Ahead and behind: a small, small RNA world

MICHAEL YARUS
Department of MCD Biology, University of Colorado, Boulder, Colorado 80309-0347, USA

Is the RNA World hypothesis pertinent to the origin of life?

The discovery of RNA catalysis (by the Cech lab 1982; and the Altman lab 1983) led quickly to the idea that the first catalysts might have been RNA-like (from W. Gilbert 1986). This notion maintains that phenotypes (chemical capability) and genotypes (capability for replication) might be combined in a single RNA. In this way, defining life as a replicator that chemically alters its environment, the problem of life’s origin might be streamlined to the origin of a seemingly familiar kind of ribonucleotide molecule.

However, the experimental demonstration of this process with specific RNA structures hits some thorny obstacles. Nucleotides are difficult to synthesize and are unstable (for example, see Benner in 2012) and their polymers are even more so (for example, see Bowler et al. in 2013). Even worse, related problems seem inevitable, rooted in the basic combinatorics of the situation. For example, there are $1.3 \times 10^{30}$ different sequences of four nucleotides that are 50 nucleotides long. Supposing that a few RNAs capable of replicating such RNAs might be among these 50-mers (but compare Wochner et al. in 2011), how would the few replicators favor a staggeringly small minority of useful 50-mer molecules? Actually, with random sequences all sequences are not likely present without multiple copies of each sequence. Five copies implies that at least one copy of all sequences is present with probability 0.9993; but this requires 180,000 metric tons of 50-mer RNA! While this problem can be softened by more congenial assumptions, it cannot be entirely avoided. For example, if we are interested in 30-mer RNAs, “having them all” requires 0.1 gram. Accordingly, comprehensive RNA accumulations containing RNAs of usual ribozymic size (several tens of nucleotides) seem unimaginable on the primitive Earth. To be more explicit, it does not seem likely that molecules of this size are likely to appear under primordial conditions, to be replicated, or to be selected via their (necessarily) extremely atypical chemical activity, even if they somehow do appear.

However, a possible solution is suggested by the same calculations that identify the problem. If present-day ribozymes are too large, then consider something smaller.

Embodiments of the oligonucleotide ribozyme

A substantial argument for this point of view is that real, minuscule nucleotide structures do have biochemical activities (suggested by Copley et al. in 2007). For example, there is RNA hydrolytic activity which accompanies the binding of metal ions to RNA. For example, the fixation of Mn$^{2+}$ within a tetramer paired to a triemer helps extract a proton from the nearby ribose 2’OH, activating hydrolysis of the tetramer (compare Kazakov and Altman 1992).

But there are also tiny RNAs that have more elaborate chemical capability; consider the nucleotide coenzymes (Yarus in 2011), which mediate the reactions conducted by $\sim$50% of modern “protein” enzymes. These have varied structures (review in Richter 2013), but often are ribodinucleotides with AMP 5’-5’ linked to a second modified ribonucleotide structure with special chemical capabilities. These coenzymes are often cited as a remnant of ancient RNA catalysts (as first suggested by White 1976; elaborated by Yarus 2011). Most particularly, because they are composed of nucleotide-like moieties, they might be descended from replicators. Modern enzymatic cofactors, like NAD$^+$, would require only slight changes to lose their replication, especially in view of their biological universality today. Such universality implies that gigayears of evolution have altered similar enzymatic cofactors.

Moreover, pure standard oligonucleotides just above this dinucleotide size range are truly enzymatic. The synthesis of aminoacyl-RNA from its biological precursor, amino acid adenylate (aa-AMP), a required step in translation, was first selected in 95-mer RNAs (by Illangasekare et al. 1995). However, using an improved, less-constrained selection, a much simpler ribozyme appears that binds this complex aa-AMP substrate molecule in trans, and transfers the amino acid specifically, in emulation of this universal translational reaction (see in Turk et al. 2010). The pentamer oligonucleotide 5’ GUNMO 3’ base pairs to tetramer oligonucleotide

