DNA AND CELL BIOLOGY
Volume 39, Number 2, 2020

Pp. 177–186
DOI: 10.1089/dna.2019.4943

Ribose Selected as Precursor to Life

Gaspar Banfalvi

The chemical or prebiotic evolution referred also to as pre-Darwinian evolution describes chemical reactions up to the origin of a self-replicating system that was capable of Darwinian evolution. These chemical processes took place on Earth between about 3.7 and 4.5 billion years ago when cellular life came into being. The pre-Darwinian chemical evolution usually assumes hereditary elements, but does not regard them as self-organizing processes. Physical and chemical self-organization led to uninterrupted pre-Darwinian and Darwinian evolution. Thus, it is not justified to distinguish between different types of evolution. From the many possible solutions, evolution selected among those reactions that generated catalytic networks incorporating chemical sequence information and under gradually changing circumstances produced a reproducible and stable living system that adapted to these conditions. Major issues in this review involve prebiotic reactions leading to genetic evolution involving (1) abiotic sources of components of ribonucleotides and xenobiotic nucleotides, (2) formation of prebiotic RNA, (3) development of genetic RNA from random-sequence noncoding RNA, (4) transition from RNA World to DNA Empire, (5) the role of oxygenic photosynthesis in genetic transitions, and (6) hierarchical arrangement of processes involved in the optimized genetic system.

Keywords: ribose conformations, selection of ribose, artificial nucleic acids, prebiotic RNA, DNA empire, processes of cellular information

Introduction

The study of the origin of life is an effort to understand how chemical reactions could have resulted in a primordial cell system (Chronin and Walker, 2016). Several theories exist that try to answer the question of how disordered chemical reactions developed into a highly organized biological system. Using genomic and fossil evidence, the possibility of a very early origin of life was described by Betts *et al.* (2018). Pre-Darwinian evolution refers to chemical evolution, whereas Darwinian evolution corresponds to the encoded genetic information upon which the natural selection acts (Tessera, 2018). This definition is not taking into consideration that Darwinian evolution is not equal to biological evolution as often selection is overridden by nonselective factors such as growing populations of small size and varying trait composition.

The distinction among different kinds of evolution would mean that processes could go in several directions despite the uncertainty regarding precellular or postcellular life or any indication that other life forms would have existed or been replaced by ribonucleic acid. There is no agreement among scientists regarding the appearance of noncoding RNA (nRNA) and coding genetic RNA (gRNA) or how this transition could have taken place. Disagreement exists even among supporters of the RNA World hypothesis, questioning whether the RNA World was a shorter (100 millions of years) or a longer time lasting for up to billions years.

The RNA World hypothesis is opposed also because it could not have taken a long time between the appearance of nRNA, and bacteria representing gRNA. Geochemical evidence supports the notion that life existed before 3.7–3.8 billion years (Mojzsis *et al.*, 1996; Rosing, 1999). There is an agreement that archaebacteria and bacteria developed from a common ancestor before 4 billion years in a reducing atmosphere in the absence of oxygen. In the last decades, the primitive prebiotic (preoxygenic) atmosphere has been considered to have been only weakly reducing or even neutral (Cleaves *et al.*, 2008).

The fossils of oxygen-producing blue-green bacteria (cyanobacteria) were among the youngest group of bacteria. Oxygenic photosynthesizer cyanobacteria may have originated before 3 billion years (Schirrmeister *et al.*, 2015). Other fossil evidence of cyanobacterial oxygen production extends back to 3.6–3.7 billion years (Schopf, 2012). Radioisotope dating confirmed that extremophile cellular metabolism existed even earlier as a carbon-based self-replicating microbial life. The comparison of the age of Earth (4.54 billion years) with the appearance of bacterial life under extreme conditions suggested that only a relatively narrow window was left for the RNA world to come into being (Manhesa *et al.*, 1980; Dalrymple, 2001). The rise of atmospheric oxygen levels about

Department of Molecular Biotechnology and Microbiology, University of Debrecen, Debrecen, Hungary.

2.32–2.45 billion years ago took place during the Great Oxygenation Event (Bekker *et al.*, 2004; Schirrmeister *et al.*, 2011) and oxygenation continues since.

We do not know how fast/slow these transitions could have been because we fundamentally do not understand them yet. According to the recent view, life on Earth came into being through chemical evolution by producing increasingly complex molecules and yielding a molecule with the properties of information storage and replication prone to mutations (Gilbert, 1986; Bada and Korenaga, 2018). Nevertheless, the critical chemical reactions yielding the first informational macromolecule have yet to be clarified.

In this review, it is postulated that the selection of β-D-ribose and the chemical reactions of ribose were critical to generate the first informational macromolecule RNA. Others defended the view that activated ribonucleotides polymerized and formed the first RNA molecule by synthesizing nucleotides with their N-glycosidic bonds already formed, rather than synthesizing the sugar and nucleobase separately (Powner *et al.*, 2008). It is assumed that under prebiotic conditions, the synthesis of nucleotides took place through abiotic reactions.

Among the contradictions of the nonenzymatic nucleotide synthesis, one could mention (1) the first biopolymer could have been RNA, without confirming the existence of RNA world, hypothesized by Gilbert (1986), (2) the formation of ribose by chemical means such as the formose reaction (Butlerow, 1861), although the formose reaction turned out to be quite efficient under certain reaction conditions (Kopetzki and Antonietti, 2011), (3) sugar backbone formation (Bernhardt and Sandwick, 2014), or by synthesizing nucleotides first and polymerizing them with the N-glycosidic bond already formed (Powner *et al.*, 2008), (4) the inefficiency of adding nucleobases to ribose (Powner *et al.*, 2008), and (*5*) the scarcity of nucleotides favoring the synthesis of RNA from simpler and available precursors such as sugars and bases (Unrau and Bartel, 1998). These contradictions can be explained in different ways.

It is suggested that ribose-phosphate could have been the basic compound for abiotic RNA synthesis. The synthesis of RNA could have started with the following: (1) the selection of ribose as the exclusive sugar component, as only this pentose fits perfectly into the known nucleotides, (2) the cyclization of ribose increased the conformational flexibility and reduced the number of reactive substituents, (3) the phosphorylation of ribose could have served as basis for its polymerization, (4) the addition of nucleobases in a random manner to the ribose-phosphate backbone could have generated nRNA, (5) the polymerization of the sugar-phosphate backbone was probably followed by the hydrolysis of nRNA to oligonucleotides and mononucleotides, (6) the rephosphorylation of mononucleotides generated the nucleoside triphosphate (NTP) pool for the synthesis of selected nucleotides brought about the gRNA (Banfalvi, 2019), and (7) the long-term evolution resulted in hierarchical processes of cellular information with DNA replacing RNA as the primary genetic information.

Supporting ideas to this explanation came from spontaneous chemical reactions between macromolecules with phosphodiester backbones resulting in chemical stability (Yakhnin, 2013) and from the polymerization of ribose phosphates to generate the ribose-phosphate backbone (Bernhardt and Sandwick, 2014).

