

Getting Past the RNA World: The Initial Darwinian Ancestor

Michael Yarus

Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, Colorado 80309-0347

Correspondence: yarus@colorado.edu

SUMMARY

A little-noted result of the confirmation of multiple premises of the RNA-world hypothesis is that we now know something about the dawn organisms that followed the origin of life, perhaps over 4 billion years ago. We are therefore in an improved position to reason about the biota just before RNA times, during the era of the first replicators, the first Darwinian creatures on Earth. An RNA congener still prominent in modern biology is a plausible descendent of these first replicators.

Outline

- | | |
|--|--|
| 1 Introduction | 10 An old variant backbone |
| 2 A new vantage | 11 An augmented alphabet, an augmented manifold of reactions |
| 3 Evidence for ancient RNA | 12 Chickens and eggs |
| 4 Conservation and persistence | 13 Continuity with an RNA world |
| 5 Persistence of CHNOPS | 14 Novel replication chemistry |
| 6 Persistence of RNA | 15 Catch-22 |
| 7 The hypothesis | References |
| 8 Earlier ideas about nucleotide cofactors | |
| 9 Pre-RNA world credentials | |

Editors: John F. Atkins, Raymond F. Gesteland, and Thomas R. Cech
Additional Perspectives on The RNA World available at www.cshperspectives.org

Copyright © 2010 Cold Spring Harbor Laboratory Press; all rights reserved; doi: 10.1101/cshperspect.a003590
Cite as *Cold Spring Harb Perspect Biol* 2010;1:a003590

1 INTRODUCTION

In the following discussion, we try to imagine the biological past over 4 Gya (Gigayears ago). Such speculation is worth new effort because of the success of the RNA-world hypothesis itself. The ultimate goal is a glimpse of the widely accepted, though hypothetical, initial replicator, whose biomolecular activity initiated Darwinian evolution on Earth. We are searching before the last universal common ancestor (LUCA), looking back toward the substantially older initial darwinian ancestor (IDA). That such a duplicating molecule existed (for example, Szathmari 2006) is therefore an indispensable assumption.

2 A NEW VANTAGE

Viewed from our time, it requires a huge leap to reach the IDA. The 4 Gy gap necessarily spanned is almost unthinkable, so great that it has carried most rock of the time away, consequently swallowing most hope of finding scientific evidence. But that is precisely the point; if 4 Gya was a time when oligoribonucleotides were prominent, we know something new, whose implications are usable as the crux of an inquiry into earlier times.

Furthermore, it requires only arithmetic to see that the RNA world can provide a decisive new vantage point. From the RNA world, the time to be spanned to the IDA is very much shortened. Because the solar system (and thus the Earth) congealed 4.5 Gya (Dalrymple 1991), in all likelihood an RNA world at 4 Gya is about 20-fold closer to the initiation of biological evolution than is the present. We want to know the Earth's biota a few hundred million years before the age of RNA. So, extrapolation from 4 Gya back to the IDA is 20-fold shorter than from today, probable molecular change is roughly 20-fold smaller, and the accuracy of any speculation roughly 20-fold greater than before acquisition of an RNA world standpoint. Instead of nearly completely spanning planetary history to reach the IDA, from the RNA world we span a time, for example, comparable to the history of Earth's mammals (Luo et al. 2001).

Moreover, the supporting data we need from the RNA world is minimal. Thus, the argument that follows does not use most of what we presently know with high probability. We need not assume that an organism existed that used RNA in every capacity for which it has been contemplated, or in even every capacity for which RNA's competence has been shown in selection experiments. Though RNA has surely proven its unanticipated versatility (Chen et al. 2007), we require only that a molecule made of something like modern ribonucleotides performed essential functions 4 Gya.

3 EVIDENCE FOR ANCIENT RNA

This indispensable contention rests on very robust experimental evidence, linking many independent observations. Such evidence is very frequent, for example, in studies of the apparent descent of translation—coded protein biosynthesis. If our immediate evolutionary predecessors were RNA-based organisms, then we plausibly require that they use RNA to invent translation and their successor catalysts, the peptides.

