, P. J. Selection of an improved RNA polymerase ribozyme with superior extension and fidelity. *RNA* **13**, 1017–1026 (2007).
20. Cheng, L. K. L. & Unrau, P. J. Closing the circle: replicating RNA with RNA. *Cold Spring Harb. Perspect. Biol.* **2**, a002204 (2010).
21. Shapiro, R. *Origins: A Skeptic's Guide to the Creation of Life on Earth* (Bantam New Age,1987).
22. Shapiro, R. A simpler origin of life. *Sci. Am.* **296**, 46–53 (2007).
23. Neveu, M., Kim, H. J. & Benner, S. A. The "strong" RNA world hypothesis: fifty years old. *Astrobiology* **13**, 391–403 (2013).
24. Joyce, G. F., Schwartz, A. W., Miller, S. L. & Orgel, L. E. The case for an ancestral genetic system involving simple analogues of the nucleotides. *Proc. Natl Acad. Sci. USA* **84**, 4398–4402 (1987).
25. Joyce, G. F. The rise and fall of the RNA world. *New Biol.* **3**, 399–407 (1991).
This often overlooked paper lists many of the pros and cons of the RNA World hypothesis before experimental data emerged in the 1990s.
26. Joyce, G. F. The antiquity of RNA-based evolution. *Nature* **418**, 214–221 (2002).
27. Benner, S. A., Kim, H.J. & Carrigan, M. A. Asphalt, water, and the prebiotic synthesis of ribose, ribonucleotides, and RNA. *Acc. Chem. Res.* **45**, 2025–2034 (2012).
This paper outlines the discontinuous synthesis model of RNA synthesis, underscoring the fact that RNA need not, and probably was not, made all at once in a single environment.
28. Ricardo, A., Carrigan, M. A., Olcott, A. & Benner, S. A. Borate minerals stabilize ribose. *Science* **303**,196 (2004).
29. Vázquez-Mayagoitia, A. *et al.* On the stabilization of ribose by silicate minerals. *Astrobiology* **11**, 115–121 (2011).
30. Anastasi, C. *et al.* RNA: prebiotic product, or biotic invention? *Chem. Biodivers.* **4**, 721–739 (2007).
31. Powner, M. W., Gerland, B. & Sutherland, J. D. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* **459**, 239–242 (2009).
This soon-to-be-classic paper shows how activated nucleotides can be made under abiotic conditions using the 'three-fer' of phosphate buffer.
32. Pasek, M. A., Harnmeijer, J. P., Buick, R., Gull, M. & Atlas, Z. Evidence for reactive reduced phosphorus species in the early Archaen ocean. *Proc. Natl Acad. Sci. USA* **110**, 10089–10094 (2013).
33. Ritson, D. & Sutherland, J. D. Prebiotic synthesis of simple sugars by photoredox systems chemistry. *Nature Chem.* **4**, 895–899 (2012).
34. Bowler, F. R. *et al.* Prebiotically plausible oligoribonucleotide ligation facilitated by chemoselective acetylation. *Nature Chem.* **5**, 383–389 (2013).
35. Wächtershäuser, G. Before enzymes and templates: theory of surface metabolism. *Microbiol. Rev.* **52**, 452–484 (1988).
36. Orgel, L. E. Did template-directed nucleation precede molecular replication? *Orig. Life Evol. Biosph.* **17**, 27–34 (1986).
37. Ferris, J. P. Mineral catalysis and prebiotic synthesis: montmorillonite-catalyzed formation of RNA. *Elements* **1**, 145–149 (2005).
38. Monnard, P.-A., Kanavarioti, A. & Deamer, D. W. Eutectic phase polymerization of activated ribonucleotide mixtures yields quasi-equimolar incorporation of purine and pyrimidine nucleobases. *J. Am. Chem. Soc.* **125**, 13734–13740 (2003).
39. Monnard, P.-A. & Szostak, J. W. Metal-ion catalyzed polymerization in the eutectic phase in water–ice: a possible approach to template-directed RNA polymerization *J. Inorg. Biochem.* **102**, 1104–1111 (2008).