Abiotic Synthesis of Nucleotides

The known nucleotides that have been developed during evolution are complex molecules themselves consisting of three components: (1) pentose sugar, (2) 1 to 3 phosphates, and (3) nucleobases attached to the pentose. The reactivity of nucleic acids revealed that DNA is a relatively inert molecule, but most suitable to store and to pass information in cells from generation to generation.

The structure-function relationship of informational macromolecules could be connected to the more reactive RNA molecule emerging first. The reactivity sequence could also explain why phosphorylation of ribose preceded the polymerization of ribose phosphate and the lower reactivity of N-glycoside formation during the attachment of nucleobases to the sugar-phosphate backbone. The lowest reactivity of deoxyribose suggests that this molecule emerged later when the stabilization of the genetic information in the form of DNA became of primary importance in an environment changing from reductive to the oxidative direction. After several arguments were given against the RNA World hypothesis, one should also mention those that are in favor of it.

To strengthen the idea that RNA existed before DNA, more than a dozen arguments were compiled, indicating that the reactive ribonucleotides and RNA could have preceded the formation of the more stable and inert deoxyribonucleotides of DNA (Banfalvi, 2009, 2019). The reigning DNA Empire theory (Banfalvi, 2009) implies the genetic code, existence of a ribosome-based translation, and replacement of RNA with DNA (Torrents, 2014). The replacement of RNA was possible by ribonucleotide reductase enzymes catalyzing the reduction of the ribose C2′-OH to C2′-H (Thelander and Reichard, 1979). Under prebiotic conditions with an anaerobic atmosphere (Martin and Sousa, 2016), class III of ribonucleotide reductase could have been the ur-reductase from which class I and II reductases were derived (Reichard, 1993). The ribonucleotide reductase reactions and the formation of DNA are supposed to be relatively late events of the nucleotide evolution.

The function of informational macromolecules suggests that the more reactive RNA molecule could have emerged first. This review focuses on the origin of informational macromolecules; thus, more attention is paid to the probably first informational macromolecule RNA. The RNA World hypothesis (Gilbert, 1986) assumes that it was RNA that could have been the oldest informational macromolecule followed by proteins and DNA was the last one that came into being. A central element of the RNA World hypothesis was provided by the catalytic activity of ribonucleic acid that empowered RNA with an enzymatical function to replicate itself (Cech, 1989). Although, recently, we do not know an RNA molecule that would catalyze the synthesis of another RNA, there are plenty of examples of self-splicing RNAs known as ribozymes.

Among the major arguments supporting the RNA World hypothesis are the several functions of RNA, including the ability to store and pass hereditary information (Miller and Urey, 1959), and the formation of sugars, primarily ribose (Butlerow, 1861). Among the best known nonenzymatic pathways to ribose formation, we find the formose (from formaldehyde and aldose words) or Butlerow reaction with

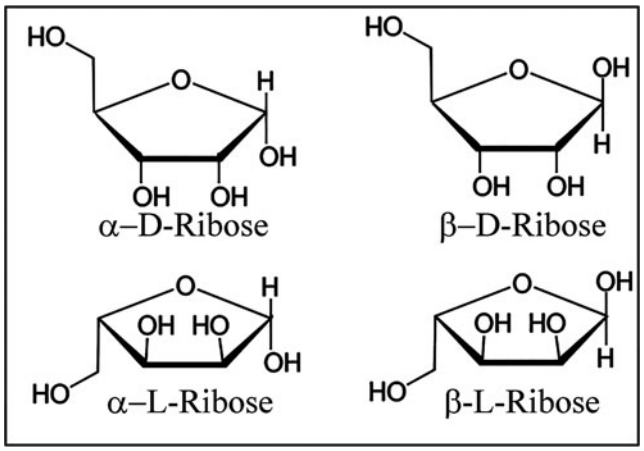

FIG. 1. Configurations of cyclic D- and L-riboses. Glycosidic hydroxyl group in α or β anomeric position.

the nonproductive Cannizzaro disproportionation of formaldehyde to methanol and formate and catalytic dimerization of formaldehyde to C2-C6 aldoses and ketoses, as well as aldose-ketose isomerizations (Butlerow, 1861). However, when the formose reaction was performed under hydrothermal temperatures or even simpler conditions, maximum yields of carbohydrates were obtained. High yields were produced without the catalytic influence of active cations, and a slightly alkaline pH was sufficient to induce the formose reaction in less than 1 min (Kopetzki and Antonietti, 2011). The higher productivity of ribose formation resulted in a more favorable attitude toward the RNA World hypothesis.

Ribose Configurations

During the cyclization, the crystalline D-ribose in aqueous solution (Angyal, 1969) is undergoing an internal aldol condensation, including the following steps: (1) delocalization of aldehyde group, first configuration change between the groups of the H atom and the hydroxymethyl group of the 5th carbon atom (from D to L). (2) The second change in configuration between the OH group and H atom of the 5th carbon atom restoring the original configuration (L to D). (3) Nucleophilic addition (A_N) reaction. (4) Cyclization of D-ribose. Similar spontaneous cyclization is taking place with L-ribose. The perspective representation of cyclic structures of D- and L-ribose and anomeric configurations are shown in Figure 1.

Ribose Conformations

The conformations of ribose can be traced back to the conformations of the 5-membered cyclopentane where two major forms, the envelope and the predominating twisted conformations, can be distinguished (Fig. 2a). Ribose is the derivative of tetrahydrofuran (Fig. 2b) with 4 atoms being in one plane of the envelope form and 3 atoms in the twisted conformation (Fig. 2b). The strong electronegativity of oxygen forces the neighboring carbon atoms to be in one plane. The configurations of ribose in D- and L-ribose are shown in Figure 2c. The twisted cyclic ribose has further so-called sugar-pucker conformations with two dominant forms: (A) C3′-endo major and C2′-exo minor conformation and (B) C2′-endo major and C3′-exo minor forms found in A- and B-DNA, respectively (Fig. 2d).

Phosphorylation of Ribose

The phosphorylation of ribose in the precellular era and protocells remained crucial during evolution as cells can use ribose only as D-ribose-5-phosphate. The initial low yield of ribose at the time of prebiotic RNA formation could have been compensated by the protective and stabilizing effect of borate-containing minerals (Benner, 2004; Ricardo *et al.*, 2004; Furukawa and Kakegawa, 2017), by phosphate

