The Hitchhiker's Guide to the RNA World

Dina Anastas

Biology, University of British Columbia
Submitted December 2002

RNA World Hypothesis

The quest to reach a plausible hypothesis concerning the origin of life is ongoing. Today, we have conceptualized many workable hypotheses; however, a vigorous controversy has evolved over the past years concerning the nature of the origin of life. We'll start first with theory of the "RNA world" that states that the first molecular systems to display the properties of self-replication and evolution were RNA molecules (Schwartz, 1993). Schwartz proposes that RNA molecules are in fact the first "living" things. What qualifies as "living" in this context is any system which is able to replicate itself and, in doing so, continually adapt by means of selection to a changing environment. Moreover, Bartel and Unrau, in their article "Constructing an RNA world" explain the popular theory of life's origin that also states that the first biocatalysts were not made of protein but were made of RNA or a very similar polymer (1999). Many scientists have accepted this theory as the most satisfactory explanation for the origin of life. According to this theory, RNA first promoted the reaction required for life with the help of metals, pyridines, amino acids and other small-molecule cofactors. Later, these RNA molecules developed various abilities to synthesize coded polypeptides that served as more sophisticated cofactors. Eventually, RNA was replaced by DNA as the genetic polymer, and by proteins as the prominent biocatalysts. However, during the history of the scientific study of the origin of life, different models have attracted attention. These different models centre their basis on the origin of life using different theories: protein-centred, nucleic acid-centred, and other "precursor" hypothesis based models.

The origin of life theory is seen as the chicken and the egg paradox; is it an RNA world or a protein world? The reason is that RNA can store information and even act as a template for synthesis of a copy of itself, but cannot alone carry out the synthesis, for which proteins (enzymes) are needed. Proteins, on the other hand, act as catalysts for the synthesis of new RNA as well as for new proteins, but cannot act as templates (Schwartz, 1993). We'll look at some of these hypotheses of the "RNA World", and hopefully we'll be able to answer some of the related questions:

· If RNA (nucleic acid) is the key player in this scenario, then can RNA catalyze the reactions needed for self-replication on the early earth?
· Can RNA-based life achieve the metabolic sophistication needed to give birth to the protein-nucleic acid world?

For Darwinian evolution to take place, a faithful replication of information is necessary. Both the nucleic acid-centered and protein-centered theories fail to provide grounds for such a requirement. It is difficult to imagine evolution beginning with a set of autocatalytic RNA molecules. Rather, it is more convenient to point out that the first self-replicating RNA molecules might have been preceded in evolution by another population of molecules which possessed the same kind of base-pairing possibilities as RNA, but built up around a simpler structural unit than ribose (Joyce et al., 1987).

The reasoning behind this is the difficulty of synthesis of nucleic acid as well as the ribose molecule itself. However, it was found that all carbohydrates can be made or synthesized by reacting formaldehyde with itself via a base catalyzed reaction called an *aldol condensation*. In Schwarts's article, the reaction is represented by the following formula:

$$n\text{Ch2O} \rightarrow \text{C}n\text{H2}n\text{O}n$$
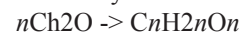
Still, in an era without enzymes, the complexity

with which the nucleosides are synthesized makes RNA synthesis somewhat difficult due to the chirality problem (the structural characteristic of a molecule that makes it impossible to superimpose it on its mirror image) that arises with the synthesis of sugars. From here emerges the theory of ever-possible-existing prebiotic molecules that lacked this chirality phenomenon. On a closer examination, glycerolbisphosphate has shown to react on a poly (C) template to form long pyrophosphate-linked chains. The prochiral nature (lacks chiral activity) of this monomer, and the flexibility of its backbone makes it easy and possible to carry out template-directed oligomerizations without having to face the problem of "enantiomeric cross-inhibition" (Schwartz, 1993). The problem still resides in the low catalytic activity in this type of oligomerization.

RNA has enzymatic activities

Back to the "RNA World" hypothesis that states that at the beginning of life there were only RNA molecules with a catalytic capacity, usually referred to as "Ribozymes", and that proteins arose much later. Indeed, the discovery of such activities made the inference of the RNA World more feasible. These ribozymes can catalyze the synthesis of a new RNA molecule, using a template RNA and a precursor (Gilbert, 1986), thus rationalizing the un-necessity behind the existence of proteins or enzymes at the beginning. Two of the ribozymes perform self-splicing reactions, four perform self-cleavage and one trims off the 5' end of the pre-tRNA. Despite the wide range of reactions performed by those different ribozymes, still the sophistication needed for RNA World is not met. These molecules evolve by self-splicing and mutation to explore new functions and to adapt to new environment. Moreover, RNA molecules, along with RNA cofactors such as: nicotinamide adenine dinucleotide and flavin mononucleotide, combine to encompass a wide range of enzymatic activities.

