Origins of Life: Chemistry and Evolution

Kavita Matange[1,2,3], Vahab Rajaei[1,2,3], Gary Schuster[2,3], Nicholas V. Hud[2,3], Cesar Menor-Salvan[2,4], Pau Capera-Aragonès[5], Moran Frenkel-Pinter[1,2,5,*], and Loren Dean Williams[1,2,3,*]

[1]NASA Center for the Origins of Life, [2]NSF-NASA Center of Chemical Evolution, [3]School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, GA 30332-0400 USA, [4]University of Alcalá, Alcalá de Henares (Madrid) Spain, [5]Institute of Chemistry, The Hebrew University of Jerusalem, Israel 91904

* address correspondence to Moran Frenkel-Pinter (moran.fp@mail.huji.ac.il) or Loren Dean Williams (ldw@gatech.edu)

Keywords: chemical evolution, Darwinian evolution, biopolymers, prebiotic chemistry, chemical origins of life

https://doi.org/10.26434/chemrxiv-2023-1jrcq-v2 **ORCID:** https://orcid.org/0000-0002-7215-4194 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

Abstract

Our understanding of the origins of life will be enhanced if models and their predictions are clearly understood and explicitly articulated. Here we outline two distinct models that are currently used to explain the origins of life. In one model, which has been pursued for a half century, inherent chemical reactivities of prebiotic chemical species produced RNA, which then invented evolution. This direct synthesis model enables the prediction that if the conditions of the ancient earth are sufficiently constrained, chemists will discover the synthetic pathways that produced RNA. In a fundamentally different model, which is more recent and less mature, RNA in concert with other biopolymers arose from prolonged, selection-based changes that occurred during chemical evolution, which transitioned smoothly into biological evolution. This evolutionary model predicts common chemistry of linkage and amazing structures, assemblies and co-assemblies, as represented by double stranded DNA, tRNA, cellulose, collagen, globular proteins, ATP synthase, and the ribosome. This evolutionary model predicts profound integration of biological subsystems as represented by ATP, which is central to and inextricable from biopolymer structure and biosynthesis and metabolic systems. In the evolutionary model, inherent chemical reactivities of biological building blocks are not necessarily relevant to the origins of life and do not predict biosynthesis. The two models of the origins of life are fundamentally different from one another and guide design of very different experimental approaches to test their underlying assumptions. It is currently undetermined which model, or a hybrid of them, is closer to reality.

https://doi.org/10.26434/chemrxiv-2023-1jrcq-v2 **ORCID:** https://orcid.org/0000-0002-7215-4194 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

Introduction

Seeking an account of our origins is a quintessential marker of human curiosity. Today, that quest is focused on the scientific search for the origin of life itself. We now know that the backbones and sidechains of life's essential biopolymers, RNA, DNA, and polypeptide were fixed in chemical structure by the time of the last universal common ancestor, around four billion years ago (1-4), and that they have been held invariant over all biological time and all speciation. Thus, the origins of biopolymers embody the central question about the origin of life. Where did biopolymers and their building blocks come from? What processes converted mixtures of prebiotic small molecules on the Hadean earth (5-7) into sophisticated informational, functional, and structural biopolymers built in living cells by condensing homochiral building blocks into specific sequences with specific linkages?

Two main models have been offered to explain the origins of biopolymers. One model, direct chemical synthesis, has been dominant over the last half century. In this model, biology incorporated and has maintained building blocks that arose as a consequence of direct synthetic success on the abiotic Hadean Earth. In this model, extant building blocks provide information on prebiotic chemistry. The other model, origins by chemical evolution, is more recent. In this model, chemical species that arose as a consequence of synthetic success on the Hadean Earth were serially replaced during chemical and early biological evolution. In this model, extant building blocks do not report in a direct way on prebiotic chemistry or early chemical evolution. We describe both models, articulating their assumptions, mechanisms, predictions, strengths, and weaknesses. We present a working definition of chemical evolution and a general framework for how it can operate. For simplicity, we focus on the origins of RNA, but the discussion incorporates other biopolymers as well.

Model 1. Origins by Direct Chemical Synthesis.

A model has been advanced in which RNA first arose on the ancient earth by direct chemical synthesis (8-14). This model envisions *stepwise reactions directed by inherent chemical reactivities.* Direct chemical synthesis starts from small molecule feedstocks that react in serial and parallel synthetic reactions to produce RNA. In this model, the origins of life is understood through the lens of organic synthetic chemistry.

