Rewarming the Primordial Soup: Revisitations and Rediscoveries in Prebiotic Chemistry

Raffaele Saladino,[a] Judit E. Šponer,[b] Jiří Šponer,[b] and Ernesto Di Mauro[a]

The development of some disciplines follows a straight path that is based on a shared body of progressing knowledge. Well-established branches of science follow this trend. Other disciplines enjoy frequent rediscoveries. Possibly due to its interdisciplinary character, prebiotic chemistry is one of these fields.

Interdisciplinarity is, in principle, an advantage. Scholars of genetics are well acquainted with the phenomenon dubbed "the strength of the hybrid": mixing distant characteristics often results in increased fitness. Prebiotic chemistry encompasses the inorganic or organic chemistry in a natural environment before the advent of life on Earth. Its main drive is to understand how the simple molecules present on the early Earth might have given rise to the complex systems and processes of contemporary biology. This quest profits from the ingenuity of scientists with backgrounds as different as astrochemistry, organic and quantum chemistry, geology, paleobiology, molecular genetics, theoretical physics, and cosmology.

One drawback of this state of things is that sometimes principles, facts, and reactions are revisited or rediscovered, the difference between the two terms being essentially that revisitation entails appropriate credit to the original discoverers and the development of their findings within the frame of prebiotic chemistry aims and purports, whereas rediscoveries encompass a loss of time (sometimes decades) and candid repetitions. With this in mind, a glimpse of a few central aspects of prebiotic chemistry could be of some interest. Here, we briefly travel over recent progress in the abiotic generation of key components of potentially biogenic compounds, such as nucleobases, nucleosides, nucleotides and oligonucleotides. However, we only focus on studies that are directly related to an RNA-based origin concept. Thus, although they are equally important, breakthrough studies related to alternative genetic systems and the design of self-replicating systems are not addressed in this work.

1. Nucleobases

Famously, the first abiotic synthesis of a nucleobase was reported by Oró from 1960 to 1962,[3] describing the reaction of a concentrated solution of ammonia and HCN to form adenine. The formation of guanine in a comparable chemical system was subsequently reported.[4] These pioneering discoveries were followed 12 years later by a report of the synthesis of pyrimine from formamide.[5] Simply heating formamide at 160 °C yields low amounts of purine; adenine can be synthesized in trace amounts when the reaction is performed in the presence of added HCN.

The first documented rediscovery in the prebiotic chemistry literature was related to the synthesis of pyrimidine bases, such as cytosine, and, curiously, involved two of the greatest minds working in the field, Stanley Miller and Leslie Orgel. Cytosine was synthesized for the first time in Orgel’s laboratory in the late 1960s from the reaction of cyanoacetylene with urea.[6] Almost 30 years later, Robertson and Miller reported essentially the same chemistry with the difference that cyanoacetaldehyde was used instead of cyanoacetylene in the synthesis.[7] This report was followed by a short correction that gave credit to the first discovery and pointed out that the earlier reported cyanoacetylene-based synthesis most likely proceeded through cyanoacetaldehyde.[8]

The field experienced little progress for decades. Scepticism about the worth of this bottom-up approach starting from one-carbon precursors prevailed to the point that the very possibility of abiotically obtaining a pyrimidine as important as cytosine became a subject of debate.[9]

The field was revisited in from 2001[10] (a gap of 23 years) in a paper that focused on the chemistry of formamide, which rapidly provided evidence for the nonfastidious synthesis of all the biotically relevant nucleobases (reviewed in ref. [12]).

2. Nucleosides

Understanding the circumstances enabling the prebiotic synthesis of nucleosides is notoriously difficult. Their formation might, in principle, be reconstructed through different approaches: 1) synthesizing the sugar starting at the appropriate position of the base; 2) synthesizing the nucleobase by step-wise assembly of the heterocycle at the appropriate position of the sugar; or 3) forming the connecting β-glicosidic bond between preformed sugar and nucleobase moieties. The diffi-
cully of linking nucleobases and carbohydrates via an N-glycose-

sicid bond has recently been reviewed and is a major prob-

lem in understanding the prebiotic origin of nucleosides.

For de novo synthetic approaches, pioneering studies sug-
gesting the potential of oxazoles for nucleoside synthesis
were reported in the late 1950s (reviewed in ref. [15]) and
were applied in a prebiotic context in 1970. In this last work,
was shown that pyrimidine ribonucleosides can be obtained
from oxazoline chemistry. This chemistry was revisited 40 years
later. In these systems, formamide acts as solvent for the thi-
olics of anhydro-nucleoside intermediates and for their phos-
phorylation. In another recent work, purine ribonucleosides
were synthesized through formamido pyrimidine (FPy) chemis-
try, which required formamide for the N-formylation of

3-aminopyrimidines intermediates for purine ring closure. This
reaction is the reverse of the previously reported formation of

FPy by treating purine nucleosides with formamide. Remark-

ably, formamide plays a role in both pathways.