© 2015 Yarus This article, published in RNA, is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at http://creativecommons.org/licenses/by-nc/4.0/.
5′ OMNU 3′ (italics indicate complementarity; Illangasekare and Yarus 2012) to assemble a three-nucleotide active site which transfers amino acid. The small active center which makes OMNU-aminoacyl is regiospecific, utilizing only the 2′OH, and stereospecific, as are the natural class I protein aminoacyl tRNA synthetases (from Eriani et al. 1990). The same active center, composed of GU and U brought together by nonspecific base pairs (Illangasekare and Yarus 2012), can be utilized in successive reactions that both activate the amino acid, using a preselected ribozyme (Xu et al. in 2014), and then transfer it to the tetramer RNA. These latter multi-stage reactions therefore emulate both translational activation and acylation steps as performed by protein enzymes today.

However, probably because the active center is too small to have significant contact with the amino acid or the nucleobase of aa-NMP, it is not specific for the amino acid or the base (see Turk et al. 2010). This is likely to be an inherent property of small catalysts, not unique to this RNA (discussed in Yarus 2011). When one can span only a bond length or two from the atoms actually reacting, opportunity for specific contacts is limited. Thus, small active centers cannot be as precise as when more architecture supports their sites. Such catalysts may not even be able to spend full time in catalytically configurations, so they will also accelerate reactions less than do active centers embedded in larger ribozymes. In fact, these tendencies are evident in the reactions of the five-nucleotide aminoacylator (as seen in Turk et al. 2011 and Yarus 2011). So, though they are unlikely to be superior catalysts, five-nucleotide RNA enzymes with three-nucleotide active centers are well known.

The sporadically fed pool

Further, beyond the five-nucleotide aminoacylating ribozyme, we can argue that oligonucleotides are capable of solving most of the problems associated with the initiation of Earthly life (compare the discussion in Szostak 2012). This notion comes from calculations on uncatalyzed RNA synthesis in the “sporadically fed pool,” which implements a type of oligonucleotide synthesis which would likely be practical, or perhaps even frequent, on a primitive planet.

A sporadically fed pool is one that gets small doses of activated and unactivated nucleotides at low concentrations; these substrates also arrive at uncontrolled random times. These nucleotides and their possible oligomerized products immediately begin to decay at realistic rates. Surprisingly, in such a pool, using rates for ribonucleotide reactions from the chemical literature, small self-complementary RNAs readily appear and even replicate (Yarus 2012). Thus short RNAs can appear (Yarus 2011) as soon as there are activated nucleotides, in contrast to specific longer sequences which seem (in the first paragraphs above) inaccessible. Further, replication seems to need nothing more exotic than ordinary solution kinetics (Yarus 2012). Later, when the ability to make longer RNAs appears, small active centers can be merged into the interiors of larger chains (Illangasekare and Yarus 2012), assuring their evolutionary continuity.

However, this description leaves unmentioned the most interesting properties of a sporadically fed pool, which are suggested by its detailed kinetics (see Yarus 2013). Because substrate arrival is unguided, all possible reaction sequences are explored. Successful syntheses therefore can come from recurrent near-ideal reactions, in which proper reagent addition and sequence is recurrently achieved with no external guidance at all. Further, because the assumed dominant form of molecular decay (first order, exponential) has a large variance (protracted tailing in time), intermediates persist and interact with later substrate pulses. Thus, long, successful complex reaction sequences are more frequent than intuition expects. Accordingly, a plausibly primitive reactor, the sporadically fed pool, can recurrently host substantial reaction sequences, perform unexpectedly complex syntheses (Yarus 2013), and potentially contribute these products to a descendant. The short oligonucleotides we have been talking about have an apparently practical route into a primordial world.

Conclusion—the promise of small oligonucleotides

In celebration of the 20-year anniversary of RNA, then, here is a prediction about its future. It will be productive to examine the chemical abilities of short oligomers of varied structures (but surely including 2′-5′, 3′-5′, and 5′-5′ connectivity) and varied nucleobase compositions. There are strong arguments for a long-ago, small, small RNA world populated with Darwinian beings who were short oligomers; perhaps we can recreate that first biosphere.

A normally referenced version of this note is available: mcdb.colorado.edu/labs/yarus/index.html