This idea in turn is strongly supported by selection of RNAs with the required translational capabilities. For example, the synthesis of aminoacyl-RNA (which activates amino acids and potentially links them to a matching coding triplet) is easily found within the RNA reaction repertoire, using amino acid adenylate (Illangasekare et al. 1995) (the universal biological substrate) or other carboxyl-activated amino acids (Lee et al. 2000) as precursor. Further, these RNA transacylators can be both faster and more accurate than modern proteins performing the same reaction (Illangasekare and Yarus 1999). In addition, aminoacyl group transfer is so simple a reaction for RNA that it can be performed by small RNAs, even in a ribozymic reaction center containing only three sparingly constrained nucleotides, or a ribozyme totaling five nucleotides in length (Turk et al. 2010). Such a catalyst would appear in an untemplated pool of ribonucleotide sequences consisting of only attograms (10^{-18} g) of RNA, unexpectedly easily reached by geochemical means. I argue, as previously (Yarus 2001; Yarus et al. 2005), that the unanticipated confirmation of translational activity, in fact, confirmation of the thorough competence of RNA in translational reactions, requires that the probability of an RNA invention of protein biosynthesis be strongly elevated. To do less than this is to ignore the theory of probabilities.

Alternatively, one could argue from the fact that the present peptidyl transferase is RNA, which exploits ribonucleotide properties of both peptidyl-tRNA (Weinger et al. 2004) and the rRNA (Nissen et al. 2000) to accelerate peptide transfer. The highly conserved rRNA peptidyl transferase sequence includes a ribonucleotide cradle whose sequence is particularly suited to poise the reacting amino acids (Welch et al. 1997), further supporting and illustrating its origin in the RNA world. Thus there is exceedingly strong support, all unexpected only a few years ago, for the origination of translation as a constellation of RNA activities. In fact, all chemical group transfers, and even the information transfers required for coded protein biosynthesis, have precedents within the chemical repertoire of pure small RNAs (Yarus 1991; Yarus et al. 2009). Any of the observations above, taken alone, is a strong argument for the existence of RNA-like molecules before the present

nucleoprotein era. Taken together, these multiple experiments are very persuasive indeed.

4 CONSERVATION AND PERSISTENCE

So, having staked out a position for RNA 4 Gya, what can we consequently see? To see anything we need to define some evolutionary tendencies that span the RNA world.

One of these is continuity, the property that links predecessors with descendants by small genetic changes. Evolutionary descent follows a path that takes finite steps, but each descendant is continuously linked to its forbear by individual practical changes in a genetic text. One frequently noted consequence of continuity on large time scales is that successful adaptations tend to persist, and are rarely discarded. For example, it seems likely to many that iron, nickel, copper, manganese, and molybdenum clusters within modern protein enzymes are a retained adaptation from a mineral chemistry (Rees 2002) that was once even more useful to life. In fact, such mineral cofactors probably predate the RNA world, surviving even beyond the gigayears that stretch between us and the era of RNA dominance. They therefore illustrate the permanence of useful biochemical adaptations. Persistence and exaptation of successful evolutionary devices therefore is a rule and not an exception. We will depend on parallel ideas in the following discussion.

A second useful idea is Bayes' Theorem, which is a rigorous statement about the probability of ideas in the light of new evidence, like new evidence of an RNA world (Yarus 2001). Bayes' application here is that your successors are the most likely outcome of your existence, or alternatively, that your predecessors were those most likely to give rise to the present situation. Bayes' Theorem again strongly reinforces the preservation of prior adaptations, not simply because they passively preexist in almost their current form (as for continuity, earlier), but because persistence in descent naturally results from the prior success of an adaptation. Continuity implies that evolution has a path, and Bayes' implies that this path has a probable functional logic. Continuity and Bayes' together constitute a mutually reinforcing argument for the persistence of prior adaptations.

Now we aggressively combine these ideas with the newly plausible existence of RNA-like molecules 4 Gya, to consider what might have existed relatively shortly before, during the time of the IDA, the primordial replicator.

5 PERSISTENCE OF CHNOPS

Continuity should extend even down to atomic composition. Therefore we assume here that the primordial replicator was composed of the most prevalent biological atoms; C, H, N, O, P, and S. This notion at first looks bland and

uncontroversial—but is worth emphasis because it wholly contradicts a frequently discussed idea, that of the “genetic takeover.” The idea, owing to Cairns-Smith (Cairns-Smith 1982), is that another kind of replicator entirely, perhaps a layer of clay, not only preceded the rise of RNA but also templated the first RNA molecules. In this case, a clay “genetic material” with prevalent Al and Si centers would give way to a CHNOPS replicator, thereby violating this implementation of continuity.

However, although clays do have useful interactions with ribonucleotides (discussed later), it is unclear whether clays can replicate any of their properties. It has been claimed that the redox and binding affinities of montmorillonite clays can be propagated from preexisting to newly grown layers of clay (Weiss 1981). But efforts to define and reproduce these observations have not been successful (Arrhenius et al. 1986). Thus replication of clay “information” has remained hypothetical, and transfer of replicated clay properties to nucleic acids even more so. In materials of unlimited complexity, the only replication that we know takes place entirely within molecules that are CHNOPS, so this continuity assumption seems apt for the present purpose.

