

# THE RNA WORLD

## PIECING TOGETHER THE HISTORICAL DEVELOPMENT OF A HYPOTHESIS

ANTONIO LAZCANO

The concept of an RNA world is a hypothesis firmly rooted in empirical data and is part of a long and complex scientific perspective that goes back more than fifty years to the discovery of the central role RNA and ribonucleotides play in protein synthesis and biochemical reactions took place. As the understanding of RNA biology progressed, several independent proposals of protein-free primordial life forms were suggested. Although this possibility was strongly reinforced with the discovery of ribozymes, there are many definitions of the RNA world, including several contradictory ones. One could say that it was an early, perhaps primordial, stage during which RNA molecules played a much more conspicuous role in heredity and metabolism and, particularly, in the origin and early evolution of protein biosynthesis. The overwhelming evidence for the catalytic, regulatory, and structural properties of RNA molecules, combined with their ubiquity in cellular processes, can only be explained with the proposal that they played a key role in early evolution and perhaps in the origin of life itself.

Keywords: origin of life, RNA world, ribonucleotidyl coenzymes, catalytic RNA.

Although it is not possible to assign a precise chronology to the appearance of life, it is generally agreed that it emerged on Earth soon after the formation of our planet. Proper understanding of the processes that led to the emergence of the first life forms has also been hindered, not only by the incompleteness of the sedimentary record, but also by the absence of fossils from prebiological systems that may have preceded the first living entities, and by the lack of an all-embracing generally-agreed definition of life.

In spite of the seemingly insurmountable obstacles surrounding the understanding of the origin of life, or perhaps because of them, there has been no shortage of discussion about how it came about. To a considerable extent our current ideas on the origin of life are based on the pioneering work of Alexander I. Oparin, who in November 1923 published a small book that attempted to explain the emergence of the first living entities from a

Darwinian perspective. He suggested that life had been preceded by a lengthy period of abiotic syntheses and accumulation of organic compounds that had led to the accumulation of what we call today the primitive soup. Oparin was well acquainted with the ideas of Ernst Haeckel and other nineteenth century scientists that assumed that the first forms of life had been autotrophic microbes. However, it was impossible for him to accept that life had emerged already endowed with an autotrophic metabolism based on enzymes, pigments like chlorophyll, and the ability to synthesize organic compounds from CO<sub>2</sub> and water.

As a young student at the University of Moscow, Oparin had joined the laboratory of Alexei N. Bakh, a distinguished scientist and political figure at the Karpov Physicochemical Institute, where he acquired a solid background in biochemistry and quickly adopted the idea that metabolism was the outcome of oxidation

**«IN SPITE OF THE SEEMINGLY  
INSURMOUNTABLE  
OBSTACLES SURROUNDING  
THE UNDERSTANDING OF  
THE ORIGIN OF LIFE, THERE  
HAS BEEN NO SHORTAGE OF  
DISCUSSION ABOUT HOW IT  
CAME ABOUT»**

and reduction reactions that were coupled inside cells. Oparin was also a convinced evolutionist. As an undergraduate, he had attended the lectures given regularly by Kliment A. Tymiriazev, a renowned plant physiologist, agronomist and the main advocate of Darwinism in Russia. Tymiriazev left the university in 1911 in the protest against the Czarist government, but kept discussing scientific issues with students and colleagues in his Moscow apartment. By the time Oparin graduated, he had an academic background that combined natural history, biochemistry, and plant physiology, a knowledge acquired within a research tradition strongly committed to integral approaches in the analysis of natural phenomena. He was not only familiar with nearly all the literature on evolution available in Russia but also, perhaps even more importantly, with those related to the Darwinian method of comparative analysis and historical interpretation of the features of natural history (Lazcano, 2012).

Since a heterotrophic anaerobe is metabolically simpler than an autotrophic one, Oparin argued that the former would necessarily have evolved first. Thus, based on the simplicity and ubiquity of fermentative reactions, he proposed that the first organisms must have been heterotrophic bacteria that could not make their own food but obtained organic material already present in the primitive milieu. Like Oparin, the geochemist Charles Lipman, the microbiologist Roger B. Harvey, and the geneticist and polymath John B. S. Haldane also advocated the heterotrophic emergence of life (Bada & Lazcano, 2003). The discovery of phages led Haldane to suggest that viruses could represent an intermediate step in the transition from the prebiotic soup to the first heterotrophic cells. Life, wrote Haldane, could have remained «in the virus stage for many millions of years before a suitable assemblage of elementary units was brought together in the first cell».