Gilbert, in his article, explains the self-splicing intron element that is found in the RNA molecule. This self-splicing reaction is reversible. So the intron can splice itself back into any nucleotide sequence in the RNA. This is quite phenomenal and significant, regarding the incapability of an RNA molecule to undergo recombination. So what have been suggested here is that any intron can splice itself in and out of an RNA molecule while it is being synthesized. The introns, here, can play the role of a transposon. Two introns flanking an exon can splice out as one unit and re-insert themselves in another sequence in the same molecule. This technique provides RNA with an evolutionary ability that it lacks_recombination. In a world like this, the structure that is replicated has the full complement of introns; whereas, the daughter

molecules that splice out some of their introns, function as ribozymes (Gilbert, 1986).

Indeed, it has been shown that some RNAs promote the chemistry of polymerization. These are exemplified by a ribozyme that ligates RNA efficiently (Kcat>1sec) using a reaction similar to the addition of a single nucleotide during RNA polymerization (Bartel et al., 1991). An important key here is that the efficiency of this process must be sufficient to produce "progeny" RNA molecules at a rate that exceeds the rate of decomposition of the parents (Gesteland, 1998). Today, we have three different ribozymes; each catalyzing a different reaction. One carries out the catalysis of the proper chemistry, a second uses nucleoside triphophates in a templeted fashion, and the third recognizes an RNA duplex without regard for sequence (Bartel et al., 1991). Bartel explains that these features have to be found in a single ribozyme. Yet there is still no example of a ribozyme with sufficient significance and efficiency that it could carry out an activity presumed by the RNA World hypothesis. So the polymerase activity is still far from that needed for self-replication.

We see RNA molecules evolving through various modifications. Many associate the nucleosides modification with the passage from the RNA World into the DNA-Protein World. This is best explained by the replacement of ribozymes by protein-based enzymes. At this stage, RNA molecules began to synthesize proteins, first by developing RNA adapter molecules that can bind activated amino acids and then by arranging them according to an RNA template using other RNA molecules such as the RNA core of the ribosome (Gilbert, 1986). This ultimately created proteins, which simply happened to be better enzymes than RNA. Moreover, it is presumed that ribosome and tRNA evolved from "a pre existing function in the RNA stage of life"(Gimenez, 1998), and that later they adapted for protein synthesis. These early or primitive ribosomes have been referred to as protoribosome. And their function was somewhat similar to that of primitive genetic translation. Many have suggested that the presence of modified nucleosides have increased the catalytic activities of the ribozymes. The origin of protein synthesis is quite complex. I It is more likely that translation evolved by the modification of a pre-existing ribozyme-catalyzed function in the RNA World.

Before we discuss the reasons behind the transition to the DNA World, we need to explain that DNA synthesis cannot proceed without An RNA primer, and that deoxyribonucleotides are synthesized, by ribonucleotide reductase, also argue for the evolutionary transition from RNA to DNA. Another thing to note is that the highly complex nature of the reduction mentioned above suggests that catalytic proteins had

to arise before the transition to DNA could occur, and that RNA ribozymes could not have carried out such free radical chemistry (Poole et al. 2000).

Why the transition from RNA to DNA?

One of the apparent differences between RNA and DNA is the presence of a hydroxyl group at the 2' position of the ribose in RNA. We also see various tertiary structures formed by RNA, not seen in DNA. The 2'-hydroxyl, we just mentioned, is crucial for the folding of the molecule. Two major, worth mentioning, reasons for the transition from RNA to DNA as the coding genetics material are: first, RNA's inability to retain large amounts of genetic information, and second, RNA's low stability as a molecule.