6 PERSISTENCE OF RNA

Focusing more closely, the primordial replicator must give way to RNA at 4 Gya. Therefore the immediate ancestor must be chemically related to RNA by a plausible transformation. At least the last in this chain of evolutionary transitions must be small, implying a predecessor anchored to RNA by change of a few atoms. Therefore it is economical of assumptions and consistent with continuity and Bayes' to extend the resemblance to RNA further back. Evolutionary succession would somehow conserve a predecessor's chemical proficiencies. It is difficult to see how to do this in a primitive context except by propagating the forerunner's RNA-like atoms and bonding. Is there a potential replicator that fits these criteria?

7 THE HYPOTHESIS

I suggest that AMP-containing enzymatic cofactors are the modern descendants of the IDA. An example (NADP, the virtually ubiquitous redox cofactor) is shown in Figure 1A. These structures meet the logic of continuity and Bayes' Theorem proposed above. NAD and NADP in particular are small 5'-5' linked ribonucleic acids with a potential stacked structure (Fig. 1B), which aligns their nucleotides for coordinated base-pairing interaction during a possible complementary replication (Fig. 1C). Their original complements/templates (Fig. 1C) may not presently be used by modern protein enzymes—it seems to

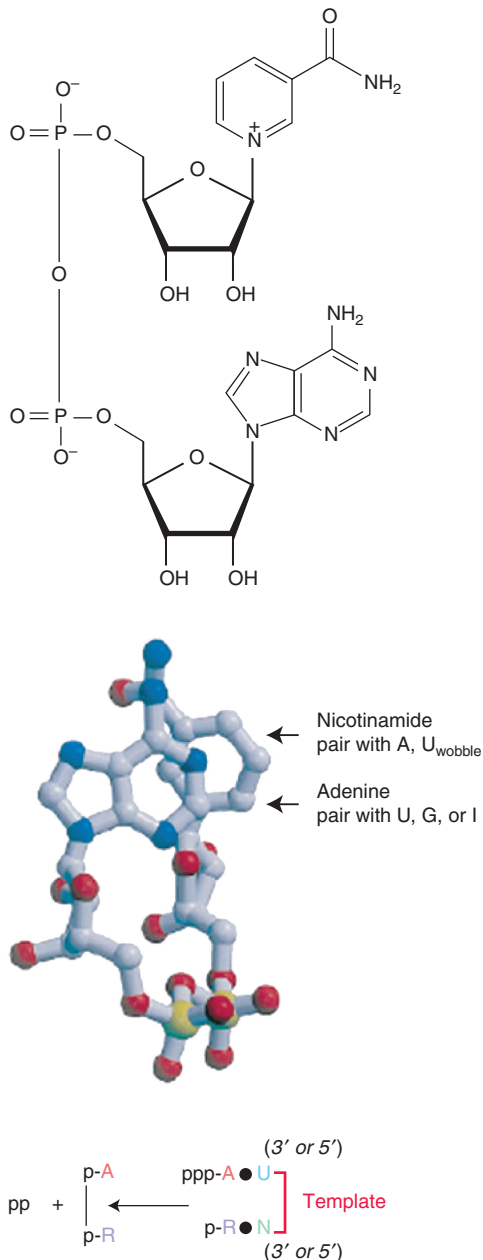


Figure 1. (A) Structure of NADP, the 3' phosphorylated derivative of nicotinamide adenine diphosphate (NAD). The image is a public domain illustration from Wikipedia. (B) A solution structure for NAD, with the nucleobases at the *top*, ribose in the *middle*, and the pyrophosphate at the *bottom*. Possible base pairings for replication are indicated. The figure is derived from the model of (Smith and Tanner 2000). (C) Replication scheme for an AMP-containing cofactor, like NAD or FAD. For simplicity, the backbone linking chemistry is shown as that for RNA catalysis (Huang and Yarus 1997), but this might vary. The complementary “template” strand and its replication on an NAD-like template are unspecified, because it is unclear what polarity it would have, or if it would be 5'-5' linked. However, 5'-5' polarity appears simplest, and is adopted in the text for clarity. R is nicotinamide nucleotide or a congener and N is its hypothetical complement.

be hoping for too much to suppose that both the primordial cofactor and its template became protein cofactors that survive today. Therefore, we are somewhat uncertain what the complements to the cofactors were or even what template backbones were because these may now be lost (Fig. 1 legend). However, it is notable that frequently, as for the chemically active nucleotides of NAD and FAD, both chemical centers and of course AMP could be complementary to existing nucleotides, as suggested by the notations within Figure 1B. For conciseness, discussion below takes both strands to be 5'-5' linked dinucleotides.

