

# The Future of Origin of Life Research: Bridging Decades-Old Divisions

**Martina Preiner**<sup>1,\*</sup>, **Silke Asche**<sup>2</sup>, **Sidney Becker**<sup>3</sup>, **Holly C. Betts**<sup>4</sup>, **Adrien Boniface**<sup>5</sup>, **Eloi Camprubi**<sup>6</sup>, **Kuhan Chandru**<sup>7,8</sup>, **Valentina Erastova**<sup>9</sup>, **Sriram G. Garg**<sup>1</sup>, **Nozair Khawaja**<sup>10</sup>, **Gladys Kostyrka**<sup>11</sup>, **Rainer Machné**<sup>12,13</sup>, **Giacomo Moggioli**<sup>14</sup>, **Kamila B. Muchowska**<sup>15</sup>, **Sinje Neukirchen**<sup>16</sup>, **Benedikt Peter**<sup>17</sup>, **Edith Pichlhöfer**<sup>16</sup>, **Ádám Radványi**<sup>18,19</sup>, **Daniele Rossetto**<sup>20</sup>, **Annalena Salditt**<sup>21</sup>, **Nicolas M. Schmelling**<sup>12,22</sup>, **Filipa L. Sousa**<sup>16</sup>, **Fernando D. K. Tria**<sup>1</sup>, **Dániel Vörös**<sup>18,19</sup> and **Joana C. Xavier**<sup>1,\*</sup>

<sup>1</sup> Institute of Molecular Evolution, University of Düsseldorf, 40225 Düsseldorf, Germany; gargs@hhu.de (S.G.G.); tria@hhu.de (F.D.K.T.)

<sup>2</sup> School of Chemistry, University of Glasgow, Glasgow G128QQ, UK; 2304694A@student.gla.ac.uk

<sup>3</sup> Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK; sb2391@cam.ac.uk

<sup>4</sup> School of Earth Sciences, University of Bristol, Bristol BS8 1RL, UK; hb1864@bristol.ac.uk

<sup>5</sup> Environmental Microbial Genomics, Laboratoire Ampère, Ecole Centrale de Lyon, Université de Lyon, 69130 Ecully, France; adrien.boniface@univ-lyon1.fr

<sup>6</sup> Origins Center, Department of Earth Sciences, Utrecht University, 3584 CB Utrecht, The Netherlands; e.camprubicasas@uu.nl

<sup>7</sup> Space Science Center (ANGKASA), Institute of Climate Change, Level 3, Research Complex, National University of Malaysia, 43600, UKM Bangi, Selangor, Malaysia; kuhan@ukm.edu.my

<sup>8</sup> Department of Physical Chemistry, University of Chemistry and Technology, Prague, Technická 5, 16628 Prague 6-Dejvice, Czech Republic

<sup>9</sup> UK Centre for Astrobiology, School of Chemistry, University of Edinburgh, Edinburgh EH9 3FJ, UK; valentina.erastova@ed.ac.uk

<sup>10</sup> Institut für Geologische Wissenschaften, Freie Universität Berlin, 12249 Berlin, Germany; nozair.khawaja@fu-berlin.de

<sup>11</sup> Lycée Colbert, BP 50620 59208 Tourcoing Cedex, France; gladys.kostyrka@wanadoo.fr

<sup>12</sup> Institute of Synthetic Microbiology, University of Düsseldorf, 40225 Düsseldorf, Germany; machne@hhu.de (R.M.); nicolas.schmelling@hhu.de (N.M.S.)

<sup>13</sup> Quantitative and Theoretical Biology, University of Düsseldorf, 40225 Düsseldorf, Germany

<sup>14</sup> School of Biological and Chemical Sciences, Queen Mary University of London, London E1 4DQ, UK; giacomo.moggioli@qmul.ac.uk

<sup>15</sup> Université de Strasbourg, CNRS, ISIS, 8 allée Gaspard Monge, 67000 Strasbourg, France; muchowska@unistra.fr

<sup>16</sup> Archaea Biology and Ecogenomics Division, University of Vienna, 1090 Vienna, Austria; sinje.neukirchen@univie.ac.at (S.N.); a00807114@unet.univie.ac.at (E.P.); filipa.sousa@univie.ac.at (F.L.S.)

<sup>17</sup> Cellular and Molecular Biophysics, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany; peter@biochem.mpg.de

<sup>18</sup> Department of Plant Systematics, Ecology and Theoretical Biology, Eötvös Loránd University, Pázmány Péter sétány 1/C, 1117 Budapest, Hungary; adamradvanyi117@gmail.com (A.R.); mrdanielred@gmail.com (D.V.)

<sup>19</sup> Institute of Evolution, MTA Centre for Ecological Research, Klebelsberg Kuno u. 3., H-8237 Tihany, Hungary

<sup>20</sup> Department of Cellular, Computational and Integrative Biology (CIBIO), University of Trento, 38123 Trento, Italy; daniele.rossetto@unitn.it

<sup>21</sup> Systems Biophysics, Physics Department, Ludwig-Maximilians-Universität München, 80799 Munich, Germany; a.salditt@physik.uni-muenchen.de

<sup>22</sup> Cluster of Excellence on Plant Sciences (CEPLAS), University of Cologne, 50674 Cologne, Germany

\*Correspondence: xavier@hhu.de (J.C.X.); preiner@hhu.de (M.P.)

Received: 31 January 2020; Accepted: 21 February 2020; Published: 26 February 2020

**Abstract:** Research on the origin of life is highly heterogeneous. After a peculiar historical development, it still includes strongly opposed views which potentially hinder progress. In the 1st Interdisciplinary Origin of Life Meeting, early-career researchers gathered to explore the commonalities between theories and approaches, critical divergence points, and expectations for the future. We find that even though classical approaches and theories—e.g. bottom-up and top-down, RNA world vs. metabolism-first—have been prevalent in origin of life research, they are ceasing to be mutually exclusive and they can and should feed integrating approaches. Here we focus on pressing questions and recent developments that bridge the classical disciplines and approaches, and highlight expectations for future endeavours in origin of life research.

**Keywords:** origins of life; prebiotic chemistry; early life; LUCA; abiogenesis; top-down; bottom-up; emergence

---

## 1. Introduction

Understanding the origin of life (OoL) is one of the major unsolved scientific problems of the century. It starts with the lack of a commonly accepted definition of the phenomenon of life itself [1], but difficulties go far beyond merely that obstacle. OoL research involves a large number of diffuse concepts cornering several natural sciences and philosophy, such as entropy, information and complexity. Despite evidence that untangling this knot will require a concerted and collaborative effort between different disciplines, technologies, individuals and groups [2], division in OoL research is still marked, concerning both theories (e.g. RNA world vs. metabolism-first) and approaches (e.g. bottom-up vs. top-down). What causes these on-going divisions, and how can heated debates be moderated?

There is some consensus on a few points. First, the earliest undisputed fossil evidence places life on Earth prior to 3.35 Ga [3] and molecular clocks suggest an origin prior to the late heavy bombardment >3.9 Ga [4]. Second, the origin of life must have resulted from a long process or a series of processes, not a sudden event, for the complexity of a cell could not have appeared instantaneously. The OoL must have started from simple abiotic processes, involving one or more sources of energy and matter (particularly CHNOPS: carbon, hydrogen, nitrogen, oxygen, phosphorus and sulfur, the six most prevalent elements in life on Earth) forming protometabolism, compartmentalization and inheritance. But strikingly, the list of agreements does not expand much further than this. Several researchers have speculated on life forms different than the ones we know and see on Earth—not based on cells [5] or even on CHNOPS [6,7]. Even if evidence for those alternative life forms has yet to be found, here we do not exclude their possibility. Throughout this text we will focus on life as we know it, which insofar as has been demonstrated is all based on individual cells with metabolism, genetic inheritance and compartmentalization.

The list of individual theories, different lines of experimental and theoretical research and diverse views on the OoL is extensive and eclectic. It is not the purpose of this article to be a comprehensive review of all of those. Rather, based on discussions that took part in the 1st Interdisciplinary Origin of Life meeting for early-career researchers (IOoL), we present a forward-looking perspective on how discontinued discourses on the OoL can be (re)united in a new mosaic with resolution and meaning. We reflect purposely on individual topics causing the most distressing divisions in OoL research, most of which result from classical separations between disciplines and theories that date to decades ago. We then portray examples of bridges being built between classically opposed views and finish by providing a roadmap for future dialogue and evidence-based research in OoL.

## 2. Classical Divisions in Origin of Life (OoL) Research

### 2.1. Top-Down versus Bottom-Up: Where To?

Classical OoL research has gained from synthetic or bottom-up approaches, aiming at synthesizing *de novo* the building blocks of life (and assemblies of them), and top-down or analytic approaches, which start from living systems, deconstruct them into their parts and survey their properties. But while most origins questions (e.g. that of the universe, atoms, continents or social structures) are usually led by ‘their’ disciplines with small contributions from others, the OoL has been peculiar in calling multiple subjects to play major roles.

Chemistry focused on the heavy lifting at the bottom-up, synthetic side of OoL research. Chemical reactions potentially relevant to life’s beginnings date back as far as 1828 with Wöhler’s urea synthesis [8], a landmark for the birth of organic chemistry. The year of 1861 brought Butlerow’s formose reaction [9], which generates a mixture of carbohydrates from formaldehyde through an autocatalytic mechanism that came to be appreciated by the OoL community nearly a century later [10]. The launch of the chemical origins field occurred with Miller’s experiment in 1953, based on Oparin and Haldane’s theoretical work (see “Section 3. Building Bridges”), synthesizing a mixture of organic compounds (including essential amino acids) from water, methane, ammonia and hydrogen using electric discharges [11,12]. Later, significant advances were made with the synthesis of phospholipid membranes under prebiotic conditions [13]. The synthetic quest continued with efforts focusing on the synthesis of amino acids (usually involving the Strecker or the Bücherer–Bergs synthesis), sugars and nucleobases (usually based on the formose reaction and the chemistry of nitriles and their derivatives); for thorough recent reviews see [14,15] and references therein. Progress is scarcer in understanding the polymerization of proto-biological monomers into oligomers. Recent achievements include RNA synthesis assisted by ribozymes [16,17] and lipids [18]; polymerization of carbohydrate monomers [19] and peptide-bond formation at the water-air interface [20], under dehydration-hydration conditions [21] and in mineral interlayers [22]. Synthetic biology worked on the assembly of more complex features, as the synthesis of a whole bacterial genome [23] and of a viral cDNA genome that was transcribed, translated and replicated without cells [24].

Given that virtually all of the origins hypotheses are non-falsifiable, the diversity of the bottom-up approaches mentioned here has led to an important question: to what extent do the experimental synthetic advances resemble what actually happened at the OoL? With the complete set of inorganics and ready-synthesized organics at its disposal (plus a wide range of experimental conditions), the bottom-up approach parallels an artistic endeavour that can paint both abstract and realistic pictures of the first biomolecules and their assemblies. For this reason, to constrain the experimental space, clearer pictures of i) the fundamental features of life, ii) ancestral life forms and iii) the environmental conditions at their origin are urgently required.

Surprisingly, biology has played a modest part in the search for the OoL [14,25]. A closer look reveals possible reasons for this (quantitatively) sparse involvement. Biology is complex and its numbers are vast: the queue of living species is long, and only recently did technology to start exploring their workings within new disciplines (e.g. biochemistry, molecular biology and genetics) become available. Looking at microbes only, recent estimates point to 1 trillion different species on Earth [26] and orders of magnitude more bacterial cells on the planet than the number of stars in the known universe. A single one of those cells can have at any given moment 2–4 million expressed proteins [27]. By looking at whole living systems, biology has much in its hands, and a gap still remains between i) experimental biology looking at specific mechanistic aspects in a cell or organism (e.g. metabolism, particular enzymes or pathways) and ii) the larger (yet microscopic) picture of how life as a whole started, which falls out of the traditional interpretation of the theory of evolution [28]. Were there mechanisms that drove variation and selection before genomes?