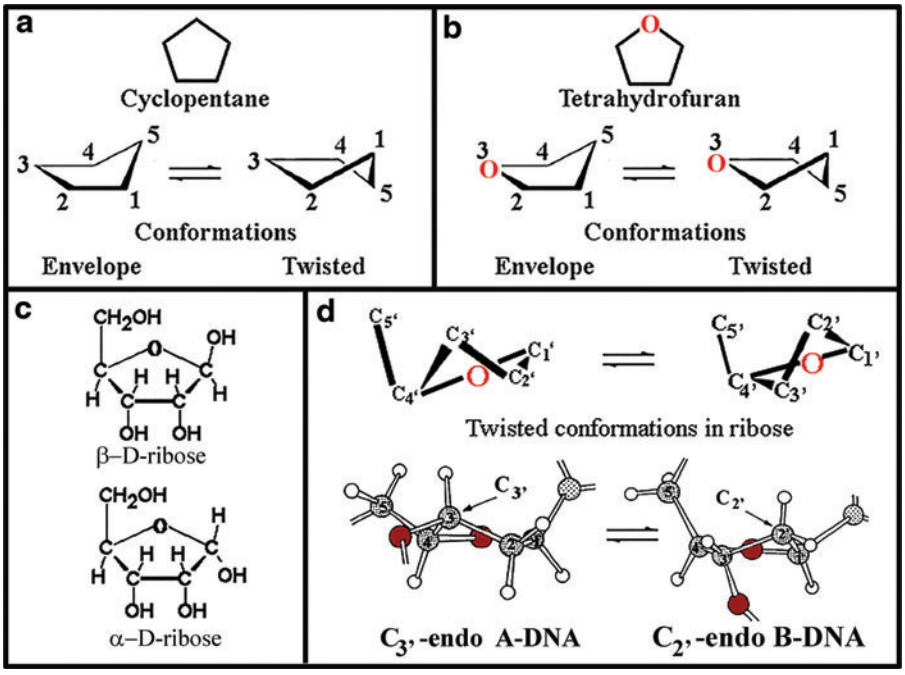

FIG. 2. Ribose conformations and configurations. **(a)** Major conformers of cyclopentane. **(b)** Envelope and twisted conformers of tetrahydrofuran. **(c)** D-configuration as well as α and β anomeric configurations of D-ribose. **(d)** Twisted conformations in ribose, C3′-endo in A-DNA and C2′-endo conformations in B-DNA. Color images are available online.

(Müller *et al.*, 1990), or by adduct formation with cyanamide (Springsteen and Joyce, 2004). Although it remains uncertain whether phosphorus limitation persisted throughout Earth's history (Reinhard *et al.*, 2017), it is assumed that a substantial quantity of phosphate may have come from photoprocessed cosmic ice (Meinert *et al.*, 2016).

Attachment of Nucleobase to Ribose

The RNA World hypothesis (Gilbert, 1986) turned the interest from macromolecular synthesis toward inorganic chemical reactions. Orgel's pathway suggested that ribose was generated from glycolaldehyde and formaldehyde in an early atmosphere containing humid CO_2. Nucleobases were coupled to the rings of ribose forming 5-membered rings, ribofuranosides, or 6-membered ribopyranosides (Orgel, 2004). The missing link in Orgel's pathway was the polymerization of ribose phosphate (Bernhardt and Sandwick, 2014). Although the polymerization of other alternate sugar backbones that form base-pairing systems such as ribopyranose RNA is not only possible in solution, D-ribose exists as a mixture of various structures predominating β-D-ribopyranose (Drew *et al.*, 1998). However, in all naturally occurring biomolecules containing D-ribose, it is present in the form of β-D-ribofuranose (Wiercigroch *et al.*, 2017). Among the structural constraints, the stability, ability to form complementary base pairs, linear covalent polymers, the encoded information, compaction in small spaces such as cells, could be mentioned. To provide a variety of structural isomers RNA could potentially function as a genetic platform.

RNA plays a central role in biochemistry as the transcriptional intermediary of genetic information as well as the mediator of the translation of mRNA messages into peptides and proteins. It has been suggested that RNA preceded DNA in biochemical evolution based on several lines of evidence, including the central role of RNA in the flow of information within the cell (Woese, 1967; Crick, 1968; Orgel, 1968), the fact that deoxyribonucleotides are often biosynthesized from ribonucleotides (Benner *et al.*, 1989), and because many enzyme cofactors are ribonucleotide derivatives (White, 1976; King, 1980)

The polymerization of informational macromolecules would have allowed the generation of polymeric systems in many ways; the arrangement of polymerizing molecules and the attachment of different nucleobases in a linear manner were necessary to generate informational macromolecules. Nucleobases attached to ribose moieties became the genetic letters of nucleic acids. The linear arrangement of nucleotides was put forward first as an *a priori* dogmatic postulate, then as an empirical fact, and finally led to the recognition that the linear arrangement in DNA stands for the genetic sequence of RNA and amino acids in proteins (Ratner, 1973; Stent, 1977).

Selection of Ribose as the Exclusive Sugar Component in Nucleotides

By accepting the RNA World hypothesis and RNA as the first informational macromolecule, the question arises why in nucleotides ribose was selected as the exclusive sugar component. For the selection of ribose as a sugar component of nucleic acids, molecular modeling experiments provided the clue (Banfalvi, 2006). These experiments tested the steric orientation and space-filling of sugar-base connections inside the nucleotide structure. Molecular modeling revealed that of the four pentoses (ribose, arabinose, xylose, and lyxose), only β-D-ribose could be inserted into nucleotides to secure the free rotation for the functional groups (OH, phosphate, and base) and the stability, including the possibility of forming a double-stranded structure (Banfalvi, 2006). Unrestricted rotation of functional groups in ribonucleotides is possible only with β-D-ribose (Fig. 3a). The conformers originating from sugar pucker folding as well as the pentose configurations indicated that the selection of ribose was not a random, but the only possible solution as only β-D-ribose is fitting into a functional ribonucleic acid.

By trying to fit any other pentose into the nucleotide, a sterical barrier is created, indicated by the black arrows in Figure 3b–d. It deserves mention that in the hypothetical lyxonucleotide, all three functional groups, namely the C2′-OH, the C3′-OH, and the C5′-phosphate groups of ribose, are within van der Waals radius distance preventing their free rotation. This could be the reason why lyxonucleotides do not exist. The advantage of selecting β-D-ribose over α-D-ribose was that, in nucleotides, the C1′-base and C2′-OH substituents of ribose became localized distantly enough to move freely and provide flexibility.

Chemical Alternatives to the Ribose Sugar Backbone

There could have been other macromolecules that could have reproduced themselves such as threose, peptide, or glycol nucleic acids (Eschenmoser, 1999, 2004; Nelson *et al.*, 2000; Orgel, 2000). Among them, threose nucleic acid (TNA) is chemically simpler than RNA and has other properties that make it a plausible alternative to RNA (Eschenmoser, 2004). The lack of rotational freedom could have prevented them from becoming self-generating molecules (Fig. 4). Unlike amino acids, nucleotides as complex compounds are not among the products of the so-called one-pot synthesis; consequently, preRNA could not have been synthesized directly from nucleotides, but only from their components. The search for new building blocks for the synthesis of artificial nucleic acids resulted in xenonucleotides (Herdewijn and Marliere, 2009) of different chemistries (Pinheiro and Holliger, 2012; Pinheiro *et al.*, 2014) and allowed even the synthesis of duplex formation (Pinheiro and Holliger, 2014) and storage of information (Yu *et al.*, 2013).

Nevertheless, the question remains whether these building blocks could have been formed under prebiotic conditions when high-energy nucleotides and proteins did not exist. Molecular modeling shows that, in RNA (Fig. 4a), the functional groups attached to ribose are distantly located securing the free movement for 2′C-OH, 3′C-OH, 5′-phosphate, and 1′C-nucleobase (G), whereas in chains of artificial nucleic acids (glycol nucleic acid [GNA]; threose nucleic acid [TNA]; peptide nucleic acid [PNA]), RNA functional groups would be in close proximity with glycol, threonine, and the peptide base, respectively (Fig. 4b–d). Also, in PNA, the guanine nucleobase and the peptide of the same nucleotide would be too far from each other and only the upper next peptide would be at the same level as the G nucleobase.