Although we came to this specific notion by a slightly exotic Continuity/Bayes' logic, the idea agrees broadly with a substantial literature about ancient molecules.

8 EARLIER IDEAS ABOUT NUCLEOTIDE COFACTORS

Casting NAD, FAD, and their related cofactors as the IDA is an extension of the influential notion of Harold White III (White 1976) that AMP-containing cofactors bear witness to the existence of a prior generation of RNA enzymes, whose AMP-containing reaction centers were appropriated by protein enzymes. In this article, we hypothetically extend the history of the cofactors further into the past and speculate that they may also have had a different and even more ancient role. More recently, Copley, Smith, and Morowitz (Copley et al. 2007) have emphasized that biocatalysis may have begun with small molecules, some of which lay on the evolutionary path to the RNA world. Though they were not thinking precisely of these same molecules, the present scheme is one implementation of such ideas.

9 PRE-RNA WORLD CREDENTIALS

The nicotinamide of NAD is easily reached by several synthetic routes beginning with plausibly prebiotic chemicals, like ethylene and ammonia (Friedmann et al. 1971) or aspartate and dihydroxyacetone phosphate (Cleaves and Miller 2001). NAD or a congener may have plausibly been one of the earliest cofactors to enter the biochemical inventory. It is therefore credible to consider nicotinamide-containing cofactors an early participant in biochemistry, before the rise of complex RNA catalysts.

10 AN OLD VARIANT BACKBONE

Desire for properties thought to be useful to primordial genetic materials, for example, novel chemical resistance or achirality, has led to varied ingenious proposals for alternate backbone composition and continuities (see Robertson and Joyce 2010; Benner et al. 2010). It is therefore doubly striking

that there exists a natural alternative ribose/phosphodiester backbone universal in modern biochemistry, the 5'-5' linked cofactors, that possesses frequently desired properties. For example, the prominent sensitivity of natural RNA to hydrolysis by any factor that stabilizes the ribose 2' oxyanion (for example, bases and metals) does not exist for NAD and its congeners. It is also conceivable (see Fig. 1B) that the cofactors, because they do not repeatedly employ their chiral atoms directly in the backbone, may be more tolerant of variant sugars, as would be prevalent in a primitive milieu. For example, because the chiral sugars are at the ends rather than in the middle of the backbone, oligomerization of the 5'-5' nucleotides may be less sensitive to sugar chirality than 5'-3' RNA (Joyce et al. 1984). Finally, the cofactors are particularly ancient on the evidence of their own ubiquitous metabolic uses, but similar old structures also mark the 5' terminus of all eukaryotic messages (reviewed in Schoenberg and Maquat 2009).

11 AN AUGMENTED ALPHABET, AN AUGMENTED MANIFOLD OF REACTIONS

Benner (Benner et al. 2010) has argued the benefits of a more diverse set of nucleobases. By synthesis of a third base pair, his, and other, laboratories have shown that a three complementary pair/six base system is capable of replication (Kimoto et al. 2009; Yang et al. 2007). It is striking that the 5'-linked dinucleotide cofactors deploy an expanded repertoire of nucleotide residues, and automatically use their more versatile reactivity. Thus it is manifest that such structural diversity enhances chemical versatility. In fact, it is quite frequent that free modern cofactors have a diminished version of the chemical activity that they display in conjunction with a protein enzyme. For example, NADH, the quintessential redox cofactor (here in its reduced form), will reduce nitro blue

tetrazolium in the presence of phenazine methosulfate. This reaction underlies a classical NAD assay and cytochemical stain (Ponti et al. 1978). NADH-dye reduction goes via superoxide aerobically, but is almost as fast by direct means in the absence of oxygen (Compare Fig. 2C).

12 CHICKENS AND EGGS

The NAD reaction just mentioned would be particularly notable in a pre-RNA and RNA world, because of the complete absence of redox centers in four-nucleotide RNA (compare (Tsukiji et al. 2003)). Such new reactions are potentially crucial because they supply one key to the Darwinian evolution of the molecule. An appeal of the RNA world hypothesis is that it solves the “chicken and egg” problem; it shows that in an earlier, simplified biota the genotype/replicator and phenotype/catalyst could have been one and the same molecule. An exciting aspect of the current notion is that the genotype and phenotype could have been joined more simply, and earlier. The cofactors suggest that genotype/phenotype junction could exist in a tiny system consisting of two complementary two-nucleotide entities. Such IDAs would be selected for both replication and activity (would have undergone Darwinian evolution) (Fig. 2) because their biochemistry became useful via their own replication to effective levels.