One could say that the top-down approach kick-started when a small number of biological species were compared, the mechanism of speciation was uncovered and the theory of evolution formulated, in parallel with the discarding of spontaneous generation [29]. Yet, the *bona fide* dawn of top-down approaches in the OoL question would not come to be until the genomics era. Modelling

early evolution and the OoL required a precise and holistic way to trace species back in time, and that came first in the form of genome sequences. Prokaryotes, the simplest forms of cellular life, are increasingly supported by evolutionary studies as the oldest lifeforms, and thus of utmost importance for OoL research [30]. The first comparison of prokaryotic genomes revealed a conserved set of 240 genes [31]. Later, the exponential growth of the number of sequenced genomes came with a daunting realization: the prokaryote world is highly diverse, full of redundancy with non-orthologous gene displacements (unrelated sequences encoding the same biochemical function) and lateral gene transfer [32], and the intersection of genomes shrank to a mere handful of ribosomal genes [33]. The search for the genetic content of LUCA—the Last Universal Common Ancestor—using the classical comparative top-down approach had thus stagnated. Some argue that the gap between the OoL and the biochemical wiring of LUCA is so vast that whatever constituted the chemistry leading to the earliest life could have been rewritten multiple times, and changed beyond recognition to then become LUCA [34]. Others argue that even if a good reconstruction of LUCA cannot determine the chemistry at the OoL, universal biochemical features can meet with OoL chemistry by advancing experimental and analytical methods [35]. Parallel but relevant to this line of thought, a new biological concept has been devised—the IDA (initial Darwinian ancestor) [36] also called FUCA (first universal common ancestor) [37], the first entity that could be considered capable of evolution.

The bridging of all the critical advances made by chemistry's bottom-up and biology's top-down approaches seems to be an indelible requirement for the advancement of the OoL field [38–40]. It is not impossible to envision these seemingly non-overlapping approaches to be reconciled by incorporating expertise from other fields. Indeed, disciplines beyond classical biology and chemistry have much to contribute. Before exploring those contributions, we elaborate further on disputes in OoL research by briefly focusing on another particular format of OoL theories: the prebiotic 'worlds'.

## 2.2. One Origin, Abundant Worlds

The relative abundance and distribution of building blocks of the main classes of biomolecules ~4 Ga ago on Earth is still elusive and a matter of strong debate [15]. Based on earlier suggestions [41], RNA emerged among all biomolecules in a privileged metaphor that would cement and permeate OoL research for decades to come—the 'RNA world' [42]. Ribozymes constituted a shocking and promising discovery: there was now genetic material that was catalytic as well [43,44]. Many found the metaphor appealing: a world with a jack-of-all-trades RNA molecule, catalyzing the formation of indispensable cellular scaffolds, from which somehow then cells emerged [45,46]. Others were quick to notice several difficulties with that scenario. These included the lack of templates enabling the polymerization of RNA in the prebiotic complex mixture [47] and RNA's extreme lability at moderate to high temperatures and susceptibility to base-catalyzed hydrolysis [48]. The 'metabolism-first' theory emerged independently [49–52], standing in favour of simpler molecular networks harnessing energy from geological disequilibrium leading to the emergence of genetic complexity. The RNA world and the metabolism-first theories have often been portrayed in stark opposition to each other [53,54], leading the OoL field to an unprecedented division. Ever since, other prebiotic worlds came to light with their own preferred class of biomolecules and significant insights, e.g. the protein (or peptide) [39,55,56], lipid [57,58], coenzyme [59], and even virus [60] worlds are some of the most popular theories for the order and/or relevance of appearance of biomolecules on Earth.

Each prebiotic world generated invaluable insight on its own class of biomolecules, but also proposes a privileged, precursory function for them. This privileged position imposes a sequential, stepwise hierarchy on the OoL timeline: subsequent takeovers by the next world(s) until a living cell originates. Still, the question of which class of biomolecules initiated the OoL is a loaded question [61]. All known living cells contain DNA, RNA, proteins, lipids, coenzymes, and other metabolites—and the earliest cells as those known on Earth would have had to fulfil these minimal cell requirements [62]. There is a strong argument to be made for the emergence of essential biomolecules to have been (at least to some extent) contemporaneous and interdependent. More importantly, the

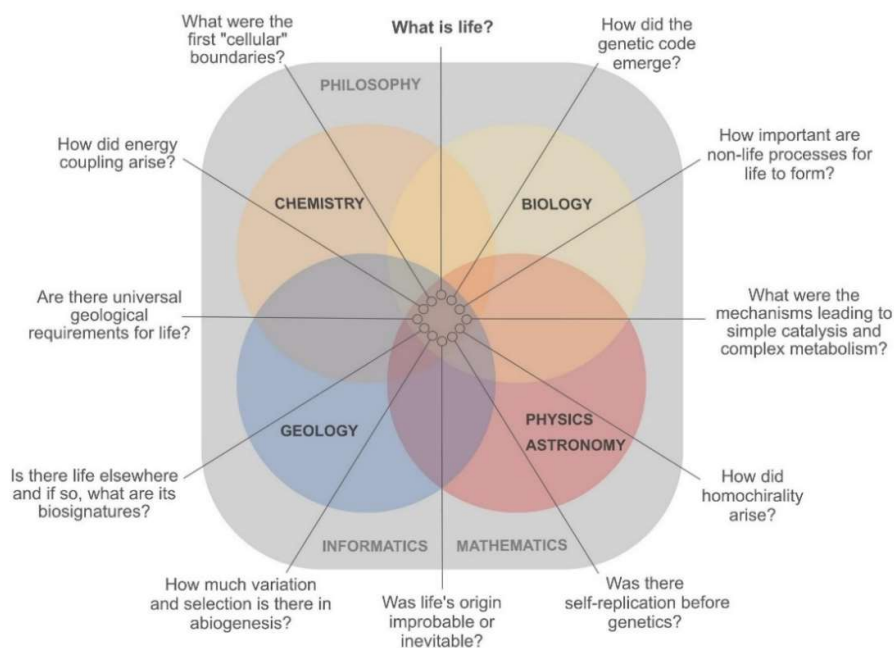
origin of biomolecules needs to be distinguished from the origin of cells, and life. Cells are not mere collections of their chemical components, but highly dynamic, complex systems with multiple interlocked processes involving those components. For this reason, the emergence of life cannot be distilled down to biomolecular retrosynthesis only. This appeal is not new, Kammaing asked biochemists to focus on processes rather than on pure retrosynthesis more than 30 years ago [54]. If common requirements between the origin of each biomolecule are found, they may provide necessary constraints for the OoL problem. Ultimately, for the emergence of biomolecules to be relevant for the OoL, it should be tightly linked with their participation in dynamic processes characteristic of life (e.g. replication, energy coupling and compartmentalization). These are some of the reasons why bridging different prebiotic worlds, with all their conceptual advances, is an urgent endeavour.

### 3. Building Bridges

#### 3.1. Pressing Questions in OoL are Interdisciplinary

Insights from classical approaches, hypotheses and ‘worlds’ have led to many advances, but they have also resulted in an ideological isolation that is possibly hindering progress in the OoL field [35,53,63–67]. A need to link different disciplines and approaches becomes evident looking at, in our view, the most central questions that should be addressed (Figure 1). These and other questions were identified before [68]. Only cooperation can push their answers forward—in both bottom-up and top-down directions. This is truly not an innovative line of thought. Aleksandr Oparin formulated it in 1924 in a comprehensive way:

*“A whole army of biologists is studying the structure and organization of living matter, while a no less number of physicists and chemists are daily revealing to us new properties of dead things. Like two parties of workers boring from the two opposite ends of a tunnel, they are working towards the same goal. The work has already gone a long way and very, very soon the last barriers between the living and the dead will crumble under the attack of patient and powerful scientific thought.”* [69]



**Figure 1.** Current central questions in origin of life (OoL) research. In order to find answers, each of the natural sciences illustrated has to play a role by converting these questions into hypotheses and theories, while constantly testing them experimentally. The fundamental roles of philosophy, mathematics and informatics are portrayed in the background.

Both Oparin and Haldane proposed independently integrative theories for the OoL in the 1920s [70]. Their ‘Heterotrophic origin of life theory’ was one of the first approaches to describe a slow transition between a non-living primordial ‘soup’ (Haldane) and living cells. Oparin was the first to integrate data from geochemistry, chemistry, planetary sciences and biochemistry in his seminal work proposing different prebiotic and biological stages [54,69]. His words, quoted above, now almost 100 years old, were optimistic. The “very, very soon” has not happened to this day—there is still a pressing need to bridge (or, in Oparin’s words “tunnel”) top-down and bottom-up approaches. Every attempt to reach over to “the opposite end of the tunnel” opens ways to new data and/or new data interpretations. Cross-disciplinary collaborations can lead to innovative discoveries when new information from discipline A challenges the status quo in discipline B [71]. This forces researchers to formulate new approaches and leads to progress. ‘Outsider information’ like that may help to find a middle ground for irreconcilable hypotheses or theories (see also “Section 4. Towards the Future”) [72].

The OoL question benefits greatly from crossovers of scientific disciplines, each of which brings its specific skills to the table. Chemistry supplies the knowledge of the construction kit, and physics, for instance, the energetic boundaries for the assembly of life [73,74]. Biology is confronted with bigger pictures, providing the roadmap from extant life to its origins via data collection and analysis (from multiple ‘omics’ technologies [75] to the study of elaborate microbial communities [76]). Geology and other geosciences help unfold the possible timeline and set the environmental constraints for the OoL (terrestrial and/or extraterrestrial) [77,78]. All of the above disciplines rely heavily on engineering, mathematics and informatics. Finally, philosophy provides a structure that allows scientists to ask the right questions following the rules of reasoning and argumentation, avoiding the traps of non-sequiturs and other logical fallacies [2,49].

An important first step in connecting disciplines is making the borders between them fully permeable. In the past 50 years, several interdisciplinary fields developed and paved the way for new research. For OoL in particular, astrochemistry, astrobiology and their search for biosignatures in the universe had a significant impact. These disciplines focused on physical and chemical processes that could lead to biotic or prebiotic systems on other planets [79] (Figure 1). One example is the discovery of subsurface oceans on icy moons, which extended the concept of habitable zone from the inner to the darker, colder places of the outer solar system [80]. Both complex, large insoluble organic material and low mass, reactive and soluble molecules, have been found in the subsurface ocean of Enceladus [81,82]. Large molecules have also been found in the atmosphere of Titan [83].

How does the need for discipline crossovers and permeability translate into practice? To illustrate this, we will use three central questions—all concerning essential cellular processes—as examples: (i) how did energy coupling arise, (ii) what mechanisms led to metabolism, and (iii) how did the genetic code emerge?

Life on Earth couples energy-releasing (spontaneous) reactions to energy-demanding (non-spontaneous) ones, capturing energy from its environment and eventually dissipating it as heat. This enables cellular processes such as growth and division. But how did this sophisticated system develop? Today, energy-coupling is mediated by enzymes which, acting as engines, funnel energy released from the cell’s diet into chemical energy. This energy is stored in a thioester linkage (as in acetyl-CoA), a phosphate-ester bond to carbon like in acetyl phosphate or a phosphate bond in the adenosine triphosphate (ATP) molecule [84]. These molecules are commonly known as energetic currencies in cells and mediate energy coupling by transferring energy between non-related biochemical processes. Are energetic currencies indispensable or exchangeable within existing metabolisms? To answer this biological insight on different species and pathways is crucial [85]. In order to find out if energy coupling was necessary for life’s first steps (still a subject of debate) several alternative energy currencies have been proposed and tested, for instance high-energy phosphates [86] or thioesters [87]—as it is known that cells use both, often sequentially (e.g. the production of ATP is sometimes mediated by thioesters). From there, further questions emerge: how did these energy-storing molecules become coupled and implemented into life as we know it (or: what was the order of entry of such heteroatoms as nitrogen, sulfur and phosphorus into proto-life’s chemistry)? Here, it would be helpful to look at different geochemical conditions to explore under which

conditions these couplings could be possible. Although the answers are still unclear [88], discovering them will be pivotal to understanding the transition between non-living and living systems [84,89].

The question of energy coupling is directly linked to the question of what geochemical mechanisms led to complex cellular metabolism. A systematic search for metabolism's beginnings (i.e. protometabolism) requires a clear and agreed-upon definition of metabolism, yet to be developed. This definition should start from the molecular level and the fundamental physical driving forces (kinetics vs. thermodynamics) [90], leading to combinations of chemical reactions, ultimately giving rise to a complex network [91,92], governed by a set of rules and boundaries [93,94]. Starting from known biochemical pathways one can deduce a conserved network that could sustain itself without enzymes [95–97]. The feasibility of such top-down analysis reducing the complexity of biological metabolic networks has to be tested in a bottom-up laboratory setting [98–100]. Conditions for these laboratory experiments can either be inspired by known geological environments, or a set of favourable conditions can be searched among existing geological settings and/or models. Such comparisons and convergences between biochemical and geochemical conditions, reactions and products can help to unveil environments where metabolic pathways and networks could emerge [101–103]. This applies not only to the early Earth but also other locations elsewhere in the universe like Enceladus [81,104]. Certainly, the same methodological rules will apply to other angles from which the OoL question is tackled: researching compartmentalisation, reconstructing the characteristics of LUCA, studying molecular replicators and precursors to genetics —hardly can we envision any of these succeeding in isolation.