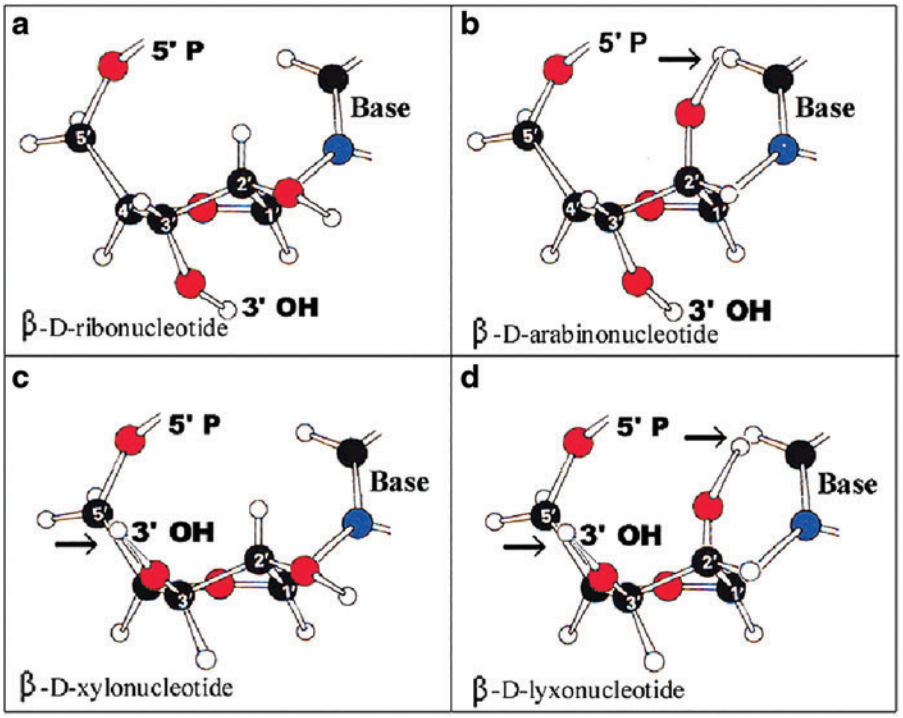

FIG. 3. β-D-ribose, the best fitting pentose in the nucleotide building block of ribonucleic acid. **(a)** In the ribonucleotide, the bulky base at C1′ and the phosphate group at C5′ position are sufficiently distant from each other to be outside the van der Waals force, responsible for a distance-dependent interaction between atoms and molecules. The C2′ and C3′ hydroxyl groups are at the opposite side of the pentose ring. Due to the major C2′-endo and the minor C3′-exo conformation, the hydroxyl groups C2′ and C3′ point to different directions allowing their free rotation. **(a)** Ribonucleotide; **(b)** Arabinonucleotide: the C2′ hydroxyl is above the plane of the ribose ring, but the base is within van der Waals distance relative to C2′-OH. The spherical hindrance of these groups prevents their free rotation indicated by the *upper right arrow* in this panel. **(c)** Xylonucleotide: the sterical vicinity of C3′ and C5′ substituents hinders the free rotation indicated by the arrow at the *left side* of this panel. **(d)** Lyxonucleotide: all functional groups are above the plane of the ring. The spatial incompatibility is caused by the lack of space. Free rotation is prevented by the C5′-phosphate, C3′-OH, and C2′-OH indicated by the *arrows*. With permission from Banfalvi (2006). Color images are available online.

It is thus questionable whether long artificial nucleic acids could have been polymerized as nucleotides and enzymes did not exist in the prebiotic era in conformity with the experimental evidence that GNA and TNA were not sequential polymers in the prebiotic evolution of RNA (Yang *et al.*, 2007). Nevertheless, artificial oligonucleotides or peptides could have been polymerized and selected from the large random sequence pool that recognizes and binds to specific target molecules, known as aptamers. Aptamers are used among others in basic research, clinical practice, and macromolecular drugs. The discovery of a nucleic acid-based genetic regulatory element of an mRNA molecule (riboswitch) (Edwards and Batey, 2010) with recognition properties similar to aptamers contributes to the credibility of the''RNA World'' hypothesis.

Formation of Prebiotic RNA

The formation of prebiotic RNA assumes that ribonucleotide components for RNA synthesis were not available on the prebiotic Earth. One of the most debated questions concerns the availability and synthesis of prebiotic ribose. The oldest theory of ribose formation comes from the formose reaction (Butlerow, 1861), but does not explain how

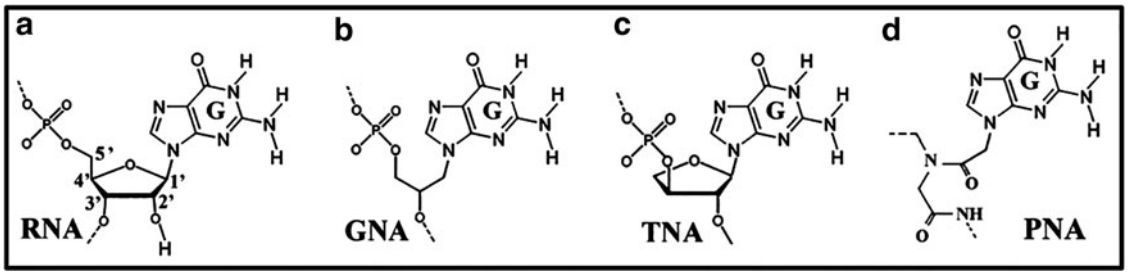

FIG. 4. Guanine nucleobase in skeletons of RNA **(a)**, GNA **(b)**, TNA **(c)**, and PNA **(d)**. GNA, glycol nucleic acid; PNA, peptide nucleic acid; TNA, threose nucleic acid.

the low yield of ribose in this reaction could lead to the macromolecular synthesis of RNA (Decker *et al.*, 1982; Shapiro, 1984). Besides, the ester linkage between ribose and phosphoric acid is known to be prone to hydrolysis (Lindahl, 1993). Among the major weaknesses of the RNA World hypothesis, the scarcity of ribose was mentioned as it was not available as a building block for the synthesis of prebiotic RNA. Chemical obstacles were also among the arguments that denied the formation of nucleic acids under abiotic conditions, but became later products of evolution (Shapiro, 1984, 1986). When the formose reaction was performed under hydrothermal temperatures, maximum yields of carbohydrates were obtained without catalytic influence. A slightly alkaline pH was sufficient to induce the formose reaction (Kopetzki and Antonietti, 2011). Among the major weaknesses of the RNA World hypothesis, the scarcity of ribose was mentioned and it was not available as a building block for the synthesis of prebiotic RNA. Chemical obstacles were also among the arguments that denied the formation of nucleic acids under abiotic conditions, but became later products of evolution (Shapiro, 1984, 1986).