13 CONTINUITY WITH AN RNA WORLD

Continuity for nucleoside diphosphate cofactors and RNA seems assured because RNA-cofactor binding and RNA-cofactor chemistry are already known in varied forms (reviewed in Jadhav and Yarus 2002b). Catalytic RNAs can synthesize A-containing cofactors—RNAs have been selected which form nucleoside 5'-5' structures via attack

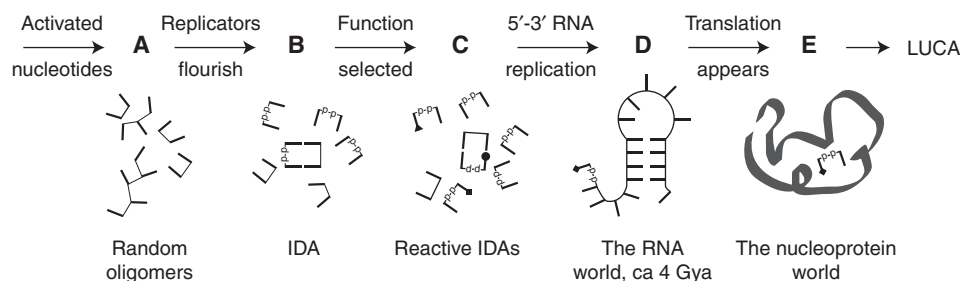


Figure 2. The IDA in context; the origin of life. (A) Activated nucleotides and compatible molecules oligomerize arbitrarily. (B) Replicators necessarily become abundant, by templating with minimal catalysis. (C) 5'-5' replicators with reactive nucleotides are selected to participate in metabolism. (D) 5'-3' RNA replicase creates an RNA world, \pm 5'-5' cofactor initiation and reactivity. (E) RNAs devise translation; 5'-5' cofactors are adopted by peptide catalysts (ribbon). The pathway is initiated by its most complex event, geochemical creation of several activated nucleotide-like materials. Thus, while “simple” is a debatable evolutionary characterization, progress might be relatively simple once begun. After panel (B), all crucial transitions depend on somewhat similar selections for enhanced chemical proficiency.

by varied phosphorylated substrates (Huang and Yarus 1997); at the α -P of their own 5' terminal triphosphate. When such 5'-5' nucleoside-forming reactions are performed using RNAs initiated with 5' ATP, attacks by phosphopantetheine, NMN and FMN create bona fide covalently linked CoA, NAD, and FAD as the 5' terminus of the catalyst (Huang et al. 2000). The ambit of RNA activities also includes the synthesis of free 5' diphosphate-linked ribonucleotides themselves, using free nucleotide substrates to form free small molecules with cofactor-like linkages (Huang et al. 1998).

5-prime cofactor RNAs may have been frequent agents in a diverse ribozyme metabolism (White 1976). Indeed, a subset of cofactors, like NADP (Fig. 1A) and CoA, have 3' A phosphates, suggesting continuation by an RNA backbone downstream, now perhaps lost. Metabolically significant activities have been measured for such 5' cofactor-RNAs. An RNA catalyst that uses a short recognition sequence near the 5' terminus can attach coenzyme A *in trans* to a recognized RNA (Jadhav and Yarus 2002a). If the CoA-tagged RNA is itself a suitable catalyst (which can be isolated from preformed CoA-randomized RNA), it will synthesize acyl-CoA-RNAs, including acetyl-CoA-RNA and butyryl-CoA-RNAs (Jadhav and Yarus 2002a). An impressively complete case is that of a ribozyme that binds either free NAD or NADH and uses the cofactor to oxidize a linked benzyl alcohol (Tsukiji et al. 2003) or to reduce a linked benzaldehyde (Tsukiji et al. 2004).

The ready chemical synthesis of 5'-5' nucleotides from activated ribonucleotides is often followed by their incorporation at the terminus of a longer 3'-5' or 2'-5' RNA backbone (Ferris and Ertem 1992). The same ease of 5' incorporation applies, as just reviewed, to RNA-catalyzed incorporation of 5'-5' structures at the RNA termini. In addition, terminal or noncovalently bound cofactor-RNAs are accessible for further biochemical transformation, in which a 5'-5' linked cofactor performs RNA enzyme chemistry. Accordingly, any activities evolved for cofactors before the rise of RNA catalysts could be quickly adopted into an emergent RNA metabolism.