It was mentioned at the beginning of this paper that a faithful replication of information is necessary for Darwinian evolution. It is known that DNA can retain much larger quantities of genetic information than RNA, and this is probably why the former was selected over the latter during evolution. That's because the accuracy of copying information places a limit on a genome's coding capacity. If accuracy is high, large amounts of information can be maintained; if accuracy is low or poor, then the amount of information maintained is low. Poole et al. describes the coding capacity by the "error threshold" or the "Eigen limit" (2000). This term is described in terms of the overall fidelity of replication. Although many 'relics' of an RNA World exist, the largest RNA organisms are viruses. In fact, Coronaviruses have the largest viral RNA genomes identified, specifically reaching 30 Kb (Atkins, 1993). It has been assumed by many that the largest modern RNA viruses are indicative of the upper size limit for an RNA world genome. Even with all the gaps in our world theory, many have suggested that the RNA organisms that evolved protein synthesis was likely to have been dangerously close to the error threshold. Those proteins in turn have improved the RNA-catalyzed reactions which allowed the increase in genome coding capacity.

In short, without proof-reading and repair, it seems highly unlikely that an RNA genome could have been large enough to support such a large collection of protein genes, making hard to see how ribonucleotide reductase could have arisen in an RNA world. From a stability standpoint, RNA molecules are highly unstable when compared to DNA. The 2'-hydroxyl on the ribose acts as a nucleophile. It can attack the adjacent 3'-phosphate, thus breaking the phosphodiester bond. This inevitably leads to self-cleavage. Having said this, many support the possibility of an ever-existing intermediate genetic material in the transition from the RNA to the DNA World. Then again, no such material has been identified.

Many authors have suggested the existence of some modified RNA molecules that played a role in the transition from RNA to DNA. One type of modification is the 2'-O-Methylation of RNA. It is not known yet whether this reaction has arisen before or after the emerging of proteins. The advantage behind the methylation is: first, silencing the 2'-OH on the ribose, thus preventing unwanted side reactions, and that's done by making sure that only specific residues are catalytic, and second, preventing 2'-OH from forming H-bonds in the tertiary structure, ultimately favoring a particular folding pathway (Poole et al., 2000). This transition has provided the genome some form of stability. In addition to stability, the 2'-OCH3 confers qualities similar to DNA as a genetic material. On the other hand, those methylated ribose molecules would produce a hydrophobic cushion in the deep groove of the helix; thus compromising genome functionally (Poole et al., 2000). An advantage for this extra step in evolution: mainly the RNA -> methyl-RNA -> DNA, is that it does not require extremely and unlikely steps. One might argue against such a hypothesis or model, but a similar argument for an intermediate stage has been used for the origins of the RNA World itself..

In summary, the unanticipated discovery of ribozymes initiated the hypothesis of the role of RNA in the origins of life. This has led to the belief of an "RNA World" which has preceded any form of life. Joyce and Orgel, in their article, include the three basic assumptions of all the RNA World hypotheses: (1) "At some time in evolution of life, genetic continuity was assured by the replication of RNA; (2) Watson-Crick base-pairing was the key to replication; (3) genetically encoded proteins were not involved as catalysts" (The RNA World, 1998). Despite the more versatile catalysis nature of proteins and their displacement of RNA, the conversion is not considered complete. RNA retains a central role in protein synthesis, perhaps including catalysis of peptidyl transfer.

So to conclude, the RNA World theme is still under considerable amount of construction. The generating of more RNA activities will hopefully attract attention to this model. Yet In the mean while, most professional advocates of an RNA world are doubtful that life began with RNA per se. Instead, they propose that life began with an RNA-like polymer, yet to be identified, that possessed the catalytic and templating features of RNA but lacked RNA's undesirable traits. The era of this RNA-like polymer is the "pre-RNA world", which presumably gave rise to the RNA world in a manner analogous to that in which the RNA world gave rise to the protein-nucleic acid world of today.

References

Bartel, David P., and Unrau, Peter J. "Constructing an RNA World". Trends in Cell Biology. Vol. 9:12. (1991).

Gesteland, raymond F. The RNA world. 2nd edition. 1993

Gilbert, Walter. "The RNA World". Nature. Vol. 319. (1986).

Martinez, Juan A., Saez, Guillermo T., and Seisdedos, Rafael Tabares. "On the Function of Modified nucleosides in the RNA World". Journal of theoretical Biology. Vol 194:4. (1998).

Moss, Eric G. "RNA Interference: It's a small World". Current biology. Vol 11: 19. (2001)

Poole, Anthony; Penny, David and Sjoberg, Britt-Marie. "Methyl-RNA: an evolutionary bridge between RNA and DNA?" Chemistry and Biology. Vol 7: 12. (2000)

Schwartz, Alan W. "The RNA World and its origins". Planet and Space Science. Vol. 43: 1-2. (1995).