Along with metabolism, life is based on another equally-important fundamental principle—inheritance, also described as “information that copies itself” [105]. As discussed (see “Section 2.2. One Origin, Abundant Worlds”), a better question than “which came first, information or metabolism?” would be: does the emergence of one have common properties with the other? In other words, perhaps we will know that acetate originated before ATP, but the emergence of metabolism and inherited information as complex systems was most likely interdependent and simultaneous. This directly leads to the fundamental question of how and when metabolism and information storage became linked. Nature's elegant solution is the genetic code, the origin of which remains a true enigma.

Different hypotheses exist for the origin of the genetic code, often depending on assumptions of the different hotly debated prebiotic conditions. Ribozymes show that RNA can harbour both catalysis and genetics, and there is support for the capability of RNAs to aminoacylate [106]. However, the code today is self-referential, that is, the mapping between amino acids and codons heavily relies on encoded proteins. When, why and how did non-coded peptides become involved? For the origin of the involvement of peptides in the genetic code, the conjecture that they are a superior catalyst is not enough, especially because they only have a supporting but not catalytic role in the ribosome [107]. Therefore, we must justify likely proximal selective advantages of using amino acids and simple peptides in the context of RNAs. These so-called ‘exaptations’ may bridge the RNA and peptide worlds, offering the basis on which later elements of the translational apparatus were built. Possible advantages of peptides include providing catalytic aid and expanding the catalytic repertoire of RNAs [107], membrane transport [108], scaffolding [109] and energy storage [110]. It has indeed been shown experimentally that non-coded peptides can potentiate the functions of RNA, which supports the coevolution of RNA and peptides [111]. Both the RNA and protein worlds ask whether a bipartite polymer setup (with nucleic acids and proteins) and the current genetic code is an inherent requirement of life, or if it is possible to envision one dominant polymer carrying information in prebiotic stages. Neither world alone can provide a clear explanation for the interlacing of the two, though, and there are relatively few endeavours to try out ‘messy emergence’, where some initial cooperation between non-coded proteins and RNAs was vital. However, regardless of the polymers at the origin of the code, another question urges: how did the code freeze to the current codon table? This question seemed for some time to be a mere statistical or even cryptographic problem [112], and a variety of explanations emerged to solve it, most popularly: stereochemical basis for the assignment between nucleic acids and amino acids [113], the

development of the code guided by the biosynthetic pathways of amino acids [114], and optimization in order to reduce the severity of mutations [115]. These are not mutually exclusive hypotheses, and the origin of the code might have simultaneously involved several of them [116]. Most likely, the development of the genetic code took place in a continuous expansion [117], a hypothesis supported by functional proteins with reduced amino acid repertoires [118]. The later expansion of the repertoire was possibly governed by a physicochemical optimization that fits well with water-based biochemistry [119].

Answers on the origin of the code still seem very distant [115,116,120–122]. In order to finally solve this puzzle it will be necessary to investigate not only the emergence of biological information [120], but also the evolution of interactions between the molecules involved in translation [117,123].

### 3.2. On the Right Track? Looking at the Past Decade

In the past 10 years, many have worked and asked for OoL research to unite [25,39,124,125]. Here we look at experimental examples from the last decade that connect different disciplines, theories or interpretations (Figure 2).

One of the first barriers to come down stood between the decades-old views of different single-biomolecule worlds (see “Section 2.2. One Origin, Abundant Worlds”). Studies merging the lipid world with others are pioneers. The requirement for compartmentalization to keep genomic molecules and their products spatially together, as well as to allow for vectorial (bio)chemistry, suggests that the potential of lipids and other amphiphilic structures to self-assemble into micelles and bilayer vesicles within an aqueous phase constituted a critical step in the emergence of life [57,126,127]. Vesicles were thought to be stable only in salt-poor aqueous environments, such as surficial freshwater ponds or hydrothermal springs, but recent work showed that these become much more resilient to extreme salinity and pH if they are composed of mixtures of amphiphiles [128], a feature which better reflects the naturally messy aspect of prebiotic chemistry. Such prebiotic compartments, also referred to as ‘protocells’, are defined as primitive precursors of modern cells which, although not yet alive, exhibited essential cellular characteristics [129]. Efforts of OoL researchers aim to establish various in vitro protocell models that mimic key features of life as we know it, including simple metabolic pathways, replication or vesicle growth and division [130]. A protometabolism leading to sugar synthesis has been assembled within lipid vesicles, with the final products diffusing through the lipid barrier and being detected by living bacteria [131]. Not long after, DNA amplification was shown to induce growth and division of lipid vesicles, linking the reproduction of an informational substance with that of a compartment [132]. These demonstrate that simpler systems than cells can model fundamental interactions between membranes and their contents. Another example is the bridge between the metabolism-first and the RNA world theories [114]. It is now clear that the building blocks for RNA and DNA are intermediates of metabolic networks; they are never directly uptaken from the environment in their ready-for-polymerization forms (i.e. as phosphorylated nucleosides), but as unphosphorylated biogenic nucleosides [133], and are also in fact moieties of several essential cofactors [134]. The idea of the simultaneous and interdependent origins for RNA and DNA’s building blocks has been given experimental evidence [135]. But bridges between the RNA world and metabolism-first can be built beyond their typical molecules. Classical approaches in OoL have often been constrained to biomolecules due to their ubiquity in biology, however, most of these biomolecules were not necessarily available at early prebiotic stages [136]. Prebiotic environments most likely included compounds not central to modern biopolymers, for example, alpha hydroxy acids (aHA) [137] among numerous others [138,139]. Recent work has shown that aHAs easily form combinatorial polyesters under wet-dry conditions that may have played a role in the catalytic landscape within which they were formed [140]. These polyesters can form membraneless compartments mimicking a cell, capable of differentially segregating various kinds of dyes, hosting a protein and even accumulating lipids in their exterior [141]. If one considers that the accumulation of biomolecules is a biological invention, optimizing for certain geochemical and/or biochemical properties [119], then compounds such as aHAs may have played a role in some form of nascent biology not conserved in modern biology [142]. Considering the

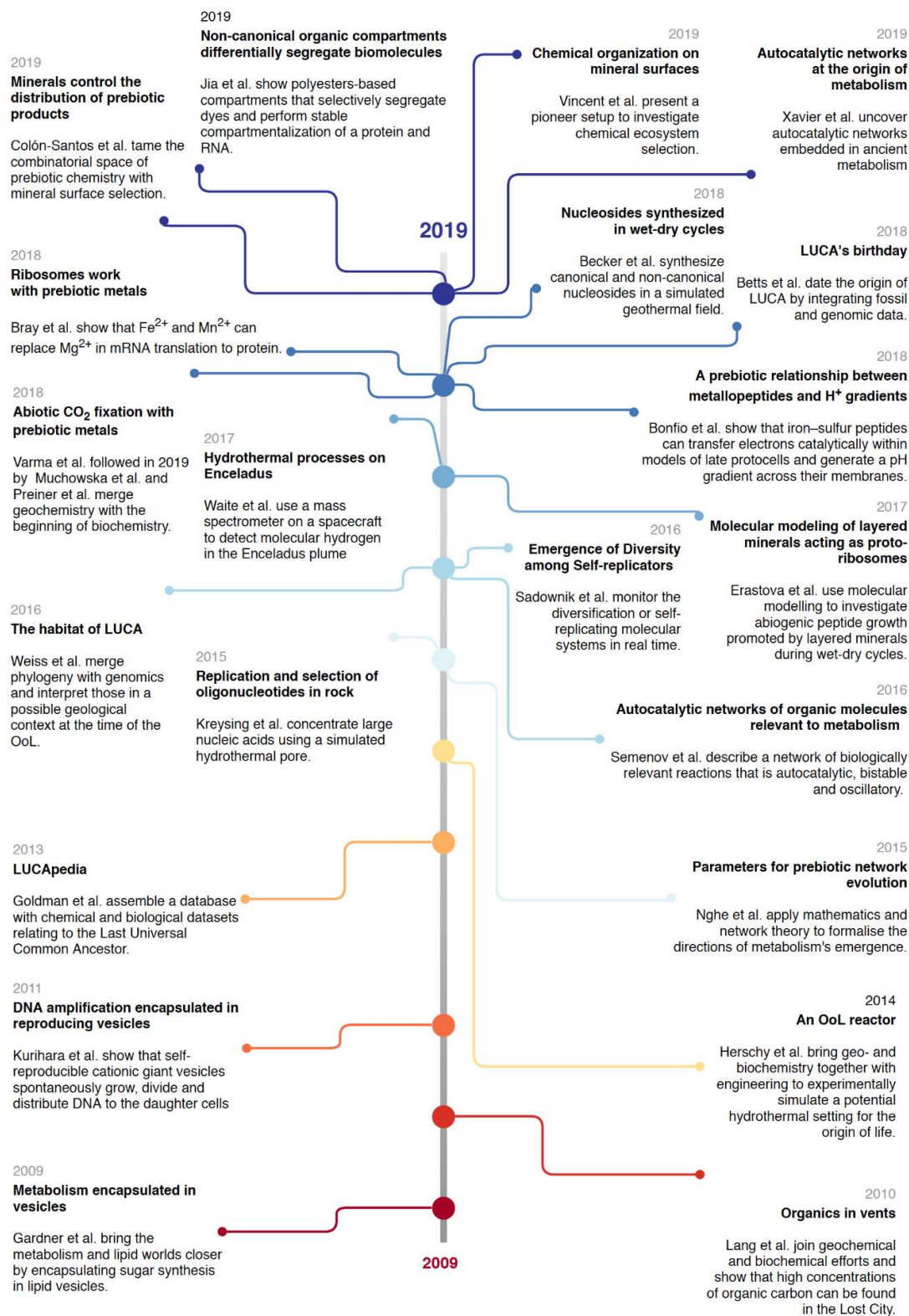


'messy' nature (meaning the inherent diversity) of prebiotic chemistry [143,144], this may have been the case, and the utility of central biomolecules such as RNA and ubiquitous metabolites to elucidate the OoL may be partial, despite their omnipresence in modern life.

Other striking examples for integration in OoL research come from geological studies providing a better picture of the Hadean world, including the likely atmospheric composition and the depths of oceans and tectonic activity [145,146]. These studies constrained and brought closer the work of both biologists and chemists. Experimentation with origins in hydrothermal vents [147–149] and geothermal fields [150–152] have taken into consideration geological insights in their OoL scenarios. In particular, the importance of relevant metals and metal clusters in the Hadean has settled in experimental work that recreates the origin of biochemistry *in vitro*, including carbon cycling [98,103,153], nitrogen fixation [154], ribosomal translation [155], and even the generation of pH gradients [156]. Simulated hydrothermal conditions, in particular pores, select for the replication of longer oligonucleotides [157]. Mineral surfaces have shown promising features in promoting biochemistry, including the selection of longer RNA molecules [100], and a variety of organic reactions including nitrogen reduction, lipid self-organization, condensation-polymerization reactions, selection and concentration of amino acids and sugars and chiral selection (see [158] and references therein). Surfaces can also help to tame the combinatorial space in a complex system of organic molecules [100,159].

By providing new ways to handle vast amounts of data in a systematic and quantitative manner, advances in computer sciences and technologies help the much-required aforementioned integrations that are starting to reflect on OoL research. Larger chemical libraries can now be monitored over a manageable timeline, and the diversification of self-replicating molecules has been observed in such systems [160]. The self-replication of small organic molecules has also been observed in an autocatalytic process displaying complex, non-linear responses to changes in environmental conditions [161]. Computational models allow us to test hypotheses unattainable in a laboratory alone, be it chemical conditions [22], timescales [4] or other routine perturbations to the model. Computational biology has now moved to mathematical simulations at increasing levels of complexity, from topological, interaction-based to constraint- and mechanism-based [162]. Recently, a database aggregating a variety of bioinformatic approaches to LUCA was compiled, allowing for testing hypotheses in investigations of ancient biochemistry [163]. Simulations can now be undertaken with geology and biology in mind, both in molecular dynamics [22] and network biology [95,96], identifying constraints and parameters that can provide direction for experimental work [98,159,164].