14 NOVEL REPLICATION CHEMISTRY

But might 5'-5' cofactors preexist to join the RNA world as replicators? Cofactor chemistry seems favorable to the hypothesis, though decisive barriers remain.

Because of the superior nucleophilicity of the 5' hydroxyl, incubation of 5' activated nucleotides leads to abundant 5'-5' products. AMP and rA with carbodiimide activation give A5'pA (Sulston et al. 1968). AMP alone under the same conditions gives A5'p-pA as the majority product (Sulston et al. 1968; Robertson and Joyce 2010) which can

then be incorporated at the 5' terminus of a 5'-3' polymer, for example, in the presence of montmorillonite clay (Ferris and Ertem 1992). Thus the nucleotide backbone in 5'-5' linked cofactors like NAD is the one most readily formed by 5' activated nucleotides.

NAD and its congeners are COSMIC LOPER, an acronym coined (Benner and Switzer 1999) to emphasize that a genetic material must have the unusual capability of tolerating changes in structure (mutations) without compromising the shared, essential ability of such varied molecules to replicate. This chemical property was envisioned as dependent on the dominance of the hydrophilic charged backbone of the nucleic acids. The 5'-diphosphate cofactors have, if anything, a more polar backbone than normal 5'-3' RNA. This is illustrated by the existence of varied cofactors (NAD, FAD, CoA, and SAM), all retaining their solubility, somewhat similar form, and biochemical roles despite differences in chemical makeup.

Accelerated phosphodiester bond formation and even complete replication in an all-nucleic acid context have been broadly observed, and are not an unprecedented conjecture. Symmetrical RNA-RNA ligases with complementary overlaps will replicate exponentially without apparent limits, each ligating the complementary ligase together from its supplied substrate fragments (Lincoln and Joyce 2009). Complimentary deoxytrinucleotides will replicate a hexanucleotide template when activated by carbodiimide, and similar reactions extend to full replication of both strands (Sievers and Von Kiedrowski 1994). Although full replication of an oligomeric template by individual nucleotide addition has not been achieved, large ribozymes do accurately fill in small, primed templates (Zaher and Unrau 2007). More simply, activated ribonucleotides (e.g., GMP imidazolide) polymerize on a template like poly C (Inoue and Orgel 1983; Orgel 1992). However, extension to other systems than oligo C template-directing GMP imidazolide has proven difficult, and good enantiopurity is essential for chain growth (because one sugar enantiomer poisons chain growth by the other) (see Robertson and Joyce 2010) (Joyce et al. 1984).

15 CATCH-22

I have found no published evaluation of whether single activated nucleotides polymerize on a dinucleotide template (Fig. 1C), especially in 5'-5' linkage. This is critically important because such chemistry potentially solves the evolutionary problem of needing a replicator to evolve an effective replicator (Robertson and Joyce 2010). That is, the miniaturization of the replicator to one internucleotide bond in each complement, as in the cofactors, may make it possible to envision replication using activated nucleotides

and a template alone, removing or diminishing the need for a specific replication catalyst. Thus Darwinian evolution of RNA would no longer be caught in the replication Catch-22.

Despite the absence of dinucleotide replication data, in an Orgel system, with a polymeric template and complementary activated nucleotide substrates, the initial two nucleotides in a new 5'-nucleotide-initiated chain are successful in a reaction of similar molecularity. Further, whatever the result with imidazolides, there are effective, new and more "biological" activated nucleotides, like 3'-5' cyclics, to try (Costanzo et al. 2009).

However, given that a 5'-5' dimer template and potential substrate nucleotides are small molecules with enhanced conformational freedom, it may well be that such systems will need the stimulation observed for activated nucleotide polymerization in ice-bound eutectic solutions (Monnard et al. 2003) or facilitation by montmorillonite clay matrices (Ferris and Ertem 1992). They may also benefit from the support provided under eutectic conditions for the activities of fragmented RNA structures (Vlassov et al. 2004). Small size therefore decisively simplifies Darwinian evolution (see *Chickens and Eggs* and *Catch-22* earlier), even though it may also slow replication chemistry. However, there seems no reason why the IDA may not have enjoyed both the cold and the clay.

Replication thus remains an essential issue to be decided. And even beyond replication, there are many footholds for a critic in the ideas above. Moreover, even if these ideas are entirely correct, we will take a step back toward the IDA, but will still not know that it is the last step. Nonetheless, in my own defense, the argument's gaps seem experimentally amenable. There may well be a new world to find just beyond the RNA world we know.