The direct chemical synthesis model is constrained by environmental scenarios of the Hadean Earth (6-8, 11, 15, 16). Specific chemical properties of small molecule feedstocks such as water, hydrogen

https://doi.org/10.26434/chemrxiv-2023-1jrcq-v2 **ORCID:** https://orcid.org/0000-0002-7215-4194 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

cyanide, cyanamide, formaldehyde, and/or glycoaldehyde caused reactions that produced nucleobases and ribose **(Figure 1)**. Then, specific properties of nucleobases and ribose caused reactions that produced nucleosides. Then, specific properties of nucleosides and phosphate caused reactions that produced nucleotides. Then, specific properties of nucleotides caused reactions that produced RNA. Then, in a profound discontinuity, RNA initiated Darwinian evolution (11).

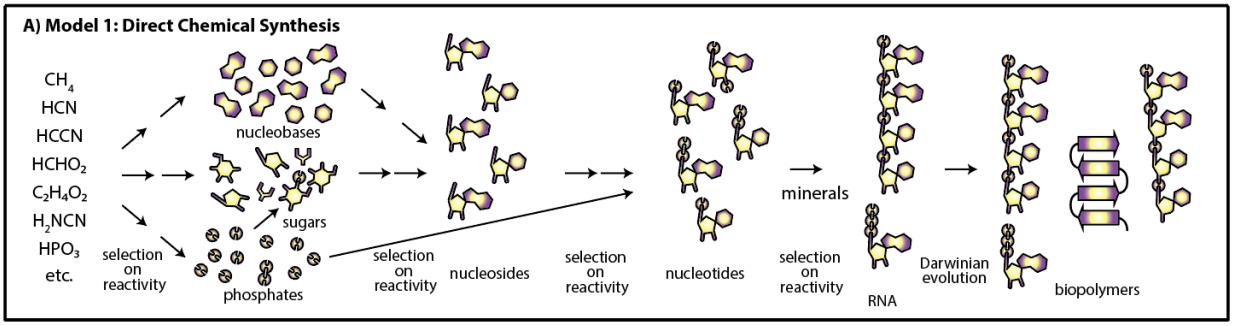

Figure 1. Direct chemical synthesis of RNA, whereby small molecule feedstocks (LH side) enter synthetic pathways that lead to RNA building blocks, then to RNA. This panel was adapted from Benner and coworkers (11). Variations on ordering of the steps and on the specific chemical reactions and intermediates have been proposed.

Variants of the direct synthesis model used altered intermediates and change the ordering of reaction steps. Benner and coworkers refer to variants of the direct chemical synthesis model as 'path hypotheses' (11). Some variants assume formation of phosphorylated carbohydrates before linkage to nucleobases (17). Some hypotheses link fragments of nucleobases and sugars before formation of nucleosides (13, 18) because chemical coupling of ribose with cytosine or uracil has proven problematic. Some variants involve facilitators such as minerals or borate anions (11, 19), while others accede the possibility of proto-RNA, a chemical and functional homolog of RNA that was ancestral to RNA.

If the direct synthesis model is broadly correct, then important goals of origins of life research are to understand prebiotic conditions, chemical inventories and inherent reactivities, and to recreate the synthetic pathways that led to biopolymers. If one can know the conditions of the ancient Earth, inherent reactivities, and the probabilities of certain stochastic events, then one can understand and hopefully recapitulate the origins of biopolymers. The RNA branch of this effort was initially led by Orgel and Oro, and has been extended by Benner, Sutherland and others [recently reviewed by Krishnamurthy (20)].

RNA is pre-evolution: The direct synthesis model is discontinuous – an initial era of non-evolutionary synthetic chemistry is distinct from a subsequent era of Darwinian evolution. RNA origins are pre-evolution (13, 21). RNA changed the world because it enabled evolution. As stated by Orgel, "natural

https://doi.org/10.26434/chemrxiv-2023-1jrcq-v2 **ORCID:** https://orcid.org/0000-0002-7215-4194 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

selection through replication and mutation was the only mechanism for evolving complex biochemical systems from simpler ones" (12).

Predictions of Origins by Direct Chemical Synthesis.

The direct synthesis model of RNA is inspired by work of Miller (22) demonstrating that organic molecules, including racimates of some biopolymer building blocks, are produced under hypothesized pre-biotic scenarios. Miller demonstrated that small molecule feedstocks such as water, methane, ammonia, and hydrogen combine directly to form biological amino acids such as glycine, alanine, and aspartic acid.

The importance of the Urey-Miller experiment is supported by phenomena in space; biological amino acids and many other organic species are found in chondrite meteorites (5, 23). The synthetic pathway does not end with monomeric amino acids. Amino acids readily link to form peptide bonds by a variety of mechanisms (24-26). Chemical pathways that lead to structures resembling racemic polypeptide have been experimentally validated. Corroboration of the direct synthesis model would be the discovery of direct synthetic pathways to nucleobases, sugars, nucleosides and nucleotides and polymers. Ribose and nucleobases but not nucleosides or nucleotides are found in meteorites (27, 28).