The progress of top-down approaches from traditional comparative genomics to more integrative approaches revives optimism in the quest for LUCA [165]. Integrating fossil data with molecular evolutionary clocks has solidified evidence for the age of LUCA at >3.9 Ga [4]. Taking into account morphological characteristics together with genetic signatures is another promising direction. Recent analyses of membranes and cell walls of prokaryotes suggest that LUCA was able to sporulate, that is, to reproduce into a dormant, non-metabolising cell that could survive long periods of harsh conditions such as the late heavy bombardment [146,166]. Modern phylogenetic considerations minimize the effect of lateral gene transfer in reconstructing the first genomes, pointing to a thermophilic autotrophic LUCA [149].



**Figure 2.** Timeline of recent multidisciplinary achievements that build bridges in OoL research. Included are examples from the past 10 years of OoL research that bridge disciplines, approaches and/or methods, biomolecules/single-world scenarios, simulations and experiments. The choice of studies does not aim to cover (exclusively) novel findings, but those that build bridges.

## 4. Towards the Future

### 4.1. General Remarks

We expect by now to have presented a convincing argument for the necessity of bridges in answering OoL's most prominent questions (Figure 1), together with examples of successful studies (Figure 2). How can the future hold more interconnections between approaches and theories? Some tools are general for all types of linking between seemingly disconnected research (Box 1). These include interdisciplinary conferences, where the focus lies not only on one molecule or one theory, but either on the whole field or on the question (Figure 1) sought to be answered. Interdisciplinary higher education and scientific collaborations should progress in the same direction in the coming decades, with more focus on questions and holistic pictures.

**Box 1:** Tools for future integration of OoL research

General		Specific
<ul style="list-style-type: none"> <li>• Interdisciplinary conferences</li> <li>• Consistent nomenclature</li> <li>• Sharing experimental designs</li> <li>• Public communication</li> <li>• Collaborations</li> <li>• Interdisciplinary higher education</li> <li>• Debate/communication training</li> <li>• Incentives for early-career researchers</li> <li>• Careful prevention of conflicts of interest</li> </ul>	Experiments <i>in vitro, in vivo, in silico</i>	<ul style="list-style-type: none"> <li>• User-friendly informatic tools</li> <li>• Better databases</li> <li>• Promote reproducibility</li> <li>• Create interface spaces for easier data exchange (between computational and experimental)</li> <li>• Writing with broad audience in mind</li> </ul>
	Theories	<ul style="list-style-type: none"> <li>• Explore both commonalities and differences</li> <li>• Explore particular parts of an alternative theory, specially the testable ones</li> <li>• Promote either double-blind or double-open peer-review of articles and grants, to reduce personal and institutional bias</li> </ul>

Regarding the limits of experimentation, computational tools can extrapolate from experimental input, and thus will play a major role in untangling complexity in OoL. However, research *in vitro*, *in vivo* or *in silico* (Box 1) needs to be attainable to everyone. This requires more user-friendly computational tools, the communication of tools and results in an accessible manner, and the creation of curated and moderated spaces where data and tools can be exchanged, faster than by the traditional scientific paper (as forums, curated databases, websites and discussion channels). Standards that can be used by different disciplines are heavily required, particularly in chemical, biochemical and genetic nomenclature—we recommend Chebi [167], KEGG [168] or BiGG [169]. Nomenclature in theoretical biology also requires attention. In particular, nomenclature regarding the first biological entities includes a great amount of confusing variability, e.g. an initial Darwinian ancestor (IDA) [36] akin to a first universal common ancestor, FUCA [37], and the more commonly used LUCA, akin to the urancestor [170], to the cenancestor [171], and to the (simpler) universal common ancestor [172]. Here, we recommend, merely based on the greater usage to this day, the use of IDA and LUCA. Semantic variability without clear conceptual justification does not ease literature review and theoretical studies, and a strong community effort is required to make different names clearly distinguishable or for standards to be adopted.

Sharing and combining experimental setups (both computational and laboratorial), can help to increase reproducibility but also to connect disciplines and approaches. For example, to narrow the current gap between prebiotic (geo)chemistry and ancient biochemistry, simulation efforts could focus on the least contested facts on the early Earth. These include a CO<sub>2</sub>-rich atmosphere and hydrosphere from volcanism, and an H<sub>2</sub>-rich hydrothermal outflow from the crust. Other parameters which are harder to characterize, as pH ranges, pressure, access to ultraviolet (UV) radiation, water activity, etc. could then be varied. There are means to attempt such simulations in the laboratory today, where collecting vast amounts of data is possible [173]. This data can be analysed and incorporated into the computational models which then, in turn, can shed more light on what laboratory experiments should be attempted. Hereby, the steady bilateral exchange between

disciplines is crucial. Not only should geology provide conditions for experiments [174], those experiments should also provide the attributes that geochemical studies could search for.

#### 4.2. Commonalities between Opposing Theories

Bridges between methodologies will allow the OoL field to advance, but ultimately theories will have to become more interconnected as well. We foresee that this can happen with an intentional focus on commonalities rather than solely on differences, and plenty of examples can be mentioned. Here we choose to focus on four of the most heated debates, heavily interrelated: (i) the geological setting for the OoL, (ii) the source of food/starting molecules, (iii) the source of energy and (iv) the RNA world vs. metabolism-first division.

##### 4.2.1. The Geological Setting

The proposed and heavily discussed locations for the OoL, including terrestrial hot springs [150], submarine hydrothermal vents [175,176], volcanic landscapes [177], hydrothermally-percolated sediments [178] and warm little ponds [179] create some of the strongest disagreements in OoL research, but have in common several requirements. These include a constant energy and CHNOPS source, heterogeneous or homogeneous catalysts, non-equilibrium conditions and conditions that enable aggregation of molecules to form life's building blocks and, eventually, polymerization of said building blocks [22,103,152,180]. One example is the requirement for wet–dry cycles for polymerization, mostly discussed in the context of geothermal fields [180], even though wet-dry cycles can also be found in hydrothermal pores [181] and even at the water-air interface [20]. On another note, some propose that different geological settings could have created different biomolecules that were then able to meet in an unchanged chemical state in the location where cells ultimately emerged [182].

##### 4.2.2. The Food Source

A historically important division drawing from the analogy with different types of metabolism in prokaryotes [125] is the heterotrophy vs. autotrophy conundrum. Each side has envisioned their abiogenic scenario from a trophic point of view: if the carbon source for organosynthesis was eminently inorganic (usually CO<sub>2</sub>), this represented an autotrophic origin (e.g. [183]), whereas if the carbon source were reduced organic compounds, a heterotrophic one (e.g. [184]). This division usually links to whether the earliest cells are considered autotrophic (e.g. [149]) or heterotrophic (e.g. [185]). It is important to note that for the first cells, the question is indeed meaningful—those cells either imported C<sub>1</sub> compounds, subsequently reducing them and forming C–C bonds (autotrophy) or imported C<sub>n</sub> compounds and used those directly in biosynthesis [186,187]. But others are wary that this division may be unhelpful before the cellular stage (see Smith and Morowitz in [68]), since trophic types apply poorly to biosynthetic pathways. All organic molecules synthesised non-enzymatically ultimately derive from an inorganic carbon source (C, CO or CO<sub>2</sub>) which was subsequently reduced. The endogenous (i.e. terrestrial origin) vs. exogenous (i.e. extraterrestrial origin) debate also entails limitations, by simply referring to the physical origin of the organic molecules, but not informing on the chemical relationship between these and the nascent proto-biochemistry. For the origin of the first biochemical networks, potentially acellular IDAs, we propose referring instead to whether all or some of the reactions leading to organic molecules occurred in situ (i.e. within, or physically contiguous to the nascent proto-biochemistry) or ex situ (i.e. physically separated from it). Note that this is not just a geographical distinction. Instead, it aims to find out whether the nascent chemical network (eventually leading to living cells) could affect (positively and/or negatively) the chemical reactions feeding it. This, we believe, is a meaningful change in language aiming to mend the limitations associated with the autotrophic vs. heterotrophic dilemma. Therefore, the in situ synthesis implies the reactions yielding biomolecules as part of the nascent protometabolism. In contrast, in the ex situ subtype the synthetic reactions rest unaffected by the

nascent chemical network, and the latter needs to invent alternative chemical pathways to eventually become independent from the former's unidirectional supply.

#### 4.2.3. The Energy Source

Non-equilibrium conditions suggest an important bridge that can be established between most OoL hypotheses [14,188]. The fact that life is a process that is by definition not in equilibrium is not debatable in physics, biology or any other of the involved disciplines [189,190]. One example of how 'disequilibration' can be observed in life as we know it, is that life builds up complexity and then breaks it down again—in biology, this is, in broad strokes, anabolism and catabolism. These complex processes in general parallel energy spending and gaining in the cell, respectively. However, some organisms (e.g. methanogens) are known to conserve energy only through anabolism [191]. The earliest life would have had to couple both the constructive and destructive regime in some way [188,192]. To fulfil their energetic requirements, prebiotic systems would need both electron donors and acceptors supplied by their environment. The source of energy for the earliest life has been strongly disputed, with plausible hypotheses ranging from pH and redox gradients to thermal energy and UV light (for a detailed discussion see [74,193,194]). It is, however, not unlikely that several energy sources played a role in different prebiotic stages. Nevertheless, other constraints than the primary energy sources are imposed as soon as more complex prebiotic systems arise. A cell-like energy-coupling system could only be persistent over time if it can be inherited, thus forging a necessary link between the offspring hypotheses of the classical 'metabolism-first' and 'genetic-first' worlds [194].

#### 4.2.4. RNA World versus Metabolism-First

Modern thinking on the OoL highlights the need for not only the synthesis of life's building blocks (themselves crucial goals), but also the reenactment of the processes they participate in [195]. Thankfully, what follows from this realization is that metabolism- and genetics-oriented hypotheses are gradually ceasing to be seen as mutually exclusive, because cellular processes involve both metabolites and genetic molecules [25,124]. A particular example regards evolution. Looking for the origin of natural selection and evolution is to look for the origin of an inheritance system that must maintain the capacity to produce new combinations, while keeping fidelity in information transfer [196,197]. But what *is* this information? Biological information is more than a collection of bits; it has contextuality, translated into functionality and is prone to evolve, as languages do [198]. In other words, genetic molecules only hold information in the right context: when they can be recognized and translated to functions that keep the system going. This definition opens different avenues to look for information storage and transmission during chemical evolution.

The earliest information could have been made from the types and quantities of molecules within a self-sustaining chemical assembly, also called 'composome' [199] or autocatalytic network [200]. This relates to the emphasis by some recent OoL research on the co-dependence of metabolism, containment and replication/information (see "Section 3.2. On the Right Track: Looking at the Last Decade"). This emphasis is central to bridging the RNA world and metabolism-first theories and points to the origin of molecular coordination or cooperation. Following this line of thought, some authors propose early 'selfish cooperators', akin to viruses, which formed stable ensembles of co-inherited genetic elements [60]. These ensembles should have been able to perform both a kind of proto-replication and a kind of proto-metabolism [201]. Still, there is a 'cooperation barrier' in the transition of non-life to life, caused by (i) molecules that could cooperatively contribute to the success of an ensemble but which are often not supported by the ensemble, and (ii) side reactions or processes that undermine cooperation [202]. To overcome this barrier, the management of those otherwise unconstrained ensembles is required [202]. This hypothesis is particularly interesting because it exposes the necessity of a digitally-coded (genetic, on/off) management of the analog (continuous) reactions of metabolism in order to overcome the cooperation barrier efficiently, reflecting the two-tiered structure of all known living cells [198,202].

If such analog information storage mechanisms as autocatalytic networks are able to undergo Darwinian evolution is a matter under debate [93,203]. Several have suggested fitness criteria for

natural chemical selection, including the rate of entropy production [204] and other kinetic or thermodynamic features of chemical reactions, such as stoichiometric catalysis, autocatalysis, and cooperativity [205]. A specific example is the greater reactivity of proteinaceous amino acids when compared to their non-proteinaceous counterparts, which naturally selects from oligomers of the former [206]. At some point, most likely quite early during the development of prebiotic systems, genetic molecules did crystallize, and information storage became based on molecular recognition. Every letter of the ‘genetic alphabet’, A, C, G, T and U, given a recognition/translation system, contains information. Early molecular recognition could have been significantly different to contemporary DNA and RNA [207,208] and even minerals have been proposed as information storage molecules [209].