ACKNOWLEDGMENTS

Many thanks to Gunter Von Kiedrowski and Hyman Hartman for a discussion of clay replication, and to Leslie Leinwand and Bill McClain for comments on a draft manuscript. During the preparation of this monograph, I am grateful for support from the National Institutes of Health (research grant GM 48080), NASA (Colorado University Astrobiology Center NCC2-1052) and the University of Colorado Council on Research and Creative Work.

REFERENCES

- Arrhenius G, Cairns-Smith AG, Hartman H, Miller SL, Orgel LE. 1986. Remarks on the Review Article "Replication and Evolution in Inorganic Systems" by Armin Weiss. *Angew Chem Int Ed Engl* **25**: 658.
- Benner SA, Switzer CY. 1999. Chance and necessity in biomolecular chemistry: Is life as we know it universal? In: H Frauenfelder, J Drenth, PG Wolynes (eds) *Simplicity and complexity in proteins and nucleic acids*. Dahlem University Press, pp 339–363.
- Benner S, Kim H-J, Yang Z. 2010. Setting the stage. The history, chemistry, and geobiology behind RNA. *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a003541.
- Cairns-Smith AG. 1982. *Genetic takeover and the mineral origins of life*. Cambridge University Press, Cambridge.
- Chen X, Li N, Ellington AD. 2007. Ribozyme catalysis of metabolism in the RNA world. *Chem Biodivers* **4**: 633–655.
- Cleaves HJ, Miller SL. 2001. The nicotinamide biosynthetic pathway is a by-product of the RNA world. *J Mol Evol* **52**: 73–77.
- Copley SD, Smith E, Morowitz HJ. 2007. The origin of the RNA world: Co-evolution of genes and metabolism. *Bioorg Chem* **35**: 430–443.
- Costanzo G, Pino S, Ciciriello F, Di Mauro E. 2009. Generation of long RNA chains in water. *J Biol Chem* **284**: 33206–33216.
- Dalrymple GB. 1991. *The age of the Earth*. Stanford University Press, Palo Alto, CA.
- Ferris JB, Ertem G. 1992. Oligomerization of ribonucleotides on montmorillonite: reaction of the 5'-phosphorimidazolide of adenosine. *Science* **257**: 1387–1389.
- Friedmann N, Miller SL, Sanchez RA. 1971. Primitive earth synthesis of nicotinic acid derivatives. *Science* **171**: 1026–1027.
- Huang F, Yarus M. 1997. 5'-RNA self-capping from guanosine diphosphate. *Biochemistry* **36**: 6557–6563.
- Huang F, Bugg CW, Yarus M. 2000. RNA-Catalyzed CoA, NAD, and FAD synthesis from phosphopantetheine, NMN, and FMN. *Biochemistry* **39**: 15548–15555.
- Huang F, Yang Z, Yarus M. 1998. RNA enzymes with two small-molecule substrates. *Chem Biol* **5**: 669–678.
- Illangasekare M, Yarus M. 1999. Specific, rapid synthesis of Phe-RNA by RNA. *Proc Natl Acad Sci* **96**: 5470–5475.
- Illangasekare M, Sanchez G, Nickles T, Yarus M. 1995. Aminoacyl-RNA synthesis catalyzed by an RNA. *Science* **267**: 643–647.
- Inoue T, Orgel LE. 1983. A nonenzymatic RNA polymerase model. *Science* **219**: 859–862.
- Jadhav VR, Yarus M. (2002a). Acyl-CoAs from coenzyme ribozymes. *Biochemistry* **41**: 723–729.
- Jadhav VR, Yarus M. (2002b). Coenzymes as coribozymes. *Biochimie* **84**: 877–888.
- Joyce GF, Visser GM, van Boeckel CA, van Boom JH, Orgel LE, van Westrenen J. 1984. Chiral selection in poly(C)-directed synthesis of oligo(G). *Nature* **310**: 602–604.
- Kimoto M, Sato A, Kawai R, Yokoyama S, Hirao I. 2009. Site-specific incorporation of functional components into RNA by transcription using unnatural base pair systems. *Nucleic Acids Symp Ser (Oxf)*: 73–74.
- Lee N, Bessho Y, Wei K, Szostak JW, Suga H. 2000. Ribozyme-catalyzed tRNA aminoacylation. *Nat Struct Biol* **7**: 28–33.
- Lincoln TA, Joyce GF. 2009. Self-sustained replication of an RNA enzyme. *Science* **323**: 1229–1232.
- Luo ZX, Crompton AW, Sun AL. 2001. A new mammaliaform from the early Jurassic and evolution of mammalian characteristics. *Science* **292**: 1535–1540.
- Monnard P, Kanavarioti A, Deamer D. 2003. Eutectic phase polymerization of activated ribonucleotide mixtures yields quasi-equimolar incorporation of purine and pyrimidine nucleobases. *J Am Chem Soc* **125**: 13734–13740.
- Nissen P, Hansen J, Ban N, Moore PB, Steitz TA. 2000. The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**: 920–930.
- Orgel LE. 1992. Molecular replication. *Nature* **358**: 203–209.
- Ponti V, Dianzani MU, Cheeseman K, Slater TF. 1978. Studies on the reduction of nitroblue tetrazolium chloride mediated through the action of NADH and phenazine methosulphate. *Chem Biol Interact* **23**: 281–291.
- Rees DC. 2002. Great metalloclusters in enzymology. *Annu Rev Biochem* **71**: 221–246.
- Robertson MP, Joyce GF. 2010. The origins of the RNA world. *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a003608.