The attachment of nucleobases to the ribose-phosphate backbone to form a random nucleotide containing nongenetic preRNA was followed presumably by the hydrolysis of the polymer chain. The released nucleoside monophosphates (NMPs) were phosphorylated to triphosphates and served as a selection pool for the synthesis of gRNA (Banfalvi, 2019) (Fig. 5).

Ribose 5′ phosphates (Fig. 5a) could polymerize both through their C2′- or C3′-OH groups. Preferentially, the 3′-OH group serves the head-to-tail elongation and phosphodiester bond formation between the two riboses. The reactive cis-diols of ribose stimulated the generation of the ribose-phosphate backbone (Fig. 5b) to which purine and pyrimidine bases (Fig. 5c) were attached randomly and the nongenetic preRNA formed (Fig. 5d).

Formation of gRNA from nRNA

It is postulated that to become a genetic molecule, the preRNA was degraded first to oligonucleotides (Fig. 5e), and further to mononucleotides (Fig. 5f). The NMPs upon rephosphorylation became nucleoside triphosphates (Fig. 5g). The 4 NTPs could have served as a pool for the selective synthesis of gRNA (Fig. 5h).

From RNA World to DNA Empire

The origin of proteins in support of nucleotide synthesis, as well as the coexistence of RNA, proteins, and prebiotic compounds are in line with the concept of RNA World (Wächtershäuser, 2007). Proteins providing more specific catalytic surfaces than RNA became the second informational macromolecules. This raises the question, when did DNA, the third player of informational macromolecules, step into service.

It is unlikely that the production of oxygen occurred in the RNA World, rather the limited coding capacity and stability of RNA was not sufficient to support the evolution of a photosynthetic apparatus, but had to wait for the emergence of DNA and membrane systems. The appearance of DNA resulted in the accumulation of environmental pollution with the production of molecular oxygen (O_2) by hydrolytic photosynthesis by cyanobacteria (blue-green algae). All plants on Earth contain cyanobacteria

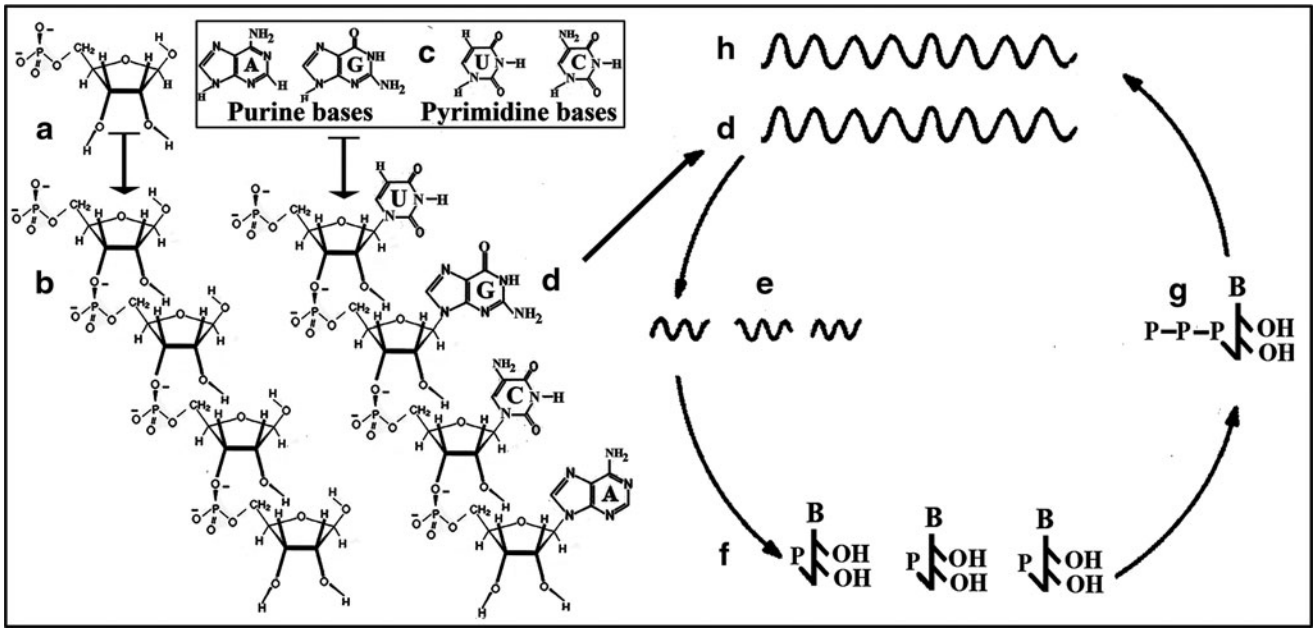

FIG. 5. Schematic view of prebiotic and gRNA formation. **(a)** Phosphorylation of β-D-ribose to β-D-ribose-5P. **(b)** Polymerization of sugar-phosphate backbone from β-D-ribose 5-phosphate units. **(c)** Abiotic purine and pyrimidine bases. **(d)** Nucleobases reacting with glycosidic hydroxyl groups of the sugar-phosphate backbone to bring about preRNA. Degradation of preRNA **(d)** to **(e)** oligonucleotides serving catalytic functions and **(f)** NMPs. Synthesis of gRNA: **(g)** phosphorylation of NMPs to nucleoside triphosphates. **(h)** Synthesis of gRNA from nucleoside triphosphate units (ATP, GTP, CTP, and UTP). P, phosphate; P-P, pyrophosphate; B, base. Phosphate ester and pyrophosphate bonds are not indicated. With permission from Banfalvi (2019). gRNA, genetic RNA; NMPs, nucleoside monophosphates.

(chloroplasts), perform photosynthesis, and produce oxygen. The concentration of oxygen in the air was constantly increasing and is now about 21% (v/v). In turn, the loss of water and the gradual increase in salt concentration in seawater become measurable, but already take an alarming proportion (Banfalvi, 2017).

According to estimations of the isotopic ratio of the sulfur transformation, the oxygen was becoming a significant portion of Earth atmosphere about 2.45 billion years ago as evidenced by the oxidized iron to appear in ancient soil (Farquhar *et al.*, 2000). Geological and geochemical evidence also indicate that the level of atmospheric oxygen was extremely low before 2.45 billion years, but reached a considerable level by 2.22 billion years ago (Bekker *et al.*, 2004). The time that passed between the initial oxygen production in the sea and its appearance in the atmosphere could be explained by the saturation (3% v/v) of dissolved oxygen in the sea. In agreement with this idea, cyanobacteria relatively rapidly occupied the surface of photic zones of the ocean, but deeper layers were distributed much later (Johnston *et al.*, 2009), probably by hydrothermal vents. In conformity with this observation, the oxygen appeared first in seawater where the oxygen was initially produced by the ~3.6–3.7 billion-year-old fossilized stromatolites found in Western Australia and southwest Greenland (Rosing, 1999).