3. Nucleotides

Notwithstanding the many successes in other areas, the possi-

bility of prebiotic phosphorylation was a focus of scepticism.

The central role of phosphates in biology (discussed by West-
heimer in ref. [23]) seemed at odds with the difficult identifica-

tion of the prebiotic source of phosphates for prebiological

chemistry (critically reviewed in ref. [25]). This feeling pre-

vailed in spite of the pioneering description of urea–inorganic

phosphate mixtures as prebiotic phosphorylating agents.

This discovery was followed by resorptive observations made

by Schoffstall between 1976 and 1988 who showed that

phosphorylation could occur through formamide chemistry,

starting from free phosphates and/or a phosphate mineral as

common as hydroxyapatite. His approach was revisited 20

years later in papers that detailed the heat dependence of

the phosphorylation reaction and showed that organic phos-

phorylation of nucleosides in formamide may occur at every

possible position of the sugar moiety and can lead to cyclic

nucleotides. Furthermore, they illustrated that numerous phos-

phate-containing minerals are plausible sources of phosphate

for the generation of nucleotides.

The face has recently been revisited and the range of pre-

biotically plausible reaction media that could drive the organic

phosphorylation of nucleosides has been extended to urea/

ammonium formate/water (UAFW) eutectic solutions. Upon

heating, the UAFW mixtures are partially converted to forma-

mide. For a summary of the studies overviewed in this and

the preceding paragraphs, see Figure 1.

4. Abiotic RNA Polymerizations

The precursors for abiotic RNA polymerization in prebiotic en-
vironments could hardly have been based on highly activated

compounds. The logic underpinning this consideration derives

from thermodynamic and kinetic considerations: the higher

the chemical potential of a compound, the lower its expected

half-life. On the other hand, when starting from much too

stable monomers, an activation step must be incorporated in

the pathway leading to oligonucleotides, thus decreasing the

chance that such chemistry would have taken place in a pre-

biotic context. Circumventing this difficulty, Orgel analyzed the

possibility that cyclic nucleotides could be appropriate precu-
sors for nonenzymatic RNA polymerization. The results, report-

ed in 1973, showed that limited oligomerization could be

achieved under alkaline conditions in dry state in the presence

of diamines or amino alcohols serving as activators. The RNA

oligomerization reaction starting from 2',3'-cyclic nucleotides

was tested and resulted in short oligomers, the monomers of

which were connected by a mixture of 3',5'- and 2',5'-phosho-
diester bonds. With the notable exception of two well-known

Figure 1. Chronology of selected historically innovative one-pot and multistep syntheses of nucleic acid building blocks. Oró (1960); ammonium hydroxide, HCN, H₂O at 90 °C. Orgel (1966); step 1: formamidine, NH₂CH(NH₂), EtOH at reflux; step 2: imidazole intermediate, formamidine, methyl cellosolve at reflux. Orgel (1970); step 1: sugar monophosphate, NCCNH₂, H₂O; step 2: oxazoline intermediate, N=C=CH; step 3: photochemistry. Schoffstall (1976); pure nucleoside, KH₂PO₄, formamide at 70 °C. Yamada (1978); formamide, HCN, sealed vial at 160 °C. Schoffstall (1982); pyrimidine nucleoside, KH₂PO₄, formamide at 125 °C. Saladino (1996); formamide at 160 °C. Saladino (2001); formamide, minerals at 160 °C. Di Mauro (2007); nucleoside, formamide, KH₂PO₄, or mineral phosphates at 90 °C. Sutherland (2009); step 1: glycolaldehyde, NCCNH₂, H₂O; step 2: amino oxazoline intermediate, glyceraldehyde, phosphate buffer; step 3: penrose amino oxazoline intermediate, N=C=CH, phosphate buffer; step 4: anhydroarabinosinucleoside intermediate, phosphate buffer. Saladino (2015); formamide, minerals, proton beam at 25 °C; Carell (2016); step 1: guanidine and aminocynocacetamide; step 2: aminopyrimidinone intermedi-
ate, formic acid or formamide at reflux; step 3: formamide pyrimidinone intermediate, ribose, dry state followed by treatment with borax at 100 °C or NH₄OH or other conditions. Sutherland (2017); step 1: amino oxazoline intermediate, and glyceraldehyde; step 2: penrose amino oxazoline intermediate, hydroxysulfide hydrate, sodium hydroxysulfide hydrate in formamide at 50 °C for 7 h; step 3: 2-thioborocytidine intermediate, hydroxysulfide hydrate, degassed water, irradiation at 254 nm for 2.5 days; step 4: NH₄H₂PO₄ urea in formamide at 100 °C for 6 days. Pu: purines, Py: pyrimidines.
subsequent studies by Usher and McHale,[31] which showed the possible prevalence of 3',5'-phosphodiester bonds, the analysis of cyclic nucleotide-based polymerization was then abandoned, and abiotic RNA syntheses starting from highly preactivated phosphoimidazolides were developed. The use of these compounds, introduced as early as 1968,[32] became the preferred methodology and stimulated the development of powerful nonenzymatic polymerization analyses. The worth of cyclic nucleotides as precursors for prebiotic RNA synthesis pioneered by Orgel was only revisited again from 2009, 36 years after it was first proposed, and is now an actively developing field of study.[33] A recent investigation demonstrated that 3',5'-cGMPs have the intrinsic propensity of selectively forming 3',5'-linked oligonucleotides.[34] Thus, the emergence of the 3',5'-linkage selectivity of modern RNAs is not necessarily the consequence of sophisticated repair mechanisms, such as those suggested in ref. [35].