40. Cafferty, B. J. *et al.* Efficient self-assembly in water of long noncovalent polymers by nucleobases analogs. *J. Am. Chem. Soc.* **135**, 2447–2450 (2013).
This paper reveals how there are far more RNA self-assembly mechanisms in an aqueous environment than previously appreciated.
41. Kuruvilla, E., Schuster, G. B. & Hud, N. V. Enhanced nonenzymatic ligation of homopurine miniduplexes: support for greater base stacking in a pre-RNA world. *ChemBioChem* **14**, 45–48 (2013).
42. Hud, N. V., Cafferty, B. J., Krishnamurthy, R. & Williams, L. D. The origin of RNA and "my grandfather's axe". *Chem. Biol.* **20**, 466–474 (2013).
43. Wu, T. & Orgel, L. E. Nonenzymatic template-directed synthesis on oligodeoxycytidylate sequences in hairpin oligonucleotides. *J. Am. Chem. Soc.* **114**, 317–322 (1992).
44. Hill, A. R. Jr., Wu, T. & Orgel, L. E. The limits of template-directed synthesis with nucleoside-5′-phosph oro(2-methyl)imidazolides. *Origins Life Evol. Biosph.* **23**, 285–290 (1993).
45. Deck, C., Jauker, M. & Richert, C. Efficient enzyme-free copying of all four nucleobases templated by immobilized RNA. *Nature Chem.* **3**, 603–608 (2011).
46. Lehman, N. & Hayden, E. J. Template-directed RNA polymerization: the taming of the milieu. *ChemBioChem* **12**, 2727–2728 (2011).
47. Moulton, V. *et al.* RNA folding argues against a hot-start origin of life. *J. Mol. Evol.* **51**, 416–421 (2000).
48. Vlassov, A. V., Johnston, B. H., Landweber, L. F. & Kazakov, S. A. Ligation activity of fragmented ribozymes in frozen solution: implications for the RNA world. *Nucleic Acids Res.* **25**, 2966–2974 (2004).
49. Wochner, A., Attwater, J., Coulson, A. & Holliger, P. Ribozyme-catalyzed transcription of an active ribozyme. *Science* **332**, 209–212 (2011).
50. Attwater, J., Wochner, A. & Holliger, P. In-ice evolution of RNA polymerase ribozyme activity. *Nature Chem.* **5**, 1011–1018 (2013).
This study demonstrated polymerization by an artificial RNA replicase beyond 200 nucleotides, which is the longest achieved so far.
51. Lehman, N. Cold-hearted RNA heats up life. *Nature Chem.* **5**, 987–989 (2011).
52. Huang, W. & Ferris, J. P. One-step, regioselective synthesis of up to 50-mers of RNA oligomers by montmorillonite catalysis. *J. Am. Chem. Soc.* **128**, 8914–8919 (2006).
53. Kauffman, S. A. *The Origins of Order: Self-Organization and Selection in Evolution* (Oxford Univ. Press, 1993).
54. Hordijk, W. & Steel, M. Detecting autocatalytic, self-containing sets in chemical reaction systems. *J. Theor. Biol.* **227**, 451–461 (2004).
55. Hordijk, W. & Steel, M. A formal model of autocatalytic sets emerging in a RNA replicator system. *J. Syst. Chem.***4**, 3 (2013).
56. Smith, J. I., Steel, M. & Hordijk, W. Autocatalytic sets in a partitioned chemical network. *J. Syst. Chem.* **5**, 2 (2014).
57. Hayden, E. J. & Lehman, N. Self-assembly of a group I intron from inactive oligonucleotide fragments. *Chem. Biol.* **13**, 909–918 (2006).
58. Hayden, E. J., von Kiedrowski, G. & Lehman, N. Systems chemistry on ribozyme self-construction: Evidence for anabolic autocatalysis in a recombination network. *Angew. Chem. Int. Ed.* **47**, 8424–8428 (2008).