What we know now, is that if we wish to search for answers on the origin of life as we know it—meaning the origin of cells—we are faced with an increasing conceptual [2] and experimental complexity [210] that will require the integration of RNA molecules with peptides, lipids, and protometabolism. Better models of complex biological systems [100], as well as better techniques to characterise them [211], will help to tackle such complexity. However, this complexity can no longer be denied. In other words, both the RNA-world and metabolism-first theories aim at the origin of life as we know it. However, we do not know it without each and both of them.

## 5. Conclusions

The main reason the OoL research field is still divided on so many issues is that it seems virtually impossible to find definite answers for all of our questions—we cannot wait hundreds of thousands of years to observe in real-time if a certain geological setting trumps another or which are the first enzymes to develop. We can only approach solutions asymptotically and, returning to the image of an OoL mosaic, add one pebble at a time, one insight that brings us closer to a more complete picture. In this article, we tried to convey how this is already being done—and hopefully, can be done even better in the future. Just as life itself is a synergistic process, in the sense that life’s biological, physical and chemical properties are intertwined, the research that tries to explain its origin has to be synergistic as well. No discipline or approach should be dismissed, as long as its claims are evidence-based. However, we should keep in mind that while single results have dominated the field in the past, these can never constitute a solution to such a complex problem. Synthesising biomolecules, for example, although informative, cannot be enough to tackle the OoL, because life is not a mere collection of biomolecules, but rather a dynamic process involving biomolecules within boundary conditions interacting with its environment. Modern thinking on origins will thrive from recreating or imitating processes, rather than focusing only on the prebiotic synthesis of biomolecules.

Finally, we must highlight bias in OoL experiments, starting from how to define ‘plausible’ prebiotic conditions to which biomolecules are more central than others [212]. It is imperative to overcome personal biases in order to make progress. To get there, we need to increase the exchange, openness and respect between all those involved despite the inherent competition and inherited bias towards different hypotheses or approaches in research these days. The bridges we propose are essential, not because scientific research must be homogeneous, but because heterogeneity must be articulated (for examples see “Section 3. Building Bridges”). This articulation will change both the way research is done but also communication (in published articles, conferences and others). At this point, it is important to emphasize how crucial classical approaches in OoL research were and still are. However, while pivotal insights were achieved in classical studies, a more complete OoL narrative can only unfold by building bridges between them. This might all appear far too obvious to mention, but the reality in origin of life research shows that it cannot be mentioned often enough.

**Author Contributions:** Conceptualization, M.P. and J.C.X.; investigation, M.P., S.A., S.B., H.C.B., A.B., E.C., K.C., V.E., S.G.G., N.K., G.K., R.M., G.M., K.B.M., S.N., B.P., E.P., Á.R., D.R., A.S., N.M.S., F.L.S., F.D.K.T., D.V., J.C.X.; writing—original draft preparation, M.P. and J.C.X.; writing—review and editing, M.P., S.A., S.B., H.C.B., A.B., E.C., K.C., V.E., S.G.G., N.K., G.K., R.M., G.M., K.B.M., S.N., B.P., E.P., Á.R., D.R., A.S., N.M.S., F.L.S., F.D.K.T., D.V., J.C.X., in particular K.B.M. and E.C. on metabolism and energy, S.B. and A.R. on the genetic code and information, K.C. on messy chemistry, B.P. on protocells, G.K. on the philosophy of OoL, F.L.S. and F.T. on the

top-down approach, H.B. on the time frame of the OoL and N.K. on astrobiology; visualization, M.P. and J.C.X.; supervision, J.C.X.; project administration, M.P. and J.C.X.; funding acquisition, M.P. and J.C.X. All authors have read and agreed to the published version of the manuscript.

**Funding:** M.P. and J.C.X. are funded by grants from the Deutsche Forschungsgemeinschaft (MA-1426/21-1), the European Research Council (666053) and the Volkswagen Foundation (93046) to William F. Martin. E.C. is thankful to the support from the Origins Center through NWA StartImpuls. K.C. is supported by the European Structural and Investment Funds Operational Programme “Research, Development and Education”-funded project “ChemJets” (C.Z. 02.2.69/0.0/0.0/16\_027/0008351) and Research Encouragement Fund UKM (GGP-2019-029). F.L.S. gratefully acknowledges funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement 803768). F.L.S., S.N. and E.P. acknowledge the Wiener Wissenschafts-, Forschungs- und Technologiefonds (grant agreement VRG15-007, awarded to F.L.S.). D.V. and Á.R. were supported by the National Research, Development and Innovation Office (NKFIH) under the grant number GINOP 2.3.2-15-2016-00057. A.S. and B.P. were supported by the CRC 235 Emergence of Life granted to Dieter Braun and Petra Schwille, respectively. H.C.B. was supported by the NERC GW4 PhD studentship. N.K. was supported by European Research Council (ERC) Consolidator Grant 724908-Habitat OASIS.

**Acknowledgements:** All authors wish to thank William F. Martin (Heinrich Heine University, Düsseldorf) for generously sponsoring the entirety of the 1st Interdisciplinary Origin of Life Meeting for early career researchers in Düsseldorf, where many insightful discussions motivated this paper. The authors are thankful to Caroline Ring for the graphic recording of the entire meeting, available in <https://www.molevol.hhu.de/iool-meeting-bilder.html>. M.P. and J.C.X. thank William F. Martin for mentoring and support. S.B. thanks Thomas Carell for mentoring and support. K.B.M. gratefully acknowledges mentoring and support from Joseph Moran (Institut de Science et d’Ingénierie Supramoléculaires, Strasbourg). S.A. acknowledges mentoring and support by Leroy Cronin (School of Chemistry, University of Glasgow). S.N. and E.P. thank Filipa L. Sousa for mentoring and support. D.R. acknowledges Sheref S. Mansy for mentoring and support. E.C. acknowledges mentoring and support by Inge Loes ten Kate (Origins Center, Department of Earth Sciences, Utrecht University).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Gayon, J.; Malaterre, C.; Morange, M.; Raulin-Cerceau, F.; Tirard, S. Defining Life: conference proceedings. *Origins. Life Evol. B.* **2010**, *40*, 119–120.
2. Mariscal, C.; Barahona, A.; Aubert-Kato, N.; Aydinoglu, A.U.; Bartlett, S.; Cárdenas, M.L.; Chandru, K.; Cleland, C.; Cocanougher, B.T.; Comfort, N.; et al. Hidden Concepts in the History and Philosophy of Origins-of-Life Studies: A Workshop Report. *Origins. Life Evol. B.* **2019**, *49*, 111–145.
3. Javaux, E.J. Challenges in evidencing the earliest traces of life. *Nature* **2019**, *572*, 451–460.
4. Betts, H.C.; Puttick, M.N.; Clark, J.W.; Williams, T.A.; Donoghue, P.C.J.; Pisani, D. Integrated genomic and fossil evidence illuminates life’s early evolution and eukaryote origin. *Nat. Ecol. Evol.* **2018**, *2*, 1556–1562.
5. Czárán, T.; Könyű, B.; Szathmáry, E. Metabolically Coupled Replicator Systems: Overview of an RNA-world model concept of prebiotic evolution on mineral surfaces. *J. Theor. Biol.* **2015**, *381*, 39–54.
6. Davies, P.C.W.; Benner, S.A.; Cleland, C.E.; Lineweaver, C.H.; McKay, C.P.; Wolfe-Simon, F. Signatures of a shadow biosphere. *Astrobiology* **2009**, *9*, 241–249.
7. McLendon, C.; Opalko, F.J.; Illangkoon, H.I.; Benner, S.A. Solubility of polyethers in hydrocarbons at low temperatures. A model for potential genetic backbones on warm titans. *Astrobiology* **2015**, *15*, 200–206.
8. Wöhler, F. Ueber künstliche Bildung des Harnstoffs. *Ann. Phys. Chem.* **1828**, *87*, 253–256.
9. Butlerow, A. Bildung einer zuckerartigen Substanz durch Synthese. *Liebigs. Ann. Chem.* **1861**, *120*, 295–298.
10. Breslow, R. On the mechanism of the formose reaction. *Tetrahedron Lett.* **1959**, *1*, 22–26.
11. Miller, S.L. A production of amino acids under possible primitive earth conditions. *Science* **1953**, *117*, 528–529.
12. Miller, S.L.; Urey, H.C. Organic compound synthesis on the primitive earth. *Science* **1959**, *130*, 245–251.



13. Hargreaves, W.R.; Mulvihill, S.J.; Deamer, D.W. Synthesis of phospholipids and membranes in prebiotic conditions. *Nature* **1977**, *266*, 78–80.
14. Ruiz-Mirazo, K.; Briones, C.; de la Escosura, A. Prebiotic systems chemistry: new perspectives for the origins of life. *Chem. Rev.* **2014**, *114*, 285–366.
15. Kitadai, N.; Maruyama, S. Origins of building blocks of life: A review. *Geosci. Front.* **2018**, *9*, 1117–1153.
16. Attwater, J.; Raguram, A.; Morgunov, A.S.; Gianni, E.; Holliger, P. Ribozyme-catalysed RNA synthesis using triplet building blocks. *eLife* **2018**, *7*.
17. Horning, D.P.; Joyce, G.F. Amplification of RNA by an RNA polymerase ribozyme. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 9786–9791.
18. Rajamani, S.; Vlassov, A.; Benner, S.; Coombs, A.; Olasagasti, F.; Deamer, D. Lipid-assisted synthesis of RNA-like polymers from mononucleotides. *Origins. Life Evol. B.* **2008**, *38*, 57–74.
19. Mamajanov, I.; MacDonald, P.J.; Ying, J.; Duncanson, D.M.; Dowdy, G.R.; Walker, C.A.; Engelhart, A.E.; Fernández, F.M.; Grover, M.A.; Hud, N.V.; et al. Ester Formation and Hydrolysis during Wet–Dry Cycles: Generation of Far-from-Equilibrium Polymers in a Model Prebiotic Reaction. *Macromolecules* **2014**, *47*, 1334–1343.
20. Griffith, E.C.; Vaida, V. In situ observation of peptide bond formation at the water-air interface. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 15697–15701.
21. Rodriguez-Garcia, M.; Surman, A.J.; Cooper, G.J.T.; Suárez-Marina, I.; Hosni, Z.; Lee, M.P.; Cronin, L. Formation of oligopeptides in high yield under simple programmable conditions. *Nat. Commun.* **2015**, *6*, 8385.
22. Erastova, V.; Degiacomi, M.T.; Fraser, D.; Greenwell, H.C. Mineral surface chemistry control for origin of prebiotic peptides. *Nat. Commun.* **2017**, *8*, 2033.
23. Gibson, D.G.; Benders, G.A.; Andrews-Pfannkoch, C.; Denisova, E.A.; Baden-Tillson, H.; Zaveri, J.; Stockwell, T.B.; Brownley, A.; Thomas, D.W.; Algire, M.A.; et al. Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* **2008**, *319*, 1215–1220.
24. Cello, J.; Paul, A.V.; Wimmer, E. Chemical synthesis of poliovirus cDNA: Generation of infectious virus in the absence of natural template. *Science* **2002**, *297*, 1016–1018.
25. Walker, S.I.; Packard, N.; Cody, G.D. Re-conceptualizing the origins of life. *Philos. Trans. A Math. Phys. Eng. Sci.* **2017**, *375*.
26. Locey, K.J.; Lennon, J.T. Scaling laws predict global microbial diversity. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 5970–5975.
27. Milo, R. What is the total number of protein molecules per cell volume? A call to rethink some published values. *BioEssays* **2013**, *35*, 1050–1055.
28. Mayr, E. Cause and effect in biology. *Science* **1961**, *134*, 1501–1506.
29. Fry, I. The origins of research into the origins of life. *Endeavour* **2006**, *30*, 24–28.
30. Cleland, C.E. Pluralism or unity in biology: could microbes hold the secret to life? *Biol. Philos.* **2013**, *28*, 189–204.
31. Mushegian, A.R.; Koonin, E.V. A minimal gene set for cellular life derived by comparison of complete bacterial genomes. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 10268–10273.
32. Koonin, E.V. Comparative genomics, minimal gene-sets and the last universal common ancestor. *Nat. Rev. Microbiol.* **2003**, *1*, 127–136.
33. Lagesen, K.; Ussery, D.W.; Wassenaar, T.M. Genome update: the 1000th genome—a cautionary tale. *Microbiology* **2010**, *156*, 603–608.



