Putting together the pieces: Evidence suggests that RNA was a product of evolution

Brian Cafferty and Nicholas V. Hud, Georgia Institute of Technology, Atlanta, GA, USA

For the past four decades, prebiotic chemists have attempted to demonstrate the formation of RNA polymers by plausible prebiotic reactions. There have been notable advances, but to be certain, the spontaneous formation of RNA remains a grand challenge in origins of life research. From a different perspective, there are reasons to seriously consider the possibility that RNA is a product of evolution. If so, there may have never been a prebiotic mechanism that produced RNA polymers. We subscribe to this latter view and hypothesize that RNA is the penultimate member of continuous lineage of genetic polymers, with DNA being the ultimate member of this lineage. In this essay, we briefly summarize the case for why RNA is likely the descendant of one or more pre-RNA polymers that spontaneous assembled on the prebiotic earth.

Nucleo*sides* are each an assemblage of a nucleobase and a ribose sugar, whereas nucleo*tides*, the monomeric units of RNA, are phosphorylated nucleosides (Figure 1). Prebiotic chemists have typically sought to form RNA in a sequential fashion, starting with the formation of nucleotides, followed by their polymerization (Figure 1). However, of the four canonical RNA bases (adenine, cytosine, guanine, uracil), only adenine has been found to react with ribose in a model prebiotic reaction to produce nucleosides in appreciable yields (i.e., about 2%). The other three canonical nucleobases do not produce nucleosides when dried and heated with ribose. This apparent roadblock in RNA synthesis motivated the Orgel laboratory and, more recently, Sutherland and co-workers, to investigate the possibility that the nucleobases were first formed on a pre-existing sugar. The successful production of the canonical pyrimidine nucleotides (cytidine and uridine) by this approach is certainly an elegant chemical achievement,[1] but it is still too early to declare RNA a plausible prebiotic molecule.

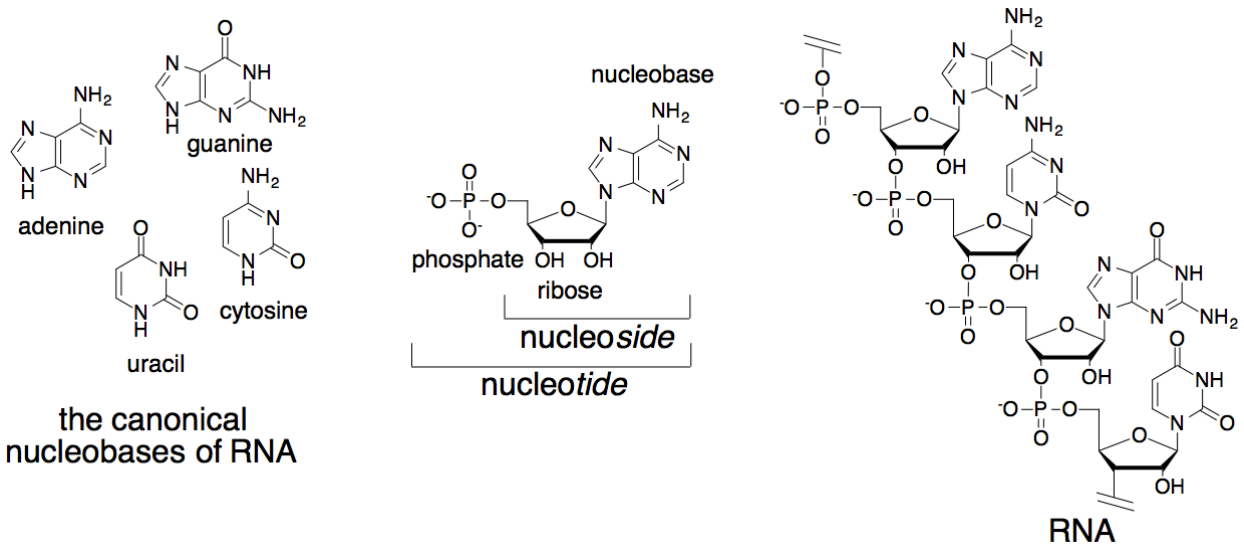

Figure 1: Chemical structures of the RNA nucleobases, nucleosides, nucleotides and polymer.