- Schoenberg DR, Maquat LE. 2009. Re-capping the message. *Trends Biochem Sci* **34**: 435–442.
- Sievers D, Von Kiedrowski G. 1994. Self-replication of complementary nucleotide-based oligomers. *Nature* **369**: 221–224.
- Smith PE, Tanner JJ. 2000. Conformations of nicotinamide adenine dinucleotide (NAD(+)) in various environments. *J Mol Recognit* **13**: 27–34.
- Sulston J, Lohrmann R, Orgel LE, Miles HT. 1968. Nonenzymatic synthesis of oligoadenylates on a polyuridylic acid template. *Proc Natl Acad Sci* **59**: 726–733.
- Szathmary E. 2006. The origin of replicators and reproducers. *Philos Trans R Soc Lond B Biol Sci* **361**: 1761–1776.
- Tsukiji S, Pattnaik SB, Suga H. 2003. An alcohol dehydrogenase ribozyme. *Nat Struct Biol* **10**: 713–717.
- Tsukiji S, Pattnaik SB, Suga H. 2004. Reduction of an aldehyde by a NADH/Zn²⁺-dependent redox active ribozyme. *J Am Chem Soc* **126**: 5044–5045.
- Turk R, Chumachenko NV, Yarus M. 2010. Multiple translational products from a five-nucleotide ribozyme. *PNAS* **107**: 4585–4589.
- Vlassov A, Johnston B, Landweber L, Kazakov S. 2004. Ligation activity of fragmented ribozymes in frozen solution: implications for the RNA world. *Nucleic Acids Res* **32**: 2966–2974.
- Weinger JS, Parnell KM, Dorner S, Green R, Strobel SA. 2004. Substrate-assisted catalysis of peptide bond formation by the ribosome. *Nat Struct Mol Biol* **11**: 1101–1106.
- Weiss A. 1981. Replication and evolution in inorganic systems. *Angewandte Chemie-International Edition in English* **20**: 850–860.
- Welch M, Majerfeld I, Yarus M. 1997. 23S rRNA similarity from selection for peptidyl transferase mimicry. *Biochemistry* **36**:6614–6623.
- White HB, III. 1976. Coenzymes as fossils of an earlier metabolic state. *J Mol Evol* **7**: 101–104.
- Yang Z, Sismour AM, Sheng P, Puskar NL, Benner SA. 2007. Enzymatic incorporation of a third nucleobase pair. *Nucleic Acids Res* **35**: 4238–4249.
- Yarus M. 1991. An RNA-amino acid complex and the origin of the genetic code. *New Biol* **3**: 183–189.
- Yarus M. 2001. On translation by RNAs alone. *Cold Spring Harb Symp Quant Biol* **66**: 207–215.
- Yarus M, Caporaso JG, Knight R. 2005. Origins of the genetic code: The escaped triplet theory. *Annu Rev Biochem* **74**: 179–198.
- Yarus M, Widmann JJ, Knight R. 2009. RNA-amino acid binding: A stereochemical era for the genetic code. *J Mol Evolution* **69**: 406–429.
- Zaher HS, Unrau PJ. 2007. Selection of an improved RNA polymerase ribozyme with superior extension and fidelity. *RNA* **13**: 1017–1026.

Getting Past the RNA World: The Initial Darwinian Ancestor

Michael Yarus

Toc Blurb: Nucleotide cofactors such as NADP could represent relics of pre-RNA life on Earth. They may have been key elements of the first replicators, subsequently co-opted by RNA and then proteins.