59. Burton, A. S. & Lehman, N. Enhancing the prebiotic relevance of a set of covalently self-assembling, autorecombining RNAs through *in vitro* selection. *J. Mol. Evol.* **70**, 233–241 (2010).
60. Vaidya, N. *et al.* Spontaneous network formation among cooperative RNA replicators. *Nature* **491**, 72–77 (2012).
61. Vaidya, N., Walker, S. I. & Lehman, N. Recycling of informational units leads to selection of replicators in a prebiotic soup. *Chem. Biol.* **20**, 241–252 (2013).
62. Levy, M. & Ellington, A. D. The descent of polymerization. *Nature Struct. Biol.* **8**, 580–582 (2001).
63. Maynard Smith, J. & Szathmary, E. *The Major Transitions in Evolution* (Oxford University Press,1995).
64. Hammerstein, P. *Genetic and Cultural Evolution of Cooperation* (MIT Press, 2003).
65. Nowak, M. A. Five rules for the evolution of cooperation. *Science* **314**, 1560–1563 (2006).
This paper outlines the various modes of cooperation that could be operative in biological scenarios, setting the stage for considering chemical cooperation mechanisms.
66. Boerlijst, M. C. & Hogeweg, P. Spiral wave structure in prebiotic evolution: hypercycles stable against parasites. *Physica. D* **48**, 17–28 (1991).
67. Sardanyes, J. & Solé, R. V. Bifurcations and phase transitions in spatially extended two-member hypercycles. *J. Theor. Biol.* **243**, 468–482 (2006).
68. Sardanyes, J. & Solé, R. V. Spatiotemporal dynamics in simple asymmetric hypercycles under weak parasitic coupling. *Physica. D* **231**, 116–129 (2007).
69. McCaskill, J. S., Füchslin, R. M. & Altmeyer, S. The stochastic evolution of catalysts in spatially resolved molecular systems. *Biol. Chem.* **382**, 1343–1363 (2001).
70. Konnyu, B., Czaran, T. & Szathmáry, E. Prebiotic replicase evolution in a surface-bound metabolic system: parasites as a source of adaptive evolution. *BMC Evol. Biol.* **8**, 267 (2008).
71. Branciamore, S., Gallori, E., Szathmáry, E. & Czaran, T. The origin of life: chemical evolution of a metabolic system in a mineral honeycomb? *J. Mol. Evol.* **69**, 485–469 (2009).
72. Brogioli, D. Marginally stable chemical systems as precursors for life. *Phys. Rev. Lett.* **105**, 058102 (2010).
73. Szabo, P., Scheuring, I., Czaran, T. & Szathmáry, E. *In silico* simulations reveal that replicators with limited dispersal evolve towards higher efficiency and fidelity. *Nature* **420**, 340–343 (2002).
74. Eigen, M. & Schuster, P. The hypercycle: a principle of natural self-organization. Part B: the abstract hypercycle. *Naturwissenschaften* **65**, 7–41 (1978).
75. Eigen, M., McCaskill, J. & Schuster, P. Molecular quasi-species. *J. Phys. Chem.* **92**, 6881–6891 (1988).
76. Takeuchi, N. & Hogeweg, P. Evolutionary dynamics of RNA-like replicator systems: a bioinformatic approach to the origin of life. *Phys. Life Rev.* **9**, 219–263 (2012).
This paper brings together a wide range of theories and models of RNA evolution and clearly explains the issues from evolutionary theory.
77. Takeuchi, N. & Hogeweg, P. The role of complex formation and deleterious mutations for the stability of RNA-like replicator systems. *J. Mol. Evol.* **65**, 668–686 (2007).
78. Wu, M. & Higgs, P. G. The origin of life is a spatially localized stochastic transition. *Biol. Direct* **7**, 42 (2012).
79. Higgs, P. G. & Wu, M. The importance of stochastic transitions for the origin of life. *Orig. Life Evol. Biosph.* **42**, 453–457 (2012).