Shortly after the discovery of the nonenzymatic oligomerization of 2',3'-cyclic nucleotides, in 1975, conditions leading to the formation of amino nucleotides and oligopeptides were described.[36] The chemistry was elaborated in 2015 when a reaction network utilizing carboxidiimide activation and leading to peptidyl-RNA chains was reported.[37] A completely different, RNA-catalyzed aminoacylation involving a phenylalanyl-adenosine monophosphate substrate was described by the Yarus group.[38] Ten years later, essentially the same chemistry was used to demonstrate the transfer of an aminoacyl group from the phosphate group of a 3'-aminoacylated nucleosidetriphosphate to the 2'-hydroxy group of the same nucleotide.[39]

5. From RNA Catalysis to the Ribosome

The largely accepted RNA-world scenario relies on the (self-)catalytic properties of RNA polymers. These were discovered by Cech and collaborators in 1981,[40] who described the in vitro splicing of the ribosomal RNA precursors of Tetrahymena, based on the involvement of a guanosine nucleotide in the excision of the intervening sequence. The reaction was modeled,[41] which showed that it could be performed by simpler model molecules. Recently, a similar reaction was observed for abiotically synthesized oligonucleotides afforded by cyclic ribonucleotide precursors.[42]

The recognition that the central catalytic core of the ribosome is a ribozyme[43] is considered to be the proof of principle that today's protein world had to be preceded by an RNA world,[44] in which RNA served as a carrier of genetic information and as a catalyst. It has been suggested that the modern ribosome evolved from an ancient protoribosome that catalyzed noncoded peptide bond formation.[45] This simple molecular machine could consist of two symmetry-related oligonucleotides consisting of 60–90 nucleotides. It has been suggested that pocket-like entities formed in this way are able to accommodate amino acids bound to short oligonucleotides.[45, 46] In this context, genetic coding (i.e., amino acid selectivity) is the consequence of the suitability of certain amino acid sequences for catalyzing some essentially important reactions or for increasing the stability of the molecular machinery. (For a chronological overview of selected studies related to the emergence of catalytically active oligonucleotides, see Figure 2.)

6. Concluding Remarks

Whereas the synthesis of building blocks usually proceeds on timescales of a couple of hours or days, self-assembling—the prerequisite of every oligomerization—apparently requires more time. Thus, when progressing towards higher molecular complexities, time and statistics become increasingly more important experimental factors. This makes prebiotic studies related to the emergence of catalytic activity or oligomerization technically more difficult and scarce compared to those aimed at reconstructing the synthesis of building blocks. The in comparably higher amount of information available on the basic chemistry leading to life’s building blocks might be one of the reasons why reinventions and revisitations are more common in this topic. Nevertheless, time and statistics should not be put aside even in this case: experts working on the synthesis of building blocks must also acknowledge that laboratory reconstructions of life’s origin inherently lack a very important parameter: the time that enabled nature to make a sufficiently large number of attempts to reach the goal. For this reason, every piece of the picture, regardless of whether it is a high-yield or low-yield process, is potentially an important contribution because it represents a snapshot of the whole pool of the numerous possibilities from which life evolved.

Acknowledgements

Financial support from the project STARS in the CAOS (Simulation Tools for Astrochemical Reactivity and Spectroscopy in the Cyber-
Conflict of Interest

The authors declare no conflict of interest.

Keywords: nucleosides • nucleotides • oligonucleotides • prebiotic chemistry • RNA