Oparin rejected this possibility as reductionist and, as reviewed elsewhere (Lazcano, 2012), in 1936 published a revised version of his original proposal, also titled *The origin of life*, that was translated into English a couple of years later (Oparin, 1938). This book is a masterpiece of evolutionary analysis, and in it Oparin critically revised his original proposal.

Based on additional astronomical data, an extensive review of organic chemistry, and the available data on cell biology and microbial metabolism, in this new book he suggested that the primitive Earth was endowed with a highly reducing environment in which iron carbides of geological origin would react with steam to form hydrocarbons. Their oxidation would yield alcohols, ketones, aldehydes and other reactive compounds that would then react with ammonia to form amines, amides, and ammonium salts. The resulting protein-like compounds and other molecules would form a dilute solution, where they would aggregate to form colloidal systems from which the first heterotrophic microbes evolved. Following H. G. Bungenberg de Jong's proposal that the colloid properties of droplets, which he termed coacervates, formed by the spontaneous aggregation of biological macromolecules could explain the

properties of the protoplasm, Oparin proposed them as precursors of the first cells. Like many of his contemporaries, Oparin's original genetics were pre-Mendelian, and he assumed that the biological inheritance was the outcome of growth and division of the coacervate drops. For him, enzymatic-based assimilation, growth, and reproduction represented the basic traits of life, but the list did not include nucleic acids, whose role as the material

basis of inheritance was not even suspected when he published his second book.

Today the situation has changed. The discovery of ribozymes has given considerable support to the hypothesis of the so called-RNA world, from which the nucleic acid-based genetic system of extant life is believed to have originated. It is of course impossible to demonstrate that this is the evolutionary pathway that led to the origin and early evolution of life. However, the available evidence from widely different scientific fields is consistent with the possibility that it happened this way.

The purpose of this short essay is to discuss the discoveries and the history of the ideas that led to the RNA world hypothesis. As summarized here, as the understanding of the properties of RNA molecules progressed, the evolutionary interpretation of their widespread distribution in intracellular environments, as well as the catalytic properties of nucleotide coenzymes and the participation of RNA monomers

**«SINCE A HETEROTROPHIC ANAEROBE IS METABOLICALLY SIMPLER THAN AN AUTOTROPHIC ONE, OPARIN ARGUED THAT THE FORMER WOULD NECESSARILY HAVE EVOLVED FIRST»**



From Oparin. Dedicated to the 120th anniversary of the birth (2014). Moscow: Geos.

Andrei N. Belozerskii presented, in the now famous 1957 Moscow meeting on the origin of life, an insightful and pioneering analysis of the role of nucleotides and RNA during the early stages of biological evolution. In the picture, Alexander I. Oparin (1894–1980), on the left, accompanied by Belozerskii (1905–1972) at the Department of Plant Biochemistry at Moscow State University in 1964.

in metabolic pathways, led to several independent proposals of protein-free primordial life forms. As argued here, current ideas and discussions about the RNA world are part of a long and complex scientific perspective whose historical analysis still has many lessons to teach us.

#### ■ BIOCHEMICAL ROOTS OF THE RNA WORLD CONCEPT

The demonstration that the tobacco mosaic virus (TMV) has RNA, which was followed by the confirmation in the late 1950s that TMV's infectivity resides in its RNA, led many to recognize that genetic information could be stored in RNA. Although many rejected the possibility that viruses represented a missing link between the non-living and the living worlds, by the time the now famous 1957 Moscow meeting on the origin of life took place, John Burdon Sanderson Haldane, Oparin, John Desmond Bernal, and many others were convinced that RNA had preceded DNA as genetic material during early cell evolution. Indeed, careful reading of the proceedings of the Moscow meeting (Oparin, Pasynskii,

Braunshetin, & Pavloskaya, 1959) shows the influence of evolutionary biochemistry and of molecular biology in the shift from a protein-based molecular view of life to an increased awareness of the role of RNA and ribonucleotides during early stages of biological evolution.