80. Takeuchi, N. & Hogeweg, P. Evolution of complexity in RNA-like replicator systems. *Biol. Direct* **3**,11 (2008).
81. Takeuchi, N. Salazar, L., Poole, A. M. & Hogeweg, P. The evolution of strand preference in simulated RNA replicators with strand displacement: implications for the origin of transcription. *Biol. Direct.* **3**, 33 (2008).
82. Shay, J. A., Huynh, C. & Higgs, P. G. The origin and spread of a cooperative replicase in a prebiotic chemical system. *J. Theor. Biol.* http://dx.doi.org/10.1016/j.jtbi.2014.09.019 (2014).
83. Szathmáry, E. & Demeter, L. Group selection of early replicators and the origin of life. *J. Theor. Biol.* **128**, 463–483 (1987).
This early classic paper sparked considerations of how spatial aggregation might have been important in the RNA World, and it led to the concepts of molecular cooperation outlined here.
84. Szathmáry, E., Sandos, M. & Fernando, C. Evolutionary potential and requirements for minimal protocells. *Top. Curr. Chem.* **259**, 167–211 (2005).
85. Takeuchi, N. & Hogeweg, P. Multilevel selection in models of prebiotic evolution II: a direct comparison of compartmentalization and spatial self-organization. *PLoS Comput. Biol.* **5**, e1000542 (2009).
86. Konnyu, B. & Czaran, T. The evolution of enzyme specificity in the metabolic replicator model of prebiotic evolution. *PLoS ONE* **6**, e20931 (2011).
87. Doudna, J. A., Couture, S. & Szostak, J. W. A multi-subunit ribozyme that is a catalyst of and template for complemtary strand RNA synthesis. *Science* **251**, 1605–1608 (1991).
88. Ferris, J. P. 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 (2002).
89. Lawless, J. G. & Yuen, G. U. Quantification of monocarboxylic acids in the Murchison carbonaceous meteorite. *Nature* **282**, 396–398 (1979).
90. Deamer, D. W. The first living systems: a bioenergetics perspective. *Microbiol. Mol. Biol. Rev.* **61**, 239–261 (1997).
91. Szostak, J. W., Bartel, D. P. & Luisi, P. L. Synthesizing life. *Nature* **409**, 387–390 (2001).
92. Solé, R. V., Munteanu, A., Rodriguez-Caso, C. & Macía, J. Synthetic protocell biology: from reproduction to computation. *Phil. Trans. Roy. Soc. B.* **362**, 1727–1739 (2007).
93. Chen, I. A, Roberts R. W, & Szostak J. W. The emergence of competition between model protocells. *Science* **305**, 1474–1476 (2004).
This paper nicely shows how chemical conflicts can be solved, or exacerbated, in cell-like structures.
94. Adamala, K. & Szostak, J. W. Competition between model protocells driven by an encapsulated catalyst. *Nature Chem.* **5**, 495–501 (2013).

95. Strulson, C. A., Molden, R. C., Keating, C. D. & Bevilacqua, P. C. RNA catalysis through compartmentalization. *Nature Chem.* **4**, 941–946 (2012).
96. Adamala, K. & Szostak, J. W. Nonenzymatic template-directed RNA synthesis inside model protocells. *Science* **342**, 1098–1100 (2013).
97. Bansho, Y. *et al.* Importance of parasite RNA species repression for prolonged translation-coupled RNA self-replication. *Chem. Biol.* **19**, 478–487 (2012).
98. Lehman, N. Evolution finds shelter in small spaces. *Chem. Biol.* **19**, 439–440 (2012).
99. Ichihashi, N. *et al.* Darwinian evolution in a translation-coupled RNA replication system within a cell-like compartment. *Nature Commun.* **4**, 2494 (2013).
100. von Kiedrowski, G., Otto, S. & Herdewijn, P. Welcome home, systems chemists! *J. Syst. Chem.* **1**, 1 (2010).
101. Eschenmoser, A. Chemical etiology of nucleic acid structure. *Science* **284**, 2118–2124 (1999).
102. Malyshev, D. A. *et al.* Solution structure, mechanism of replication, and optimization of an unnatural base pair. *Chem. Eur. J.* **16**, 12650–12659 (2010).