34. Sutherland, J.D. The Origin of Life--Out of the Blue. *Angew. Chem. Int. Ed Engl.* **2016**, *55*, 104–121.
35. Ralser, M. An appeal to magic? The discovery of a non-enzymatic metabolism and its role in the origins of life. *Biochem. J* **2018**, *475*, 2577–2592.
36. Yarus, M. Getting past the RNA world: the initial Darwinian ancestor. *Cold Spring Harb. Perspect. Biol.* **2011**, *3*.
37. Prosdocimi, F.; José, M.V.; Farias, S.T. de Be Introduced to the First Universal Common Ancestor (FUCA): The Great-Grandmother of LUCA (Last Universal Common Ancestor). **2018**. DOI: 10.20944/preprints201806.0035.v1. Available online: <https://europepmc.org/article/ppr/ppr49297> (accessed on: 25th February 2020)
38. Lahav, N.; Nir, S.; Elitzur, A.C. The emergence of life on Earth. *Prog. Biophys. Mol. Biol.* **2001**, *75*, 75–120.
39. Ikehara, K. Evolutionary Steps in the Emergence of Life Deduced from the Bottom-Up Approach and GADV Hypothesis (Top-Down Approach). *Life* **2016**, *6*.
40. Vicens, J.; Vicens, Q. Emergences of supramolecular chemistry: from supramolecular chemistry to supramolecular science. *J. Incl. Phenom. Macrocycl. Chem.* **2011**, *71*, 251–274.
41. Rich, A. On the problems of evolution and biochemical information transfer. In *Horizons In Biochemistry*; Kash, M., Pullman, B., Eds.; Academic Press: New York, NY, USA, 1962; pp. 103–126.
42. Gilbert, W. Origin of life: The RNA world. *Nature* **1986**, *319*, 618–618.
43. Kruger, K.; Grabowski, P.J.; Zaug, A.J.; Sands, J.; Gottschling, D.E.; Cech, T.R. Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* **1982**, *31*, 147–157.
44. Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **1983**, *35*, 849–857.
45. Robertson, M.P.; Joyce, G.F. The origins of the RNA world. *Cold Spring Harb. Perspect. Biol.* **2012**, *4*.
46. Orgel, L.E. Prebiotic chemistry and the origin of the RNA world. *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 99–123.
47. Shapiro, R. A Replicator Was Not Involved in the Origin of Life. *IUBMB: Life* **2000**, *49*, 173–176.
48. Oivanen, M.; Kuusela, S.; Lönnberg, H. Kinetics and Mechanisms for the Cleavage and Isomerization of the Phosphodiester Bonds of RNA by Brønsted Acids and Bases. *Chem. Rev.* **1998**, *98*, 961–990.
49. Wächtershäuser, G. The origin of life and its methodological challenge. *J. Theor. Biol.* **1997**, *187*, 483–494.
50. Dyson, F. *Origins of Life*, Cambridge University Press: Cambridge, UK, 1999.
51. de Duve, C. A Research Proposal on the Origin Of Life. Closing Lecture given at the ISSOL Congress in Oaxaca, Mexico, on July 4, 2002. *Origins Life Evol. B.* **2003**, *33*, 559–574.
52. Smith, E.; Morowitz, H.J. Universality in intermediary metabolism. *Proc. Natl. Acad. Sci. U.S.A* **2004**, *101*, 13168–13173.
53. Orgel, L.E. The implausibility of metabolic cycles on the prebiotic Earth. *PLoS Biol.* **2008**, *6*, e18.
54. Kamminga, H. Historical perspective: the problem of the origin of life in the context of developments in biology. *Origins. Life Evol. B.* **1988**, *18*, 1–11.
55. Caetano-Anollés, G.; Wang, M.; Caetano-Anollés, D.; Mitterthaler, J.E. The origin, evolution and structure of the protein world. *Biochem. J.* **2009**, *417*, 621–637.
56. Plankensteiner, K.; Reiner, H.; Rode, B. Prebiotic Chemistry: The Amino Acid and Peptide World. *Curr. Org. Chem.* **2005**, *9*, 1107–1114.
57. Segré, D.; Ben-Eli, D.; Deamer, D.W.; Lancet, D. The lipid world. *Origins Life Evol. B.* **2001**, *31*, 119–145.
58. Tessera, M. Origin of evolution versus origin of life: A shift of paradigm. *Int. J. Mol. Sci.* **2011**, *12*, 3445–3458.

59. Sharov, A.A. Coenzyme world model of the origin of life. *Biosystems*. **2016**, *144*, 8–17.
60. Koonin, E.V.; Senkevich, T.G.; Dolja, V.V. The ancient Virus World and evolution of cells. *Biol. Direct.* **2006**, *1*, 29.
61. Lanier, K.A.; Williams, L.D. The Origin of Life: Models and Data. *J. Mol. Evol.* **2017**, *84*, 85–92.
62. Xavier, J.C.; Patil, K.R.; Rocha, I. Systems biology perspectives on minimal and simpler cells. *Microbiol. Mol. Biol. Rev.* **2014**, *78*, 487–509.
63. Pross, A. Causation and the Origin of Life. Metabolism or Replication First? *Origins Life Evol. B.* **2004**, *34*, 307–321.
64. Wächtershäuser, G. In Praise of Error. *J. Mol. Evol.* **2016**, *82*, 75–80.
65. Sutherland, J.D. Opinion: Studies on the origin of life — the end of the beginning. *Nat. Rev. Chem.* **2017**, *1*.
66. Ross, D.S. It is Neither Frankenstein Nor a Submarine Alkaline Vent, It is Just the Second Law. *BioEssays* **2018**, *40*, 1800149.
67. Branscomb, E.; Russell, M.J. Frankenstein or a Submarine Alkaline Vent: Who is Responsible for Abiogenesis? *BioEssays* **2018**, *40*, 1700182.
68. Emergence, O. Workshop OQOL'09: Open Questions on the Origins of Life 2009. *Origins Life Evol. B.* **2010**, *40*, 347–497.
69. Oparin, A. I. The origin of life (A. Synge, transl) In *The Origin of Life*, Bernal, J. D., Weidenfeld & Nicolson: London, U.K. 1967 pp. 197–234.
70. Haldane, J.B.S. The Origin of Life. *Ration. Annu.* **1929**, 3–10.
71. Pennington, D.D.; Simpson, G.L.; McConnell, M.S.; Fair, J.M.; Baker, R.J. Transdisciplinary Research, Transformative Learning, and Transformative Science. *BioScience* **2013**, *63*, 564–573.
72. Polanco, C. Why interdisciplinary research matters. *Nature* **2015**, *525*, 305.
73. Morowitz, H.; Smith, E. Energy flow and the organization of life. *Complexity* **2007**, *13*, 51–59.
74. Kaufmann, M. On the free energy that drove primordial anabolism. *Int. J. Mol. Sci.* **2009**, *10*, 1853–1871.
75. Zhang, W.; Li, F.; Nie, L. Integrating multiple “omics” analysis for microbial biology: application and methodologies. *Microbiology* **2010**, *156*, 287–301.
76. Larsen, P.; Hamada, Y.; Gilbert, J. Modeling microbial communities: current, developing, and future technologies for predicting microbial community interaction. *J. Biotechnol.* **2012**, *160*, 17–24.
77. Waite, J.H.; Glein, C.R.; Perryman, R.S.; Teolis, B.D.; Magee, B.A.; Miller, G.; Grimes, J.; Perry, M.E.; Miller, K.E.; Bouquet, A.; et al. Cassini finds molecular hydrogen in the Enceladus plume: Evidence for hydrothermal processes. *Science* **2017**, *356*, 155–159.
78. Zahnle, K.; Arndt, N.; Cockell, C.; Halliday, A.; Nisbet, E.; Selsis, F.; Sleep, N.H. Emergence of a Habitable Planet. *Space Sci. Rev.* **2007**, 35–78.
79. Ehrenfreund, P.; Irvine, W.; Becker, L.; Blank, J.; Brucato, J.R.; Colangeli, L.; Derenne, S.; Despois, D.; Dutrey, A.; Fraaije, H.; et al. Astrophysical and astrochemical insights into the origin of life. *Rep. Prog. Phys.* **2002**, *65*, 1427–1487.
80. Vance, S.D. The Habitability of Icy Ocean Worlds in the Solar System. In *Handbook of Exoplanets*; Springer: Berlin, Germany, 2018; pp. 1–23.
81. Postberg, F.; Khawaja, N.; Abel, B.; Choblet, G.; Glein, C.R.; Gudipati, M.S.; Henderson, B.L.; Hsu, H.-W.; Kempf, S.; Klenner, F.; et al. Macromolecular organic compounds from the depths of Enceladus. *Nature* **2018**, *558*, 564–568.
82. Khawaja, N.; Postberg, F.; Hillier, J.; Klenner, F.; Kempf, S.; Nölle, L.; Reviol, R.; Zou, Z.; Srama, R. Low-mass nitrogen-, oxygen-bearing, and aromatic compounds in Enceladean ice grains. *Mon. Not. R. Astron.*

- Soc.* **2019**, *489*, 5231–5243.
83. Hörst, S.M.; Yelle, R.V.; Buch, A.; Carrasco, N.; Cernogora, G.; Dutuit, O.; Quirico, E.; Sciamma-O'Brien, E.; Smith, M.A.; Somogyi, Á.; et al. Formation of Amino Acids and Nucleotide Bases in a Titan Atmosphere Simulation Experiment. *Astrobiology* **2012**, *12*, 809–817.
  84. Sousa, F.L.; Thiergart, T.; Landan, G.; Nelson-Sathi, S.; Pereira, I.A.C.; Allen, J.F.; Lane, N.; Martin, W.F. Early bioenergetic evolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2013**, *368*, 20130088.
  85. Müller, V.; Chowdhury, N.P.; Basen, M. Electron Bifurcation: A Long-Hidden Energy-Coupling Mechanism. *Annu. Rev. Microbiol.* **2018**, *72*, 331–353.
  86. Whicher, A.; Camprubi, E.; Pinna, S.; Herschy, B.; Lane, N. Acetyl Phosphate as a Primordial Energy Currency at the Origin of Life. *Origins. Life Evol. B.* **2018**, *48*, 159–179.
  87. Duve, C.D. Clues from present-day biology: the thioester world. In *The Molecular Origins of Life*; Brack, A., Ed.; Cambridge University Press: Cambridge, UK, 1998; pp. 219–236.
  88. Chandru, K.; Gilbert, A.; Butch, C.; Aono, M.; Cleaves, H.J. The Abiotic Chemistry of Thiolated Acetate Derivatives and the Origin of Life. *Sci. Rep.* **2016**, *6*, 29883.
  89. Schoepp-Cothenet, B.; van Lis, R.; Atteia, A.; Baymann, F.; Capowiez, L.; Ducluzeau, A.-L.; Duval, S.; ten Brink, F.; Russell, M.J.; Nitschke, W. On the universal core of bioenergetics. *Biochim. Biophys. Acta* **2013**, *1827*, 79–93.
  90. Jinich, A.; Sanchez-Lengeling, B.; Ren, H.; Goldford, J.E.; Noor, E.; Sanders, J.N.; Segrè, D.; Aspuru-Guzik, A. A thermodynamic atlas of carbon redox chemical space. *BioRxiv* **2019**, 245811.
  91. Mattia, E.; Otto, S. Supramolecular systems chemistry. *Nat. Nanotechnol.* **2015**, *10*, 111–119.
  92. Kim, H.; Smith, H.B.; Mathis, C.; Raymond, J.; Walker, S.I. Universal scaling across biochemical networks on Earth. *Sci Adv* **2019**, *5*, eaau0149.
  93. Markovitch, O.; Lancet, D. Excess mutual catalysis is required for effective evolvability. *Artif. Life* **2012**, *18*, 243–266.
  94. Pascal, R.; Pross, A. Stability and its manifestation in the chemical and biological worlds. *Chem. Commun.* **2015**, *51*, 16160–16165.
  95. Xavier, J.C.; Hordijk, W.; Kauffman, S.; Steel, M.; Martin, W.F. Autocatalytic chemical networks preceded proteins and RNA in evolution. *bioRxiv* doi: 10.1101/693879 . Available online: <https://www.biorxiv.org/content/10.1101/693879v1.abstract>. (accessed on: on: 25th February 2020)
  96. Goldford, J.E.; Hartman, H.; Smith, T.F.; Segrè, D. Remnants of an Ancient Metabolism without Phosphate. *Cell* **2017**, *168*, 1126–1134.e9.
  97. Goldford, J.E.; Segrè, D. Modern views of ancient metabolic networks. *Curr. Opin. Syst. Biol.* **2018**, *8*, 117–124.
  98. Muchowska, K.B.; Varma, S.J.; Moran, J. Synthesis and breakdown of universal metabolic precursors promoted by iron. *Nature* **2019**, *569*, 104–107.
  99. Keller, M.A.; Turchyn, A.V.; Ralser, M. Non-enzymatic glycolysis and pentose phosphate pathway-like reactions in a plausible Archean ocean. *Mol. Syst. Biol.* **2014**, *10*, 725–725.
  100. Vincent, L.; Berg, M.; Krismer, M.; Saghaei, S.S.; Cosby, J.; Sankari, T.; Vetsigian, K.; Li, H.J.C.; Baum, D.A. Chemical Ecosystem Selection on Mineral Surfaces Reveals Long-Term Dynamics Consistent with the Spontaneous Emergence of Mutual Catalysis. *Life* **2019**, *9*.
  101. Amend, J.P.; LaRowe, D.E.; McCollom, T.M.; Shock, E.L. The energetics of organic synthesis inside and outside the cell. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2013**, *368*, 20120255.
  102. Russell, M.J.; Martin, W. The rocky roots of the acetyl-CoA pathway. *Trends Biochem. Sci.* **2004**, *29*, 358–363.