This is particularly true of the contributions by the Belgian embryologist Jean A. Brachet, and the Soviet microbiologist Andrei Nikolaevich Belozerskii, who both had a deep understanding of RNA biology. Their independent analyses of the main traits of metabolism allowed them to develop novel insights on the role of ribonucleotides in cell evolution. Brachet speculated that RNA, whose structure was assumed to be much simpler than that of proteins, could have formed by a non-enzymatic polymerization processes. «But all these data, are indeed quite weak and do not answer in full the question of “What came first, RNA or proteins”», argued Brachet, and concluded that «the collaboration of chemists, physicists and theoreticians may demonstrate if the synthesis via an spontaneous polymerization of a nucleic acid is more likely than that of a protein» (Brachet, 1959).

Belozerskii went further.

Together with his associates he had demonstrated the occurrence of both DNA and RNA in plants and bacteria, had analyzed the nucleic acid content in a wide range of bacterial taxonomic groups, and had measured the levels of RNA and DNA in prokaryotes at different stages of activity and proliferation. As reviewed elsewhere (Lazcano, 2012), Belozerskii presented an insightful and pioneering analysis of the role of nucleotides

and RNA during the early stages of biological evolution at the Moscow meeting. As he wrote:

There is no shadow of doubt about the fact that nucleic acids played a significant part in the evolution of the organic world. However, it is doubtful that both RNA and DNA appeared simultaneously at the early stages of the development of life. The author of the present communication is inclined to think that the appearance of ribonucleotides, followed by RNA, was primary. The DNA appeared much later, parallel with more complicated functions and an ever-increasing differentiation of protoplasm.

(Belozerskii, 1959)

After discussing the genetic role of RNA, its direct involvement in protein synthesis, its wide distribution

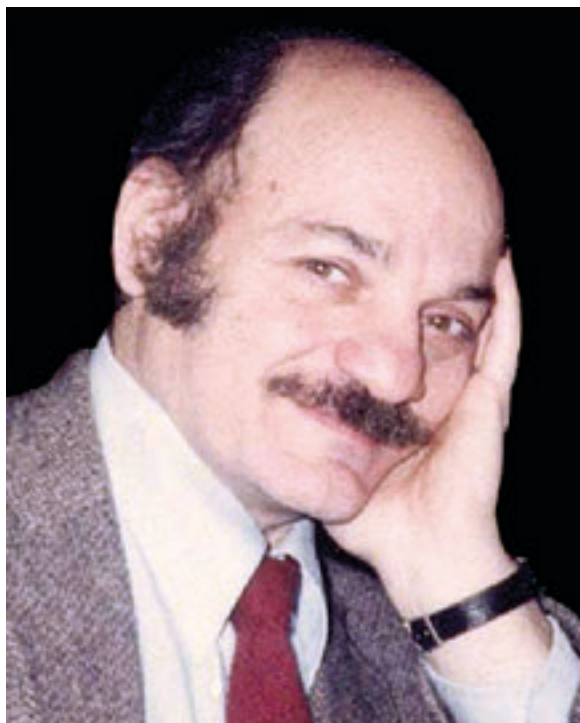
**«CURRENT IDEAS AND DISCUSSIONS ABOUT THE RNA WORLD ARE PART OF A LONG AND COMPLEX SCIENTIFIC PERSPECTIVE WHOSE HISTORICAL ANALYSIS STILL HAS MANY LESSONS TO TEACH US»**

in cells, and the participation of ribonucleotides in metabolic pathways, Belozerskii added that «the very fact of greater specialization and differentiation appears to lend support to the belief that DNA had originated later than RNA», and concluded that «[It] seems that RNA, being associated with the most general processes of life, was formed at an earlier evolutionary stage, while the origin of DNA was associated with the development of more specialized and phylogenetically later features of organisms» (Lazcano, 2012).