From an informational point of view, oxygen, the ''by-product'' of photolysis produced by DNA-containing cyanobacteria, did not arise during the reducing environment of the RNA World period, but speeded up the evolution of DNA Empire. The sequence hypothesis (Crick, 1968), the central dogma (Crick, 1970), and the reverse transcription (Baltimore, 1970; Temin and Mizutani, 1970) have been accepted as the basic rules of molecular biology summarized as DNA makes RNA makes protein, but protein never makes DNA and RNA makes DNA only in retroviruses. There are several competing theories of how life on Earth came into being. The RNA World theory and its transition to DNA Empire became popular theories of the evolution of life on Earth depicted in Figure 6.

It is assumed that there could have been no RNA-encoded organism since RNA could not code for large molecules such as those required for the cellular machinery. The RNA World was strictly abiotic and cells could have evolved only after the invention of DNA. The saturation of oxygen in the oceans (~3%) and its accumulation in the atmosphere could have taken a much longer time than the RNA World. The primary information could have been subjugated to the DNA Empire during the Great Oxygenation Event.

Development of DNA Empire

Bacteria played a major role in the events after the transition from RNA World to DNA Empire. Photosynthesis originated in bacteria, with some bacteria possessing photosystems that produced oxygen similar to how green plants do today. How did oxygen production impact the transition from RNA World to DNA Empire? It is now generally believed that the archaea and bacteria developed separately from a common ancestor less than 4 billion years ago. Non-oxygen-producing bacterial species such as purple and green bacteria are the oldest photosynthetic bacteria. Blue-green

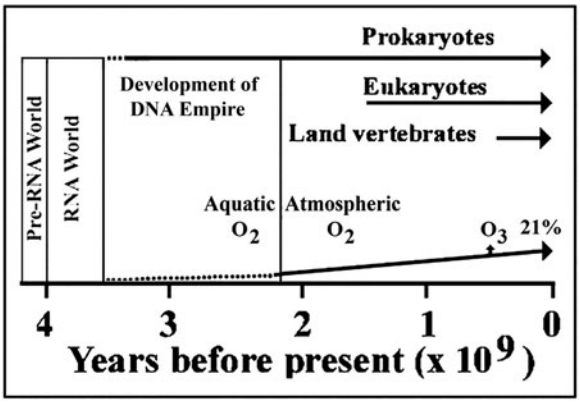

FIG. 6. Likely scenario of the evolution of life on Earth. Prebiotic RNA formation could have taken place between 4.2 and 4.0 billion years ago. Earliest possible appearance of cellular life took place when oxygen production started some 3.6–3.7 billion years ago in cyanobacteria. Atmospheric oxygen appeared 2.2 billion years ago, eukaryotes 1.5 billion years ago, land vertebrates 500 million years ago. Modified with permission from Banfalvi (2009).

algae (also called blue-green bacteria or cyanobacteria) are known from sedimentation data are ~3.6 billion years old. Cyanobacteria were among the last prokaryotic bacterial groups to form (Fig. 7).

Hierarchical Processes in the Transfer of Cellular Information

The recognition that among the genetic molecules, RNA came first, protein second, and DNA only third shows that RNA and probably an even earlier ancestor molecule, namely ribose, could be the first molecule related to and initiating the genetic system. Processes belonging to the intracellular transfer of genetic information are summarized in Figure 8. DNA↔DNA transfer processes: (1) Mutation:

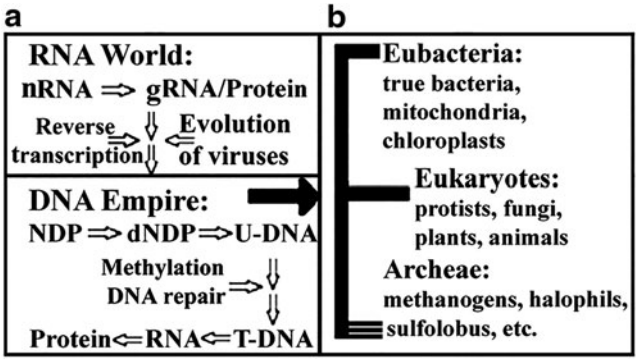

FIG. 7. Rise of DNA Empire and fall of RNA World. **(a)** Major engagements in the battlefield of cellular information. **(b)** The winner T-DNA took it all and the DNA Empire reigns ever since. With permission from Banfalvi (2009). gRNA, genetic RNA containing coding nucleotide sequence; nRNA, noncoding random nucleotide containing RNA; T-DNA, dTMP containing DNA; U-DNA, dUMP containing DNA.

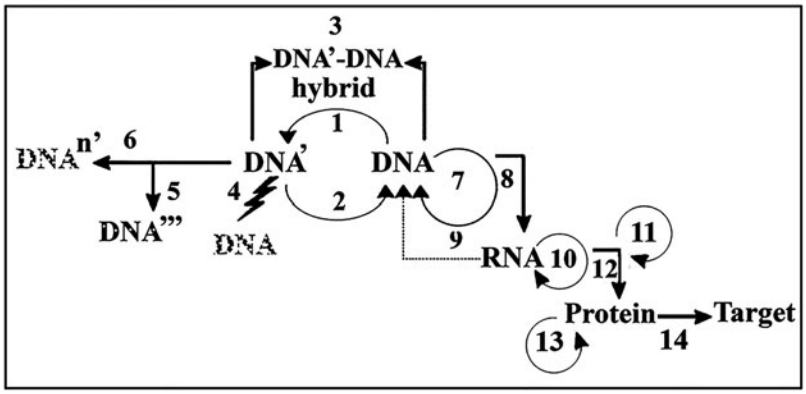

FIG. 8. Hierarchical arrangement of processes involved in the transfer of genetic information. Modified with permission from Banfalvi (2009).

DNA→DNA′. (2) DNA repair: DNA′→DNA. (3) Recombination: crossover, gene conversion (DNA′-DNA hybridization). Recombination can take place intracellularly, but is mainly an intercellular process. (4) Apoptosis (programmed cell death, high levels of DNA damage). (5) Aging, several mutations, persistent DNA damage: DNA′→DNA″→DNA‴. (6) Malignant transformation with multiple mutations: DNA→DNA$^{n\prime}$ (persistent DNA damage, many mutations, mutant p53). (7) DNA replication: DNA↔DNA reduplication (high fidelity, HiFi process, $1{:}10^{10}$ misincorporated deoxyribonucleotide). (8–9) DNA-RNA transfer. (8) Transcription: DNA→RNA (medium-fidelity, MeFi process, $1{:}10^{5}$ misincorporated ribonucleotide). (9) Reverse transcription: RNA→DNA (in retroviruses). (10) RNA replication: RNA↔RNA (in RNA viruses). Processes belonging to posttranscriptional modifications (Fig. 8): (11) 5′-cap formation, 3′-polyA formation, splicing. (12) Translation: RNA→protein (low fidelity, LoFi process, $1{:}10^{4}$ misincorporated amino acid). Processes belonging to protein modification: (13) protein splicing, transglutamination. (14) Protein targeting: information reaches intracellular or extracellular destination.