103. Preiner, M.; Igarashi, K.; Muchowska, K.B.; Yu, M.; Varma, S.J.; Kleinermanns, K.; Nobu, M.K.; Kamagata, Y.; Tüysüz, H.; Moran, J.; Martin, W.F. A hydrogen-dependent geochemical analogue of primordial carbon and energy metabolism. *Nat. Ecol. Evol.* **2020**, in press; doi: 10.1038/s41559-020-1125-6
104. Barge, L.M.; White, L.M. Experimentally Testing Hydrothermal Vent Origin of Life on Enceladus and Other Icy/Ocean Worlds. *Astrobiology* **2017**, *17*, 820–833.
105. Adami, C.; Labar, T. From Entropy to Information: Biased Typewriters and the Origin of Life. *arXiv* preprint arXiv:1506.06988. Available online: <https://arxiv.org/abs/1506.06988> (accessed on: 25th February 2020)
106. Turk, R.M.; Chumachenko, N.V.; Yarus, M. Multiple translational products from a five-nucleotide ribozyme. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 4585–4589.
107. Szathmáry, E. The origin of the genetic code: amino acids as cofactors in an RNA world. *Trends Genet.* **1999**, *15*, 223–229.
108. Morris, C.E. How did cells get their size? *Anat. Rec.* **2002**, *268*, 239–251.
109. Noller, H.F. The driving force for molecular evolution of translation. *RNA* **2004**, *10*, 1833–1837.
110. de Vlar, H.P. Amino acid fermentation at the origin of the genetic code. *Biol. Direct* **2012**, *7*, 6.
111. Tagami, S.; Attwater, J.; Holliger, P. Simple peptides derived from the ribosomal core potentiate RNA polymerase ribozyme function. *Nat. Chem.* **2017**, *9*, 325–332.
112. Maddox, J. The genetic code by numbers. *Nature* **1994**, *367*, 111–111.
113. Yarus, M. The Genetic Code and RNA-Amino Acid Affinities. *Life* **2017**, *7*, 13.
114. Wong, J.T. A co-evolution theory of the genetic code. *Proc. Natl. Acad. Sci. U. S. A.* **1975**, *72*, 1909–1912.
115. Knight, R.D.; Freeland, S.J.; Landweber, L.F. Selection, history and chemistry: the three faces of the genetic code. *Trends Biochem. Sci.* **1999**, *24*, 241–247.
116. Kun, Á.; Radványi, Á. The evolution of the genetic code: Impasses and challenges. *Biosystems.* **2018**, *164*, 217–225.
117. Higgs, P.G.; Pudritz, R.E. A thermodynamic basis for prebiotic amino acid synthesis and the nature of the first genetic code. *Astrobiology* **2009**, *9*, 483–490.
118. Müller, M.M.; Allison, J.R.; Hongdilokkul, N.; Gaillon, L.; Kast, P.; van Gunsteren, W.F.; Marlière, P.; Hilvert, D. Directed evolution of a model primordial enzyme provides insights into the development of the genetic code. *PLoS Genet.* **2013**, *9*, e1003187.
119. Ilardo, M.; Bose, R.; Meringer, M.; Rasulev, B.; Grefenstette, N.; Stephenson, J.; Freeland, S.; Gillams, R.J.; Butch, C.J.; Cleaves, H.J., 2nd Adaptive Properties of the Genetically Encoded Amino Acid Alphabet Are Inherited from Its Subsets. *Sci. Rep.* **2019**, *9*, 12468.
120. Cartwright, J.H.E.; Giannerini, S.; González, D.L. DNA as information: at the crossroads between biology, mathematics, physics and chemistry. *Philos. Trans. A Math. Phys. Eng. Sci.* **2016**, *374*.
121. Woese, C.R.; Dugre, D.H.; Saxinger, W.C.; Dugre, S.A. The molecular basis for the genetic code. *Proc. Natl. Acad. Sci. U. S. A.* **1966**, *55*, 966–974.
122. Copley, S.D.; Smith, E.; Morowitz, H.J. A mechanism for the association of amino acids with their codons and the origin of the genetic code. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 4442–4447.
123. Koonin, E.V.; Novozhilov, A.S. Origin and Evolution of the Universal Genetic Code. *Annu. Rev. Genet.* **2017**, *51*, 45–62.
124. Scharf, C.; Virgo, N.; Cleaves, H.J., 2nd; Aono, M.; Aubert-Kato, N.; Aydinoglu, A.; Barahona, A.; Barge, L.M.; Benner, S.A.; Biehl, M.; et al. A Strategy for Origins of Life Research. *Astrobiology* **2015**, *15*, 1031–1042.
125. Gutekunst, K. Hypothesis on the Synchronistic Evolution of Autotrophy and Heterotrophy. *Trends Biochem. Sci.* **2018**, *43*, 402–411.

126. Mansy, S.S.; Szostak, J.W. Reconstructing the emergence of cellular life through the synthesis of model protocells. *Cold Spring Harb. Symp. Quant. Biol.* **2009**, *74*, 47–54.
127. Cornell, C.E.; Black, R.A.; Xue, M.; Litz, H.E.; Ramsay, A.; Gordon, M.; Mileant, A.; Cohen, Z.R.; Williams, J.A.; Lee, K.K.; et al. Prebiotic amino acids bind to and stabilize prebiotic fatty acid membranes. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 17239–17244.
128. Jordan, S.F.; Ramm, H.; Zheludev, I.N.; Hartley, A.M.; Maréchal, A.; Lane, N. Promotion of protocell self-assembly from mixed amphiphiles at the origin of life. *Nat. Ecol. Evol.* **2019**, *3*, 1705–1714.
129. Monnard, P.-A.; Walde, P. Current Ideas about Prebiological Compartmentalization. *Life* **2015**, *5*, 1239–1263.
130. Dzieciol, A.J.; Mann, S. Designs for life: protocell models in the laboratory. *Chem. Soc. Rev.* **2012**, *41*, 79–85.
131. Gardner, P.M.; Winzer, K.; Davis, B.G. Sugar synthesis in a protocellular model leads to a cell signalling response in bacteria. *Nat. Chem.* **2009**, *1*, 377–383.
132. Kurihara, K.; Tamura, M.; Shohda, K.-I.; Toyota, T.; Suzuki, K.; Sugawara, T. Self-reproduction of supramolecular giant vesicles combined with the amplification of encapsulated DNA. *Nat. Chem.* **2011**, *3*, 775–781.
133. Gumpenberger, T.; Vorkapic, D.; Zingl, F.G.; Pressler, K.; Lackner, S.; Seper, A.; Reidl, J.; Schild, S. Nucleoside uptake in *Vibrio cholerae* and its role in the transition fitness from host to environment. *Mol. Microbiol.* **2016**, *99*, 470–483.
134. White, H.B. 3rd Coenzymes as fossils of an earlier metabolic state. *J. Mol. Evol.* **1976**, *7*, 101–104.
135. Bhowmik, S.; Krishnamurthy, R. The role of sugar-backbone heterogeneity and chimeras in the simultaneous emergence of RNA and DNA. *Nat. Chem.* **2019**, *11*, 1009–1018.
136. Trifonov, E.N. Consensus temporal order of amino acids and evolution of the triplet code. *Gene* **2000**, *261*, 139–151.
137. Parker, E.T.; Cleaves, H.J.; Bada, J.L. Quantitation of  $\alpha$ -hydroxy acids in complex prebiotic mixtures via liquid chromatography/tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **2016**, *30*, 2043–2051.
138. Sephton, M.A. Organic compounds in carbonaceous meteorites. *Nat. Prod. Rep.* **2002**, *19*, 292–311.
139. Johnson, A.P.; Cleaves, H.J.; Dworkin, J.P.; Glavin, D.P.; Lazcano, A.; Bada, J.L. The Miller volcanic spark discharge experiment. *Science* **2008**, *322*, 404.
140. Chandru, K.; Guttenberg, N.; Giri, C.; Hongo, Y.; Butch, C.; Mamajanov, I.; Cleaves, H.J. Simple prebiotic synthesis of high diversity dynamic combinatorial polyester libraries. *Commun. Chem.* **2018**, *1*, 30.
141. Jia, T.Z.; Chandru, K.; Hongo, Y.; Afrin, R.; Usui, T.; Myojo, K.; Cleaves, H.J. Membraneless polyester microdroplets as primordial compartments at the origins of life. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 15830–15835.
142. Chandru, K.; Mamajanov, I.; Cleaves, H.J., 2nd; Jia, T.Z. Polyesters as a Model System for Building Primitive Biologies from Non-Biological Prebiotic Chemistry. *Life* **2020**, *10*.
143. Guttenberg, N.; Virgo, N.; Chandru, K.; Scharf, C.; Mamajanov, I. Bulk measurements of messy chemistries are needed for a theory of the origins of life. *Philos. Trans. Royal Soc. A.* **2017**, 375.
144. Schmitt-Kopplin, P.; Gabelica, Z.; Gougeon, R.D.; Fekete, A.; Kanawati, B.; Harir, M.; Gebefuegi, I.; Eckel, G.; Hertkorn, N. High molecular diversity of extraterrestrial organic matter in Murchison meteorite revealed 40 years after its fall. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 2763–2768.
145. Arndt, N.T.; Nisbet, E.G. Processes on the Young Earth and the Habitats of Early Life. *Annu. Rev. Earth Planet. Sci.* **2012**, *40*, 521–549.
146. Nisbet, E.G.; Sleep, N.H. The habitat and nature of early life. *Nature* **2001**, *409*, 1083–1091.