Weaknesses of Origins by Direct Chemical Synthesis.

Miller-Urey getting blurry. Some weaknesses of the direct synthesis model have been discussed by Krishnamurthy (20), Shapiro (29) and others. Most fundamentally in our view, is that the Miller-Urey series of experiments have been misinterpreted. The results of Miller-Urey are understood by many to suggest that not only organic molecules in general (30, 31), but most or all the basic building blocks of life (8-14, 32-34) emerged through direct synthesis. This interpretation has shaped research into the origins of life by a broad scientific community.

Researchers have postulated direct synthetic routes from small molecule feedstocks to complex biologically relevant species including proteogenic amino acids tyrosine, phenylalanine, lysine, histidine, asparagine, glutamine, cysteine, methionine and arginine (9, 34), metabolites such as nicotinamide (32), glutathione and iron sulfur clusters (33), and nucleotides (8-14). A nucleotide, which is a three-component oligomer of a nucleobase, a ribose, and pyrophosphate, is profoundly more complex than any amino acid, with more atoms, more functional groups, more hydrolysis products, more elementary components, more chiral centers, and less stable bonds. The high energy phosphate bonds of a nucleotide have no analog in an amino acid.

https://doi.org/10.26434/chemrxiv-2023-1jrcq-v2 **ORCID:** https://orcid.org/0000-0002-7215-4194 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

The simplicity of the Miller-Urey paradigm, one pot conversion of small molecule feedstocks to organic molecules, has metamorphosed to extraordinary complexity. Proposed pathways to nucleotides are specific to a given type of nucleotide, and are composed of compounded branches, each with multiple synthetic steps under varying conditions, in precise order, interleaved by purifications. In practice, one designs numerous syntheses aimed at achieving the molecular target, and explores different reactants, temperatures, solvents, stoichiometries, and order of addition, and purifies intermediates, then explores again until the target is reached. The phrase 'prebiotically plausible' is retroactively fit (35, 36) to reagents and conditions that combine to give the target. The pathways require coincidences of multiple events that independently have low probabilities of occurrence, such as meteor impacts, volcanic eruptions, large scale movements of materials, and transient interactions of organic compounds with salts and minerals (10, 16, 37). Many of the branches can be characterized as "workarounds" to obtain desired products from available compounds.

Proposed synthetic pathways must be balanced against contingency and likelihood. It might be that chemists will eventually discover synthetic pathways from feedstocks, to nucleotides, to RNA. We wonder if these complex laboratory synthetic pathways, so remote from the Miller-Urey paradigm, are relevant to the origins of life on the ancient earth.

Rewired. Direct synthesis models require that the synthetic routes to building blocks and biopolymers developed during a prebiotic phase and were later re-written. Proposed reaction sequences in synthetic pathways are distinctly different from the biosynthetic reaction sequences observed in cells that produce building blocks and biopolymers (20, 40, 41). The orthogonality of direct synthetic pathways and biosynthesis, combined with the lack of direct synthetic pathways for most biochemical species, presents a challenge to the importance of direct synthesis of biological molecules during the origins of life. Is it reasonable that synthetic pathways were rewritten but the final products held constant?

No-go co-evolution. The direct chemical synthesis model is besieged by "chicken and egg" dilemmas; one must place multiple improbable events in chronological order because their simultaneous occurrence appears impossible. What came first, RNA or protein? Information or metabolism? In most [but not all (9)] direct chemical synthesis models, different components and systems have distinct and unrelated origins. For example, phosphorylated sugars can arise by one pathway, nucleobases by another, and amino acids by still another (11). Metabolism is not generally part of these models at all.

In biology, everything is linked to everything, nothing is independent. Biopolymers and metabolism are deeply integrated symbiotic systems that live and die together (44). For example, ATP is required to drive protein synthesis, which is catalyzed by RNA. ATP is a building block of RNA which is

https://doi.org/10.26434/chemrxiv-2023-1jrcq-v2 **ORCID:** https://orcid.org/0000-0002-7215-4194 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

synthesized consumption of amino acids, in reactions catalyzed by proteins. RNA, DNA, protein, and biological metabolism are all impossible without ATP. Translation is impossible with transcription. Replication is impossible without translation.

Chicken and egg dilemmas dissolve in models in which origins of systems and molecules are linked and are integrated from the ground up. Scenarios that are highly improbable if events are independent are likely or even unavoidable if events are linked. In linked scenarios it is reasonable that a building block of RNA would be required for protein synthesis, and also be the basis of metabolism. Therefore, predictions of direct chemical synthesis models, with distinct origins of various systems and molecules, appear to differ from the deep integration and dependencies observed of biological systems.