Conclusions

Biological reactions take place primarily among compounds without disturbing the compatibility of bioelements known as CHNOPS group, standing for the chemical symbols of the most frequently occurring elements in living organisms. Bioelements belong to the first three periods of the periodic table and belong to the light elements. According to the new definition, light bioelements contain exclusively s and p electrons, but no d or f electrons (Banfalvi, 2011). The most abundant atomic elements in the Universe are the lightest hydrogen and helium. The lightest atoms are less abundant in the atmosphere of the Earth due to their loss through the Planetary Air Leak to the outer space (Catling and Zahnle, 2009; Pope *et al.*, 2012). The most abundant element on Earth and in living organisms is oxygen in water molecules. The ancient and most important component of the building block of RNA was ribose, phosphorylated to 5′-ribose phosphate. It was assumed (Bernhardt and Sandwick, 2014) that ribose phosphate was polymerized to the ribose-phosphate backbone and nucleobases were randomly attached to the backbone to bring about non-gRNA.

To accumulate a selectable pool of ribonucleotides, the nRNA at the bonds between ribose and phosphates was hydrolyzed to catalytic RNA fragments and NMPs. NMPs were phosphorylated to NTPs catalyzing several processes themselves, and by the polymerization of rephosphorylated dNTPs, deoxyribonucleic acids came into being by the transition of NDPs to dNTPs, and the replacement of 2′C-OH with 2′C-H. Taking into consideration that the 2′-OH limits the stability of RNA and therefore its functional size, the evolution of large informational macromolecules could not have been established and had to await the replacement of RNA to DNA. This could have shortened the duration of RNA World and the earlier appearance of DNA Empire as suspected.

Acknowledgments

The critical reading of the article by prof. Henry Paulus is gratefully acknowledged.

Disclosure Statement

No competing financial interests exist.

Funding Information

Research related to molecular modeling was supported by the OTKA grant T0 42276 to G.B.

References

Angyal, S.J. (1969). The composition and conformations of sugars in solution. Ang Chem **8,** 157–166.

Bada, J.L., and Korenaga, J. (2018). Exposed areas above sea level on earth >3.5 Gyr ago: implications for prebiotic and primitive biotic chemistry. Life (Basel) **4,** 8. pii: E55.

Baltimore, D. (1970). RNA-dependent DNA polymerase in virions of RNA tumour viruses. Nature **226,** 1209–1211.

Banfalvi, G. (2006). Why ribose was selected as the sugar component of nucleic acids. DNA Cell Biol **25,** 189–196.

Banfalvi, G. (2009). DNA empire. In *Apoptotic Chromatin Changes*. G. Banfalvi, ed. (Springer, Dordrecht), pp. 1–29.

Banfalvi, G. (2011). Heavy metals, trace elements and their cellular effects. In *Cellular Effects of Heavy Metals*. G. Banfalvi, ed. (Springer, Dordrecht), pp. 3–28.

Banfalvi, G. (2017). Global loss of freshwater and salination of the sea. J Marine Sci Res Dev **7,** 2.

Banfalvi, G. (2019). Origin of coding RNA from random sequence RNA. DNA Cell Biol **38,** 223–228.

Bekker, A., Holland, H.D., Wang, P.L., Rumble, D. 3rd, Stein, H.J., Hannah J.L., *et al.* (2004). Dating the rise of atmospheric oxygen. Nature **427,** 117–120.

Benner, S.A. (2004). Understanding nucleic acids using synthetic chemistry. Acc Chem Res **37,** 784–797.

Benner, S.A., Ellington, A.D., and Tauer, A. (1989). Modern metabolism as a palimpsest of the RNA world. Proc Natl Acad Sci USA **86,** 7054–7058.

Bernhardt, H.S., and Sandwick, R.K. (2014). Purine biosynthetic intermediate-containing ribose-phosphate polymers as evolutionary precursors to RNA. J Mol Evol **79,** 91–104.

Betts, H.C., Puttick, M.N., Clark, J.W., Williams, T.A., Donoghue, P.C.J., and Pisani, D. (2018). Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin. Nat Ecol Evol **2**, 1556–1562.

Butlerow, A. (1861). Synthetic formation of a sugary substance [In German]. Justus Liebig's Annales of Chemistry **120,** 295–298.

Catling, D.C., and Zahnle, K.J. (2009). The planetary air leak. Sci Am **300,** 36–43.

Cech, T.R. (1989). RNA as enzyme. Biochem Int **18,** 7–14.

Chronin, L., and Walker, S.I. (2016). Beyond prebiotic chemistry. Science **352,** 1174–1175.

Cleaves, H.J., Chalmers, J.H., Lazcano, A., Miller, S.L., and Bada, J.L. (2008). A reassessment of prebiotic organic synthesis in neutral planetary atmospheres. Orig Life Evol Biosph **38,** 105–115.

Cleaves, H.J. 2nd, Meringer, M., and Goodwin, J. (2015). 227 Views of RNA: is RNA unique in its chemical isomer space? Astrobiology **15,** 538–558.

Crick, F.H. (1968). The origin of the genetic code. J Mol Biol **38,** 367–379.

Crick, F.H. (1970). Central dogma of molecular biology. Nature **227,** 561–563.

Dalrymple, G.B. (2001). The age of the Earth in the twentieth century: a problem (mostly) solved. Special Publications, Geological Society of London **190,** 205–221.

Decker, P., Schweer, H., and Pohlmann, R. (1982). Bioids: X. Identification of formose sugars, presumable prebiotic metabolites, using capillary gas chromatography-mass spectrometry on n-butoxime trifluoroacetates on OV-225. Chromatogr A **244,** 281–291.

Drew, K.N., Zajicek, J., Bondo, G., Bose, B., and Serianni, A.S. (1998). ^{13}C-labeled aldopentoses: detection and quantification of cyclic forms by heteronuclear 1D and 2D NMR spectroscopy. Carbohydr Res **303,** 199–209.

Edwards, A.L., and Batey, R.T. (2010). Riboswitches: a common RNA regulatory element. Nat Educ **3,** 9.

Eschenmoser, A. (1999). Chemical etiology of nucleic acids. Science **284,** 2118–2124.

Eschenmoser, A. (2004). The TNA-family of nucleic acid systems: properties and prospects. Orig Life Evol Biosph **34,** 277–306.

Farquhar, J., Bao, H., and Thiemens, M. (2000). Atmospheric influence of Earth's earliest sulfur cycle. Science **289,** 756–758.

Furukawa, Y., and Kakegawa, T. (2017). Borate and the origin of RNA: a model for the precursors to life. Elements **13,** 261–265.

Gilbert, W. (1986). The origin of life: the RNA World. Nature **319,** 618.

Herdewijn, P., and Marliere, P. (2009). Toward safe genetically modified organisms through the chemical diversification of nucleic acids. Chem Biodivers **6,** 791–808.

Johnston, D.T., Wolfe-Simon, F., Pearson, A., and Knoll, A.H. (2009). Anoxygenic photosynthesis modulated Proterozoic oxygen and sustained Earth's middle age. Proc Natl Acad Sci U S A **106,** 16925–16929.

King, G. (1980). Evolution of the coenzymes. Biosystems **13,** 23–45.

Kopetzki, D., and Antonietti, M. (2011). Hydrothermal formose reaction. N J Chem **35,** 1787–1794.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. Nature **362,** 709–715.