147. Lang, S.Q.; Butterfield, D.A.; Schulte, M.; Kelley, D.S.; Lilley, M.D. Elevated concentrations of formate, acetate and dissolved organic carbon found at the Lost City hydrothermal field. *Geochim. Cosmochim. Acta* **2010**, *74*, 941–952.
148. Herschy, B.; Whicher, A.; Camprubi, E.; Watson, C.; Dartnell, L.; Ward, J.; Evans, J.R.G.; Lane, N. An origin-of-life reactor to simulate alkaline hydrothermal vents. *J. Mol. Evol.* **2014**, *79*, 213–227.
149. Weiss, M.C.; Sousa, F.L.; Mrnjavac, N.; Neukirchen, S.; Roettger, M.; Nelson-Sathi, S.; Martin, W.F. The physiology and habitat of the last universal common ancestor. *Nat Microbiol* **2016**, *1*, 16116.
150. Mulikidjanian, A.Y.; Bychkov, A.Y.; Dibrova, D.V.; Galperin, M.Y.; Koonin, E.V. Origin of first cells at terrestrial, anoxic geothermal fields. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, E821–30.
151. Milshteyn, D.; Damer, B.; Havig, J.; Deamer, D. Amphiphilic Compounds Assemble into Membranous Vesicles in Hydrothermal Hot Spring Water but Not in Seawater. *Life* **2018**, *8*, 11.
152. Becker, S.; Schneider, C.; Okamura, H.; Crisp, A.; Amatov, T.; Dejmek, M.; Carell, T. Wet-dry cycles enable the parallel origin of canonical and non-canonical nucleosides by continuous synthesis. *Nat. Commun.* **2018**, *9*, 163.
153. Varma, S.J.; Muchowska, K.B.; Chatelain, P.; Moran, J. Native iron reduces CO to intermediates and end-products of the acetyl-CoA pathway. *Nat Ecol Evol* **2018**, *2*, 1019–1024.
154. Garcia, A.K.; McShea, H.; Kolaczowski, B.; Kaçar, B. Reconstructed ancient nitrogenases suggest Mo-specific ancestry. *Evol. Biol.* **2019**, 520.
155. Bray, M.S.; Lenz, T.K.; Haynes, J.W.; Bowman, J.C.; Petrov, A.S.; Reddi, A.R.; Hud, N.V.; Williams, L.D.; Glass, J.B. Multiple prebiotic metals mediate translation. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 12164–12169.
156. Bonfio, C.; Godino, E.; Corsini, M.; de Biani, F.F.; Guella, G.; Mansy, S.S. Prebiotic iron–sulfur peptide catalysts generate a pH gradient across model membranes of late protocells. *Nat. Catal.* **2018**, *1*, 616–623.
157. Kreysing, M.; Keil, L.; Lanzmich, S.; Braun, D. Heat flux across an open pore enables the continuous replication and selection of oligonucleotides towards increasing length. *Nat. Chem.* **2015**, *7*, 203–208.
158. Hazen, R.M.; Sverjensky, D.A. Mineral surfaces, geochemical complexities, and the origins of life. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, a002162.
159. Colón-Santos, S.; Cooper, G.J.T.; Cronin, L. Taming the Combinatorial Explosion of the Formose Reaction via Recursion within Mineral Environments. *ChemSystemsChem* **2019**, *54*, 104.
160. Sadownik, J.W.; Mattia, E.; Nowak, P.; Otto, S. Diversification of self-replicating molecules. *Nat. Chem.* **2016**, *8*, 264–269.
161. Semenov, S.N.; Kraft, L.J.; Ainla, A.; Zhao, M.; Baghbanzadeh, M.; Campbell, V.E.; Kang, K.; Fox, J.M.; Whitesides, G.M. Autocatalytic, bistable, oscillatory networks of biologically relevant organic reactions. *Nature* **2016**, *537*, 656–660.
162. Stelling, J. Mathematical models in microbial systems biology. *Curr. Opin. Microbiol.* **2004**, *7*, 513–518.
163. Goldman, A.D.; Bernhard, T.M.; Dolzhenko, E.; Landweber, L.F. LUCApedia: a database for the study of ancient life. *Nucleic Acids Res.* **2013**, *41*, D1079–82.
164. Nghe, P.; Hordijk, W.; Kauffman, S.A.; Walker, S.I.; Schmidt, F.J.; Kemble, H.; Yeates, J.A.M.; Lehman, N. Prebiotic network evolution: six key parameters. *Mol. Biosyst.* **2015**, *11*, 3206–3217.
165. Woese, C. The universal ancestor. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 6854–6859.
166. Tocheva, E.I.; Ortega, D.R.; Jensen, G.J. Sporulation, bacterial cell envelopes and the origin of life. *Nat. Rev. Microbiol.* **2016**, *14*, 535–542.
167. Hastings, J.; de Matos, P.; Dekker, A.; Ennis, M.; Harsha, B.; Kale, N.; Muthukrishnan, V.; Owen, G.; Turner, S.; Williams, M.; et al. The ChEBI reference database and ontology for biologically relevant chemistry:

- enhancements for 2013. *Nucleic Acids Res.* **2013**, *41*, D456–63.
168. Kanehisa, M.; Furumichi, M.; Tanabe, M.; Sato, Y.; Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **2017**, *45*, D353–D361.
  169. King, Z.A.; Lu, J.; Dräger, A.; Miller, P.; Federowicz, S.; Lerman, J.A.; Ebrahim, A.; Palsson, B.O.; Lewis, N.E. BiGG Models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Res.* **2016**, *44*, D515–22.
  170. Kim, K.M.; Caetano-Anollés, G. The proteomic complexity and rise of the primordial ancestor of diversified life. *BMC Evol. Biol.* **2011**, *11*, 140.
  171. Delaye, L.; Becerra, A. Cenancestor, the Last Universal Common Ancestor. *Evo Edu Outreach* **2012**, *5*, 382–388.
  172. Theobald, D.L. A formal test of the theory of universal common ancestry. *Nature* **2010**, *465*, 219–222.
  173. Surman, A.J.; Rodriguez-Garcia, M.; Abul-Haija, Y.M.; Cooper, G.J.T.; Gromski, P.S.; Turk-MacLeod, R.; Mullin, M.; Mathis, C.; Walker, S.I.; Cronin, L. Environmental control programs the emergence of distinct functional ensembles from unconstrained chemical reactions. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 5387–5392.
  174. Sleep, N.H. Geological and Geochemical Constraints on the Origin and Evolution of Life. *Astrobiology* **2018**, *18*, 1199–1219.
  175. Martin, W.; Russell, M.J. On the origin of biochemistry at an alkaline hydrothermal vent. *Philos. Trans. R. Soc. B.* **2007**, *362*, 1887–1926.
  176. Martin, W.; Baross, J.; Kelley, D.; Russell, M.J. Hydrothermal vents and the origin of life. *Nat. Rev. Microbiol.* **2008**, *6*, 805–814.
  177. Deamer, D. The Role of Lipid Membranes in Life's Origin. *Life* **2017**, *7*.
  178. Westall, F.; Hickman-Lewis, K.; Hinman, N.; Gautret, P.; Campbell, K.A.; Bréhéret, J.G.; Foucher, F.; Hubert, A.; Sorieul, S.; Dass, A.V.; et al. A Hydrothermal-Sedimentary Context for the Origin of Life. *Astrobiology* **2018**, *18*, 259–293.
  179. Pearce, B.K.D.; Pudritz, R.E.; Semenov, D.A.; Henning, T.K. Origin of the RNA world: The fate of nucleobases in warm little ponds. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 11327–11332.
  180. Ross, D.S.; Deamer, D. Dry/Wet Cycling and the Thermodynamics and Kinetics of Prebiotic Polymer Synthesis. *Life* **2016**, *6*.
  181. Morasch, M.; Liu, J.; Dirscherl, C.F.; Ianeselli, A.; Kühnlein, A.; Le Vay, K.; Schwintek, P.; Islam, S.; Corpinot, M.K.; Scheu, B.; et al. Heated gas bubbles enrich, crystallize, dry, phosphorylate and encapsulate prebiotic molecules. *Nat. Chem.* **2019**, *11*, 779–788.
  182. Becker, S.; Feldmann, J.; Wiedemann, S.; Okamura, H.; Schneider, C.; Iwan, K.; Crisp, A.; Rossa, M.; Amatov, T.; Carell, T. Unified prebiotically plausible synthesis of pyrimidine and purine RNA ribonucleotides. *Science* **2019**, *366*, 76–82.
  183. Ooka, H.; McGlynn, S.E.; Nakamura, R. Electrochemistry at Deep-Sea Hydrothermal Vents: Utilization of the Thermodynamic Driving Force towards the Autotrophic Origin of Life. *ChemElectroChem* **2019**, *6*, 1316–1323.
  184. Mansy, S.S.; Schrum, J.P.; Krishnamurthy, M.; Tobé, S.; Treco, D.A.; Szostak, J.W. Template-directed synthesis of a genetic polymer in a model protocell. *Nature* **2008**, *454*, 122–125.
  185. Lazcano, A. Alexandr I. Oparin and the Origin of Life: A Historical Reassessment of the Heterotrophic Theory. *J. Mol. Evol.* **2016**, *83*, 214–222.
  186. Wood, A.P.; Aurikko, J.P.; Kelly, D.P. A challenge for 21st century molecular biology and biochemistry:

- what are the causes of obligate autotrophy and methanotrophy? *FEMS Microbiol. Rev.* **2004**, *28*, 335–352.
187. Schönheit, P.; Buckel, W.; Martin, W.F. On the Origin of Heterotrophy. *Trends Microbiol.* **2016**, *24*, 12–25.
188. Ashkenasy, G.; Hermans, T.M.; Otto, S.; Taylor, A.F. Systems chemistry. *Chem. Soc. Rev.* **2017**, *46*, 2543–2554.
189. Pascal, R.; Pross, A.; Sutherland, J.D. Towards an evolutionary theory of the origin of life based on kinetics and thermodynamics. *Open Biol.* **2013**, *3*, 130156.
190. Semenov, S.N.; Wong, A.S.Y.; van der Made, R.M.; Postma, S.G.J.; Groen, J.; van Roekel, H.W.H.; de Greef, T.F.A.; Huck, W.T.S. Rational design of functional and tunable oscillating enzymatic networks. *Nat. Chem.* **2015**, *7*, 160–165.
191. Thauer, R.K.; Kaster, A.-K.; Seedorf, H.; Buckel, W.; Hedderich, R. Methanogenic archaea: ecologically relevant differences in energy conservation. *Nat. Rev. Microbiol.* **2008**, *6*, 579–591.
192. Duim, H.; Otto, S. Towards open-ended evolution in self-replicating molecular systems. *Beilstein J. Org. Chem.* **2017**, *13*, 1189–1203.
193. Deamer, D.; Weber, A.L. Bioenergetics and Life's Origins. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, a004929–a004929.
194. Boiteau, L.; Pascal, R. Energy sources, self-organization, and the origin of life. *Origins. Life Evol. B.* **2011**, *41*, 23–33.
195. Krishnamurthy, R. Life's Biological Chemistry: A Destiny or Destination Starting from Prebiotic Chemistry? *Chemistry* **2018**, *24*, 16708–16715.
196. Szathmáry, E.; Maynard Smith, J. From replicators to reproducers: the first major transitions leading to life. *J. Theor. Biol.* **1997**, *187*, 555–571.
197. Eigen, M. Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **1971**, *58*, 465–523.
198. Walker, S.I.; Davies, P.C.W. The algorithmic origins of life. *J. R. Soc. Interface* **2013**, *10*, 20120869.
199. Segré, D.; Lancet, D. Composing life. *EMBO Rep.* **2000**, *1*, 217–222.
200. Hordijk, W.; Steel, M.; Kauffman, S.A. Molecular Diversity Required for the Formation of Autocatalytic Sets. *Life* **2019**, *9*.
201. Kostyrka, G. What roles for viruses in origin of life scenarios? *Stud. Hist. Philos. Biol. Biomed. Sci.* **2016**, *59*, 135–144.
202. Stewart, J.E. The Origins of Life: The Managed-Metabolism Hypothesis. *Found. Sci.* **2019**, *24*, 171–195.
203. Vasas, V.; Szathmáry, E.; Santos, M. Lack of evolvability in self-sustaining autocatalytic networks constraints metabolism-first scenarios for the origin of life. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 1470–1475.
204. Sharma, V.; Annala, A. Natural process–Natural selection. *Biophys. Chem.* **2007**, *127*, 123–128.
205. Meléndez-Hevia, E.; Montero-Gómez, N.; Montero, F. From prebiotic chemistry to cellular metabolism—The chemical evolution of metabolism before Darwinian natural selection. *J. Theor. Biol.* **2008**, *252*, 505–519.
206. Frenkel-Pinter, M.; Haynes, J.W.; C, M.; Petrov, A.S.; Burcar, B.T.; Krishnamurthy, R.; Hud, N.V.; Leman, L.J.; Williams, L.D. Selective incorporation of proteinaceous over nonproteinaceous cationic amino acids in model prebiotic oligomerization reactions. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 16338–16346.
207. Hud, N.V.; Cafferty, B.J.; Krishnamurthy, R.; Williams, L.D. The Origin of RNA and “My Grandfather's Axe.” *Chem. Biol.* **2013**, *20*, 466–474.
208. Joyce, G.F. Molecular evolution: booting up life. *Nature* **2002**, *420*, 278–279.
209. Calvin, M. Mineral Origins of Life Genetic Takeover and the Mineral Origins of Life A. G. Cairns-Smith. *BioScience* **1983**, *33*, 596–596.



210. Schmitt-Kopplin, P.; Hemmler, D.; Moritz, F.; Gougeon, R.D.; Lucio, M.; Meringer, M.; Müller, C.; Harir, M.; Hertkorn, N. Systems chemical analytics: introduction to the challenges of chemical complexity analysis. *Faraday Discuss.* **2019**, *218*, 9–28.
211. Geisberger, T.; Diederich, P.; Steiner, T.; Eisenreich, W.; Schmitt-Kopplin, P.; Huber, C. Evolutionary Steps in the Analytics of Primordial Metabolic Evolution. *Life* **2019**, *9*, 50.
212. Richert, C. Prebiotic chemistry and human intervention. *Nat. Commun.* **2018**, *9*, 5177.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).