A notable success of prebiotic chemistry is the demonstration that all four nucleobases can be produced by reactions from simple starting materials, such as HCN and formamide.[2-4] However, these nucleobases are produced along with many other closely related heterocycles (e.g., other pyrimidines and purines). Likewise, ribose can be produced from formaldehyde by an abiotic reaction, but it is produced with an array of other sugars.[5,6] The great complexity of product mixtures resulting from most model prebiotic reactions can be seen as an additional challenge *or* as something that aided the chemistry at the origin of life. On the one hand, complex mixtures would complicate some proposed prebiotic routes to RNA, as there would have been molecules in such mixtures that reacted more readily with nucleotide precursor molecules, thereby depleting the building blocks of RNA. On the other hand, some heterocycles that formed nucleotides with ribose (or another sugar) and were capable of their own mutual selection from a complex mixture would have been better suited for the formation of an early informational polymer (a "pre-RNA"). For this reason, it is vital to test how robust proposed pathways to RNA or pre-RNA monomers and polymers are when carried out in chemically diverse (i.e., non-idealized) and thus more prebiotically relevant reaction conditions.

If a different informational polymer came before RNA, then it is possible that the nucleobases other than those used in modern nucleic acids formed nucleosides much more easily with ribose. Miller and coworkers were the first to experimentally demonstrate the feasibility of this hypothesis by showing that a five-membered ring called urazole (which shares structural similarity with uracil) readily forms nucleosides with ribose. In our laboratory, we later showed that 2-pyrimidinone, a pyrimidine base that is similar to uracil and cytosine, also forms nucleosides when dried and heated with ribose. While these experiments demonstrate the important point that some heterocycles (similar to the current nucleobases) will form

nucleosides in high yield in an abiotic reaction, neither urazole nor 2-pyrimidinone is very attractive as a pre-RNA base, since neither molecule forms base pairs that are as stable as the canonical Watson-Crick base pairs.

It is important to recognize that the canonical nucleobases *do not* form Watson-Crick base pairs as monomers in aqueous solution; RNA strands must be at least four or five nucleotides in length to make stable duplexes that are held together with base pairs. This fact presents what we have called "The Paradox of Base Paring." [7] Briefly: How would prebiotic reactions have anticipated that the canonical bases would pair *after* the formation of RNA *polymers* if the bases were not pairing already? Since so many other molecules are created by the model prebiotic reactions that produce the canonical nucleobases, it is likely that any hypothetical process for prebiotic nucleotide polymerization (that does not require base pairing) would result in polymers with many non-pairing units – a major obstacle for starting the RNA world!

We have hypothesized that the original nucleobases were selected from a complex pool of heterocycles because they were capable of pairing with each other even at the monomer level, or that another class of molecules promoted base pairing.[8] Recently, we have found that two nucleobase analogs, cyanuric acid and a modified triaminopyrimidine, will form base pairs and highly ordered supramolecular structures in aqueous solution (Figure 2). Such a property would have provided a means to locally concentrate and organize these pre-nucleobases before they were joined by a common backbone. If the same nucleobases can be shown to form nucleosides readily with ribose, then the realization of a one-pot prebiotic reaction for the formation of pre-RNA polymers could be within reach.