103. Yang, Z., Chen, F., Alvarado, J. B. & Benner, S. A. Amplification, mutation, and sequencing of a six-letter synthetic genetic system. *J. Am. Chem. Soc.* **133**, 15105–15112 (2011).
104. Kim, E.-K. & Switzer, C. Polymerase recognition of a Watson–Crick-like metal-mediated base pair: purine-2,6-dicarboxylate·copper(II)·pyridine. *ChemBioChem* **14**, 2403–2407 (2013).
105. Ebert, M. O., Mang, C., Krishnamurthy, R., Eschenmoser, A. & Jaun, B. The structure of a TNA–TNA complex in solution: NMR study of the octamer duplex derived from α-(L-threofuranosyl-)3′-2′-CGAATTCG. *J. Am. Chem. Soc.* **130**, 15105–15115 (2008).
106. Yang, Y.-W., Zhang, S., McCullum, E. O. & Chaput, J. C. Experimental evidence that GNA and TNA were not sequential polymers in the prebiotic evolution of RNA. *J. Mol. Evol.* **65**, 289–295 (2007).
107. Yu, H., Zhang, S. & Chaput, J. C. Darwinian evolution of an alternative genetic system provides support for TNA as an RNA progenitor. *Nature Chem.***10**, 183–187 (2012).
108. Pinheiro, V. B. *et al.* Synthetic genetic polymers capable of heredy and evolution. *Science* **336**, 341–344 (2012).
109. Schlosser, K. & Li, Y. Biologically inspired synthetic enzymes made from *DNA*. *Chem. Biol.* **16**, 315–322 (2009).
110. Silverman, S. K. DNA as a versatile chemical component for catalysis, encoding, and stereocontrol. *Angew. Chem. Int. Ed.* **49**, 7180–7201 (2010).
111. Lazcano, A., Guerrero, R., Margulis, L. & Oro, J. The evolutionary transition from RNA to DNA in early cells. *J. Mol. Evol.* **27**, 283–290 (1988).
112. Burton, A. S. & Lehman, N. DNA before proteins? Recent discoveries in nucleic acid catalysis strengthen the case. *Astrobiology.* **9**, 125–130 (2009).
113. Ohtsuki, H. & Nowak, M. A. Prelife catalysts and replicators. *Proc. Roy. Soc. B* **276**, 3783–3790 (2009).
114. Manapat, M. L., Chen, I. A. & Nowak, M. A. The basic reproductive ratio of life. *J. Theor. Biol.* **263**, 317–327 (2010).
115. Chen, I. A. & Nowak. M. A. From prelife to life: how chemical kinetics become evolutionary dynamics. *Acc. Chem. Res.* **45**, 2088–2096 (2012).
116. Wu, M. & Higgs, P. G. Origin of self-replicating biopolymers: autocatalytic feedback can jump-start the RNA world. *J. Mol. Evol.* **69**, 541–554 (2009).
117. Wu, M. & Higgs, P. Comparison of the roles of momoner synthesis, polymerization, and recombination in the origin of autocatalytic sets of biopolymers. *Astrobiology* **11**, 895–906 (2011).
118. Ma, W., Yu, C., Zhang, W. & Hu, J. A simple template-dependent ligase ribozyme as the RNA replicase emerging first in the RNA World. *Astrobiology* **10**, 437–447 (2010).
119. Ma, W., Yu, C., Zhang, W. & Hu, J. Nucleotide synthetase ribozymes may have emerged first in the RNA World. *RNA* **13**, 2012–2019 (2007).
120. Ma, W., Yu, C. & Zhang, W. Monte Carlo simulation of early molecular evolution in the RNA World. *Biosystems* **90**, 28–39 (2007).

Acknowledgements
The authors thank H. Lonsdale's Origin-of-Life Challenge for supporting the ideas and research mentioned in this work.

Competing interests statement
The authors declare no competing interests.