The times were ripe for a novel interpretation of biochemical data to understand the early evolution of life. Like Belozerskii and Brachet before them, in the USA both Philip Handler (1963) and Robert E. Eakin (1963) were able to pull together biochemical and metabolic data, to independently interpret the evolutionary significance of the reports on the distribution and catalytic abilities of coenzymes, and to conclude that they were possible vestiges of an earlier stage in cell evolution. Both Handler and Eakin proposed that the earliest catalysts were ribonucleotide cofactors. A dozen years later Harold B. White III (1976) argued that the presence of a ribonucleotidyl moiety among many cofactors suggested that they are the evolutionary remnants of early epochs, prior to the development of protein synthesis. Although he stopped short of the idea of an RNA world, White argued that the wide distribution of ribonucleotidyl coenzymes reflected primitive recruitment processes that had diversified the catalytic abilities of RNAs. He also proposed that the peculiarities of histidine biosynthesis, which is the only amino acid whose biosynthesis starts with ATP, together with the nucleotide coenzymes, could be understood as a biochemical fossil of a primitive RNA-based metabolic apparatus.

#### ■ THE BIRTH OF THE «RNA WORLD» CONCEPT

It was not until the 1960s, however, when Alexander Rich (1962), Carl R. Woese (1967), Francis H. C. Crick (1968), and Leslie Orgel (1968) explicitly suggested that the first living entities were devoid both of DNA and proteins and were based on RNA both as a hereditary material and a catalyst. Like many of their contemporaries, the four of them took



MÉTODOE

Alexander Rich (1924–2015) was the first to suggest that life started with molecules similar to RNA.

**«BRACHET SPECULATED THAT RNA, WHOSE STRUCTURE WAS ASSUMED TO BE MUCH SIMPLER THAN THAT OF PROTEINS, COULD HAVE FORMED BY A NON-ENZYMATIC POLYMERIZATION PROCESSES»**

the existence of a prebiotic soup rich in amino acids, sugars, nucleotides, and different polymers for granted, but their ideas reflect the hegemony of nucleic acid studies over biochemical approaches that has shaped mainstream origin-of-life research ever since.

To the best of my knowledge, Rich (1962) was the first to propose that life had started with RNA-like molecules. Based on the scant evidence available at the time, he wrote that «there are significant stereochemical reasons why the polynucleotides can act as their own catalysts for self-replication. However, there are no analogous reasons for believing the poly-amino acids have this ability to reproduce themselves,» and added that:

The primitive polynucleotide chains are able to act as a template or a somewhat inefficient catalyst for promoting the polymerization of the complementary nucleotide residues to build up an initial two-stranded molecule... It may be reasonable to speculate that the hypothetical stem or parent polynucleotide molecule was initially an RNA-like polymer...

(Rich, 1962)



Courtesy of Alice Orgel

The work by Francis H. C. Crick (1916–2004), on the right of the picture, and Leslie E. Orgel (1927–2007), on the left, should be read as the basis of what became a major conceptual revolution that has replaced proteins and DNA as the main actors of the processes that led to the origin of life.

Carl R. Woese proposed somewhat similar views in his 1967 book, where he asserted that neither the exquisite structures of ribosomes or aminoacyl-tRNA synthetases could have existed in primordial times (Woese, 1967). He assumed that the structure of primitive tRNA-like molecules allowed them to bring together activated amino acids, which would then react and give rise to small peptides. The only RNA whose structure had been determined at the time was that of tRNA, and its complexity had led several researchers to speculate on the catalytic potential of RNA. This is shown, for instance, in Crick's 1968 comment about the role of tRNA:

An attractive idea (suggested to us by Dr. Oliver Smithies) is that primitive tRNA was its own activating enzyme. That is, that its structure had a cavity in it which specifically held the side chain of the appropriate amino acid in such a position that the carboxyl group could be easily joined on the terminal ribose of the tRNA.