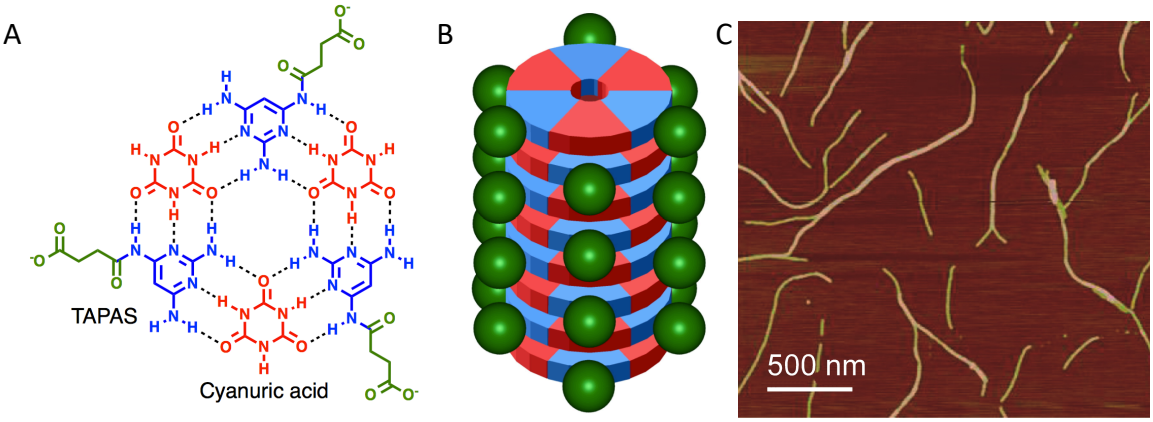

Figure 2: A) Hydrogen-bonded hexad formed by cyanuric acid and TAPAS (2,4,6-triaminopyrimidine with succinate conjugate). B) Schematic representation of the self-assembled non-covalent polymers formed by cyanuric acid and TAPAS in aqueous solution (colors correspond to chemical structures in A). C) Atomic force microscopy (AFM) image of non-covalent

polymers that spontaneously formed upon the mixing of cyanuric acid and TAPAS at 15 mM in each monomer.[9]

The presence of ribose in RNA presents its own reasons to question the prebiotic origin of RNA. The synthesis of ribose is not particularly robust compared to other sugars. As one example, ribose is only a minor component of the formose reaction, suggesting that ribose would have been scarce on a prebiotic earth. Decades of research into the chemical etiology of RNA, through the tireless synthesis of non-natural nucleic acid backbones with alternative sugars, has also shown that ribose is not unique in its ability to support the pairing of nucleobases.[10] Thus, there may be numerous possible pre-RNA polymers. This view is further supported by recent findings that nucleic acids containing even simpler sugars than ribose, such as the four-carbon sugar tetrose, can fold into functional three-dimensional structures.[11] More dramatic deviations from the RNA structure have also shown that base-pairing capacity is still maintained when ribose is replaced by something as different as an amino acid.[12] Nevertheless, in comparison to these other backbone linkers, ribose looks optimal in its ability to form both duplexes and more complex structures, such as those found in the active core of the ribosome. Overall, the combination of the difficulty of synthesizing ribose under prebiotic conditions, juxtaposed with its perfect functionality in extant life, make ribose look more like a product of evolution than a "frozen accident" (i.e., a sugar that was initially incorporated because it was simply available).

The presence of phosphorus in the backbone of RNA raises even more questions about the prebiotic plausibility of this organic-inorganic polymer. Like ribose, it is not obvious how phosphate would have entered into RNA by purely abiotic mechanisms. The availability of free phosphate during the prebiotic epoch was likely quite low, as the strong association of phosphate and divalent metal ions (e.g., Mg^{2+}) would have held most of the Earth's phosphate in the mineral state. It has recently been argued that phosphite, which has one less oxygen bound to phosphorus than phosphate, may have provided an easier route to prebiotic phosphorylation.[13] However, even if a plausible prebiotic mechanism can be found for the efficient phosphorylation of nucleosides, there are additional reasons to expect that phosphate was a later addition to informational polymers.

The phosphodiester linkage, which connects the nucleotides, is kinetically stable but thermodynamically unstable. These attributes, respectively, limit spontaneous hydrolysis and allow facile cleavage by enzymes. While important for phosphate's functions in contemporary biology, these same attributes would come with severe challenges for the formation of polymers during the earliest stages of life. For example, it is not obvious how phosphate-linked nucleosides would have been cycled between their monomeric and polymeric states without regular activation, and reactivation, by high-energy compounds.

One potential resolution to the difficulty of having phosphate part of the first informational polymers is possibility that other nucleoside linkages came before the phosphodiester linkage. For example, we have proposed that acetal linkages, like the chemical bond that link the sugars of polysaccharides, could have originally linked the nucleosides of a pre-RNA polymer. Such linkages are of sufficiently low energy of formation that they spontaneously form when nucleosides are dried and heated with glyoxylate, a small organic molecule that is similar in size and shape to phosphate.[14] These linkages would also be less kinetically stable compared to the phosphodiester linkages but, combined with their lower energy of formation in water, they would have facilitated the rate at which pre-RNA polymers were recycled — properties that would have enabled the first informational polymers to search through sequence space for functional activity (i.e., enzymes).