Manhesa, G., Allègre, C.J., Dupréa, B., and Hamelin, B. (1980). Lead isotope study of basic-ultrabasic layered complexes: speculations about the age of the earth and primitive mantle characteristics. Earth Planet Sci Lett **47,** 370–382.

Martin, W.F., and Sousa, F.I. (2016). Early microbial evolution: the age of anaerobes. Cold Spring Harbor Perspect Biol **8,** a018127.

Meinert, C., Myrgorodska, I., de Marcellus, P., Buhse, T., Nahon, L., Hoffmann, S.V., *et al.* (2016). Response on 'Ribose and related sugars from ultraviolet irradiation of interstellar ice analogs'. Science **352,** 141.

Miller, H., and Urey S.L. (1959). Organic compound synthesis on the primitive Earth. Science **130,** 245–251.

Mojzsis, S.J., Arrhenius, G., McKeegan, K.D., Harrison, T.M., Nutman, A.P., and Friend, C.R.L. (1996). Evidence for life on Earth before 3,800 million years ago. Nature **384,** 55–59.

Müller, D., Pitsch, S., Kittaka, A., Wagner, E., Wintner, C.E., and Eschenmoser, A. (1990). Chemistry of α-aminonitriles. Helv Chim Acta **73,** 1410–1468.

Nelson, K.E., Levy, M., and Miller, S.L. (2000). Peptide nucleic acids rather than RNA may have been the first genetic molecule. Proc Natl Acad Sci U S A **97,** 3868–3871.

Orgel, L.E. (1968). Evolution of the genetic apparatus. J Mol Biol **38,** 381–393.

Orgel, L.E. (2000). A simpler nucleic acid. Science **290,** 1306–1307.

Orgel, L.E. (2004). Prebiotic chemistry and the origin of the RNA World. Crit Rev Biochem Mol Biol **39,** 99–123.

Pinheiro, V.B., Arangundy-Franklin, S., and Holliger, P. (2014). Compartmentalized self-tagging for in vitro-directed evolution of XNA polymerases. Curr Protoc Nucleic Acid Chem **57,** 9.9.1–18.

Pinheiro, V.B., and Holliger, P. (2012). The XNA world: progress towards replication and evolution of synthetic genetic polymers. Curr Opin Chem Biol **16,** 245–252.

Pinheiro, V.B., and Holliger, P. (2014). Towards XNA nanotechnology: new materials from synthetic genetic polymers. Trends Biotechnol **32,** 321–328.

Pope, E.C., Bird, D.K., and Rosing, M.T. (2012). Isotope composition and volume of Earth's early oceans. Proc Natl Acad Sci U S A **109,** 4371–4376.

Powner, M.W., Gerland, B., and Sutherland, J.D. (2008). Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. Nature **459,** 239–242.

Ratner, V.A. (1973). The genetic language. Progr Theor Biol **13,** 143–227.

Reichard, P. (1993). From RNA to DNA, why so many ribonucleotide reductases? Science **260,** 1773–1777.

Reinhard, C.T., Planavsky, N.J., Gill, B.C., Ozaki, K., Robbins, L.J., Lyons, T.W., *et al.* (2017). Evolution of the global phosphorus cycle. Nature **541,** 386–389.

Ricardo, A., Carrigan, M.A., Olcott, A.N., and Benner, S.A. (2004). Borate minerals stabilize ribose. Science **303,** 196.

Rosing, M.T. (1999). ^{13}C-Depleted carbon microparticles in >3700-Ma sea-floor sedimentary rocks from west Greenland. Science **283,** 674–676.

Schirrmeister, B.E., Anisimova, M., Antonelli, A., and Bagheri, H.C. (2011). Evolution of cyanobacterial morphotypes: taxa required for improved phylogenomic approaches. Commun Integr Biol **4,** 424–427.

Schirrmeister, B.E., Gugger, M., and Donoghue, P.C. (2015). Cyanobacteria and the great oxidation event: evidence from genes and fossils. Palaeontology **58,** 769–785.

Schopf, J.W. (2012). The fossil record of cyanobacteria. In *Ecology of Cyanobacteria II*, B. Whitton, ed. (Springer, Dordrecht), pp. 15–36.

Shapiro, R. (1984). The improbability of prebiotic nucleic acid synthesis. Orig Life **14,** 565–570.

Shapiro, R. (1986). Prebiotic ribose synthesis: a critical analysis. Orig Life Evol Biosph **16,** 283–284.

Springsteen, G., and Joyce, G.F. (2004). Selective derivatization and sequestration of ribose from a prebiotic mix. J Am Chem Soc **126,** 9578–9583.

Stent, G.S. (1977). Explicit and implicit semantic content of the genetic information. In *Foundational Problems in the Special Sciences*. The University of Western Ontario Series in Philosophy of Science, Vol. 10. R.E. Butts, Hintikka, J., eds. (Springer, Dordrecht), pp. 131–149.

Temin, H.M., and Mizutani, S. (1970). RNA-dependent DNA polymerase in virions of Rous sarcoma virus. Nature **226,** 1211–1213.

Tessera, M. (2018). Is pre-Darwian evolution plausible? Biol Direct **13,** 18.

Thelander, L., and Reichard, P. (1979). Reduction of ribonucleotides. Annu Rev Biochem **48,** 133–158.

Torrents, E. (2014). Ribonucleotide reductases: essential enzymesfor bacterial life. Front Cell Infect Microbiol **4,** 52.

Unrau, P.J., and Bartel, D.P. (1998). RNA-catalysed nucleotide synthesis. Nature **395,** 260–263.

Wächtershäuser, G. (2007). On the chemistry and evolution of the pioneer organism. Chem Biodivers **4,** 584–602.

White, H.B., III. (1976). Coenzymes as fossils of an earlier metabolic state. J Mol Evol **7,** 101–104.

Wiercigroch, E., Szafraniec, E., Czamara, K., Pacia, M.Z., Majzner, K., Kochan, K., *et al.* (2017). Raman and infrared spectroscopy of carbohydrates: a review. Spectrochim Acta A Mol Biomol Spectrosc **185**, 317–335.

Woese, C. (1967). *The Genetic Code: The Molecular Basis for Genetic Expression*. (Harper & Row, New York).

Yakhnin, A.V. (2013). A model for the origin of life through rearrangements among prebiotic phosphodiester polymers. Orig Life Evol Biosph **43,** 39–47.

Yang, Y.-W., Zhang, S., McCullum, E.O., and Chaput, J.C. (2007). Experimental evidence that GNA and TNA were not sequential polymers in the prebiotic evolution of RNA. J Mol Evol **65**, 289–295.

Yu, H., Zhang, S., Dunn, M.R., and Chaput, J.C. (2013). An efficient and faithful in vitro replication system for threose nucleic acid. J Am Chem Soc **135,** 3583–3591.

Address correspondence to:
Gaspar Banfalvi, PhD, DSc
Department of Molecular Biotechnology and Microbiology
University of Debrecen
Life Sciences Building 1.102
1 Egyetem Square
Debrecen 4010
Hungary

E-mail: gaspar.banfalvi@gmail.com; bgaspar@unideb.hu

Received for publication July 5, 2019; received in revised form October 18, 2019; accepted November 4, 2019.