(Crick, 1968)

While Crick did not address the issue of the origin of ribonucleotidyl coenzymes, Rich (1962), Woese (1967), and Orgel (1968) also noted that a significant number of the most important coenzymes such as NAD involve an organic catalyst bound via a pyrophosphate link to a nucleotide, raising again the possibility that they are fossils of an epoch in which metabolism depended on RNA. Orgel's (1968)

analysis of the widespread presence of nucleotide coenzymes in metabolic pathways (Orgel, 1968; Orgel & Sulston, 1971) recognizes that genetic replication and metabolism were not different properties waiting to be conjoined, but presupposes the deep intertwining of genetic material and biochemical catalysis. As Orgel (1968) wrote «Can polynucleotide chains with well-defined secondary structures act as primitive enzymes? I doubt they alone could exhibit extensive catalytic activity, although one cannot be quite sure» adding that:

It seems to be quite possible that polynucleotide chains could make a primitive selection among organic molecules such as amino acids, by forming stereospecific complexes stabilized by hydrogen-bonding and hydrophobic interactions. This will be important in any discussion of the evolution of the genetic code. Even more speculative questions can be asked. Could a polynucleotide attached by base-pairing to a diphosphopyridine nucleotide act as a primitive catalyst for dehydrogenation?

(Orgel, 1968)

The main ideas behind the hypotheses developed independently by Woese (1967), Crick (1968), and Orgel (1968) were (a) that the original ribosomes may have been

made solely of RNA and lacked proteins, and (b) that RNA, besides acting as a template, might also behave like an enzyme and would thus be able to catalyze its own replication. As argued elsewhere (Lazcano, 2012) the most refined schemes are those of Crick and Orgel, which should be read as the basis of what became a major conceptual revolution that has replaced proteins and DNA as the main players of the processes that led to the origin of life. In spite of their different emphasis, both essays are complementary publications, while Crick (1968) attempted to understand the origin of the amino acid assignments of the code, and explicitly assumed that both ribosomal RNA and transfer RNAs are the oldest components of the translation machinery, which, as he wrote, «had no protein at all and consisted entirely of RNA,» adding quite explicitly that «the first enzyme was an RNA molecule with replicase properties.»

#### ■ CATALYTIC RNA AND THE ORIGIN OF LIFE: PENDING ISSUES

During the twenty years that followed the first suggestions of an RNA world, (a) many accepted



Photograph by Fernando Rincón

Rebeca Plana. *There, there*, 2014. Mixed technique on linen, 150×190 cm.

the idea that cellular RNA genomes were older than DNA genomes; (b) the possibility that coenzymes were vestiges of an RNA world stage was acknowledged, but few worked on it; (c) some accepted White's suggestion that histidine was a molecular vestige of an RNA world stage, but with the exception of Marie-Christine Maurel and Jaques Ninio (1987) it was not addressed experimentally; and (d) no one was really searching for catalytic RNA molecules.

In fact, the discovery of catalytic RNA took everyone by surprise. The term «the RNA world» coined by Gilbert (1986) rapidly permeated scientific literature and became a unifying term. However, there are many definitions of the RNA world, including several contradictory ones. One could say that it is an early, perhaps primordial, stage during which RNA molecules played a much more conspicuous role in heredity and metabolism and, particularly, in the origin and early evolution of protein biosynthesis. Indeed, extant natural ribozymes are mostly involved in RNA processing and in protein biosynthesis. As underlined by many, the ribosome is a ribozyme. Both the catalytic and structural properties of the ribosome strongly suggest that protein synthesis first evolved in an RNA world, which implies that the origin of the genetic code and the origin of life are not necessarily synonymous.

As argued here, the hypothesis of an early RNA world is firmly rooted in empirical data and is part of a long and complex scientific perspective that goes back more than fifty years to when the discovery of the central role of RNA and ribonucleotides in protein synthesis and biochemical reactions first took place. As described here, this proposal did not result from a continuous scientific tradition, but one that has been broken up over time due to the demise of certain areas of research like coenzyme chemistry and the lack of an evolutionary framework to explain the discoveries in molecular biology. The overwhelming evidence on the catalytic, regulatory, and structural properties of RNA molecules, combined with their ubiquity in cellular processes, can only be explained by the proposal that they played a key role in early evolution and perhaps in the origin of life itself. ☺

**«THE HYPOTHESIS OF AN  
EARLY RNA WORLD IS FIRMLY  
ROOTED IN EMPIRICAL DATA  
AND IS PART OF A LONG  
AND COMPLEX SCIENTIFIC  
PERSPECTIVE THAT GOES  
BACK MORE THAN FIFTY  
YEARS»**

## ACKNOWLEDGEMENTS

I am indebted to the staff of the Mandeville Special Collections in the Geisel Library, University of California San Diego campus, who have facilitated the consultation of the archives of S. L. Miller, Leslie Orgel, and Francis Crick. This paper is based on previously published work (Lazcano, 2012). The assistance of Sara Islas and Ricardo Hernandez Morales in preparing the manuscript is gratefully acknowledged.