Implicit in the RNA-world hypothesis is a transition that took place sometime in our distant past when DNA superseded RNA to become the predominant information storage polymer of life. Removal of a single atom (the 2′ oxygen of ribose) converts the RNA backbone into the DNA backbone, a closely related polymer that could hardly be more different in its chemical stability. Because RNA is more labile, from a chemical perspective, it is obvious that this change in the polymer's architecture was necessary to enable life to store larger genomes. Therefore, if functional pressures gave rise to DNA from RNA, then we should not find it hard to accept that an earlier polymer, or set of polymers, predated RNA.

In conclusion, as we stated at the beginning of this essay, there have certainly been impressive advances in the quest to form RNA through model prebiotic reactions. Nevertheless, declaring RNA a plausible prebiotic molecule might be premature given the chemical challenges that face the spontaneous formation of RNA, as well as the numerous reasons that each component of RNA looks like a product of evolution. Thus, we feel that the search for feasible pre-RNA polymers that assemble through robust and efficient prebiotic reactions is a necessary and potentially fruitful component of the overall search for the origin of nucleic acids and the origin of life.

1 Powner, M. W., Gerland, B. & Sutherland, J. D. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* **459**, 239-242 (2009).

2 Oró, J. Synthesis of adenine from ammonium cyanide. *Biochem. Biophys. Res. Comm.* **2**, 407-412 (1960).

3 Saladino, R., Crestini, C., Costanzo, G., Negri, R. & Di Mauro, E. A possible prebiotic synthesis of purine, adenine, cytosine, and 4(3H)-pyrimidinone from formamide: Implications for the origin of life. *Bioorganic & Medicinal Chemistry* **9**, 1249-1253 (2001).

4 Barks, H. L. *et al.* Adenine, and hypoxanthine production in UV-irradiated formamide solutions: relaxation of the requirements for prebiotic purine nucleobase formation. *ChemBioChem* **11**, 1240–1243 (2010).

5 Decker, P., Schweer, H. & Pohlmann, R. Bioids .10. Identification of Formose Sugars, Presumable Prebiotic Metabolites, Using Capillary Gas-Chromatography Gas-

Chromatography Mass-Spectrometry of Normal-Butoxime Trifluoroacetates on Ov-225. *Journal of Chromatography* **244**, 281-291 (1982).

6 Kim, H. J. *et al.* Synthesis of carbohydrates in mineral-guided prebiotic cycles. *J. Am. Chem. Soc.* **133**, 9457-9468, doi:10.1021/ja201769f (2011).

7 Engelhart, A. E. & Hud, N. V. Primitive genetic polymers. *Cold Spring Harbor Perspect. Biol.* **2**, 10.1101/cshperspect.a002196 (2010).

8 Hud, N. V., Jain, S. S., Li, X. & Lynn, D. G. Addressing the Problems of Base Pairing and Strand Cyclization in Template-Directed Synthesis. *Chemistry & Biodiversity* **4**, 768-783 (2007).

9 Cafferty, B. J. *et al.* Efficient self-assembly in water of long noncovalent polymers by nucleobase analogues. *J. Am. Chem. Soc.* **135**, 2447-2450, doi:10.1021/ja312155v (2013).

10 Eschenmoser, A. The search for the chemistry of life's origin. *Tetrahedron* **63**, 12821-12844 (2007).

11 Yu, H. Y., Zhang, S. & Chaput, J. C. Darwinian evolution of an alternative genetic system provides support for TNA as an RNA progenitor. *Nat. Chem.* **4**, 183-187, doi:10.1038/nchem.1241 (2012).

12 Mittapalli, G. K. *et al.* Mapping the landscape of potentially primordial informational oligomers: oligodipeptides and oligodipeptoids tagged with triazines as recognition elements. *Angew. Chem., Int. Ed. Engl.* **46**, 2470-2477 (2007).

13 Pasek, M. A., Harnmeijer, J. P., Buick, R., Gulla, M. & Atlasa, Z. Evidence for reactive reduced phosphorus species in the early Archean ocean. *Proc Natl. Acad. Sci. USA* **110**, 10089-10094 (2013).

14 Bean, H. D., Anet, F. A. L., Gould, I. R. & Hud, N. V. Glyoxylate as a backbone linkage for a prebiotic ancestor of RNA. *Origins Life Evol. Biospheres* **36**, 39-63 (2006).