## REFERENCES

- Bada, J. L., & Lazcano, A. (2003). Prebiotic soup: Revisiting the Miller experiment. *Science*, *300*, 745–746. doi: 10.1126/science.1085145
- Belozerskii A. N. (1959). On species specificity of nucleic acids in bacteria. In A. I. Oparin, A. G. Pasynskii, A. E. Braunschtein, & T. E. Pavloskaya (Eds.), *The origin of life on Earth* (pp. 322–331). New York: Pergamon Press/MacMillan Company.
- Brachet, J. (1959). Les acides nucléiques et l'origine des protéines. In A. I. Oparin, A. G. Pasynskii, A. E. Braunschtein, & T. E. Pavloskaya (Eds.), *The origin of life on Earth* (pp. 361–367). New York: Pergamon Press/MacMillan Company.
- Crick F. H. C. (1968). The origin of the genetic code. *Journal of Molecular Biology*, *39*, 367–379.
- Eakin, R. E. (1963). An approach to the evolution of metabolism. *Proceedings of the National Academy of Sciences*, *49*, 360–366.
- Gilbert, W. (1986). The RNA World. *Nature*, *319*, 618. doi: 10.1038/319618a0
- Handler, P. (1963). Evolution of the coenzymes. In A. I. Oparin (Ed.), *Proceedings of the Fifth International Congress of Biochemistry, Vol. III. Biochemistry* (pp. 149–157). New York: Pergamon Press/Macmillan Company.
- Lazcano, A. (2012). The biochemical roots of the RNA world: From zymonucleic acid to ribozymes. *History and Philosophy of the Life Sciences*, *34*, 407–424.
- Maurel, M. C., & Ninio, J. (1987). Catalysis by a prebiotic nucleotide analog of histidine. *Biochimie*, *69*, 551–553.
- Oparin, A. I. (1938). *The origin of life*. New York: MacMillan.
- Oparin, A. I., Pasynskii, A. G., Braunschtein, A. E., & Pavloskaya, T. E. (Eds.). (1959). *The origin of life on Earth*, New York: Pergamon Press/MacMillan Company.
- Orgel, L. E. (1968). Evolution of the genetic apparatus. *Journal of Molecular Biology*, *38*, 381–393.
- Orgel, L. E., & Sulston, J.E. (1971). Polynucleotide replication and the origin of life. In A. P. Kimball, & J. Oró (Eds.), *Prebiotic and biochemical evolution* (pp. 89–94). Amsterdam: North-Holland.
- Rich, A. (1962). On the problems of evolution and biochemical information transfer. In M. Kasha, & B. Pullman (Eds.), *Horizons in biochemistry* (pp. 103–126). New York: Academic Press.
- White III, H. B. (1976). Coenzymes as fossils of an earlier metabolic state. *Journal of Molecular Evolution*, *7*, 101–104.
- Woese, C. R. (1967). *The Genetic Code: The molecular basis for gene expression*. New York: Harper and Row.

**Antonio Lazcano**. Full Professor of the origin of life at the National Autonomous University of Mexico. He is the author of scientific papers and books on the early stages of the evolution of life and has a long career in science communication in written media, radio, and television. He is the director of the Lynn Margulis Galapagos Centre. He was the President of the ISSOL twice and was made an Honorary Doctor by the universities of Milan, Valencia, and Michoacán. Since 2014, he has been a member of El Colegio Nacional, the highest cultural and scientific distinction in Mexico. His current research interests include the chemistry of the origin of life and the use of phylogenomic databases to reconstruct the early stages of cell evolution.