Foresight. Foresight is not a property of chemical or biological processes. Direct synthesis models appear to require foresight – gratification in the form of natural selection is delayed until completion of long, branching and undulating chemical pathways, with no selection until the final phase. Monomers are incapable of maintaining or transmitting polymeric information or performing complex catalysis. The properties of biopolymers are emergent upon homochirality and polymerization. Neither fragments of biological nucleosides, nor monomeric nucleosides nor nucleotides, assemble as base pairs in aqueous solution.

By contrast, non-canonical nitrogen heterocycles, such as melamine or barbituric acid, form glycosidic linkages with ribose and combine to form linear supramolecular assemblies containing thousands of monomeric paired nucleosides (42). In sum, direct chemical synthesis of RNA requires the establishment of pathways in the absence of a stepwise-driving force. Success occurs only at the conclusion of a long and complex series of disconnected synthetic steps. The ancient earth, unlike modern organic chemists, did not experience an imperative to discover chemical pathways to RNA.

The fortuitous creation of DNA after RNA, a feature of most direct synthesis models, implies additional foresight. RNA has many useful properties but is chemically labile; the 2’ hydroxyl group of ribose is a nucleophile that catalyzes self-cleavage. Because of its lability, RNA genomes (beyond viruses) are problematic. Nature’s solution is DNA, in which the 2’-hydroxyl of RNA has been replaced by hydrogen atom. DNA is persistent chemically, it has been isolated from mammoths that died over 1 million years ago (43). RNA before DNA implies that Nature, before the invention of genomes, produced a polymer (RNA) with the potential to radically change chemical properties (increased persistence, and decreased structural complexity) via a subtle modification, while maintaining base-pairing and formation of double helices. Again, gratification (increased persistence of DNA) awaits the conclusion and is not selected for in intermediate stages.

https://doi.org/10.26434/chemrxiv-2023-1jrcq-v2 **ORCID:** https://orcid.org/0000-0002-7215-4194 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

Continuity. The direct synthesis model assumes a discontinuity between abiotic chemistry and biology. In this model, non-evolutionary direct chemistry produced RNA, which initiated evolution and biology. The chemical and biological eras are distinct and the space between them is discontinuous. The abruptness of the transition, with essentially no intermediate stages, appears to violate the principle of continuity (45, 46), which stipulates modest, consecutive, contingent, ad hoc, and opportune steps. The continuity principle stipulates many intermediary steps that share characteristics of non-evolutionary chemistry and biological evolution - a continuum between chemistry and biology.

Model 2. Origins by Chemical Evolution.

If the molecules of life did not emerge from direct chemical synthesis on the prebiotic earth, then where did they come from? We suggest a process of gradual chemical evolution. But what is evolution before biological molecules? In the chemical evolutionary model proposed here, the transformation of chemistry to biology is progressive, incremental, and continuous **(Figure 2)**. In this model the line between chemistry and biology is blurred and indistinct; prebiotic chemistry is continuous with biology. Darwinian evolution is a special case of chemical evolution and is not a discontinuous phase.

The power of evolution to create and sculpt molecules is documented by invention of tyrosine, phenylalanine, lysine, histidine, asparagine, glutamine, cysteine, methionine, arginine and tryptophan (38, 39), and invention of a myriad chemical variants of adenosine (50). Evolution created 1-, 2-, 6-, 7-, and 8-methyl adenosine, 6-dimethyl adenosine, inosine, 6-isopentenyl adenosine (hydroxylated and unhydroxylated), 2-thiomethylated adenosine variants, 6-glycinylcarbamoyl adenosine, cyclic 6-threonylcarbamoyl adenosine and 2'-O-methyl adenosine, 2'-deoxyadenosine, 2'-O-ribosyladenosine (phosphate), and more.

The evolutionary model maps concepts of biological evolution onto chemical processes. We say that during environmental wet-dry cycling: (a) a *generation* is a single cycle; (b) *heredity* is information passed from one generation to the next; (c) *information* is associated with non-random chemical composition; (d) *selection* is preferential inheritance of certain molecular compositions; (e) *fitness* is persistence of molecules and specific molecular assemblies; (f) *variation* is spatiotemporal differences in information; (g) an *individual* is a chemically isolated molecular ensemble; and (h) water is the "*energy currency"* that thermodynamically links and drives reactions. These models integrate evolutionary concepts of continuity (60), lack of foresight (4), exaptation (50, 61), symbiosis and co-evolution (44) into chemistry. Several alternative models of chemical evolution have been proposed (47, 62-65) that have critical features in common.

https://doi.org/10.26434/chemrxiv-2023-1jrcq-v2 **ORCID:** https://orcid.org/0000-0002-7215-4194 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

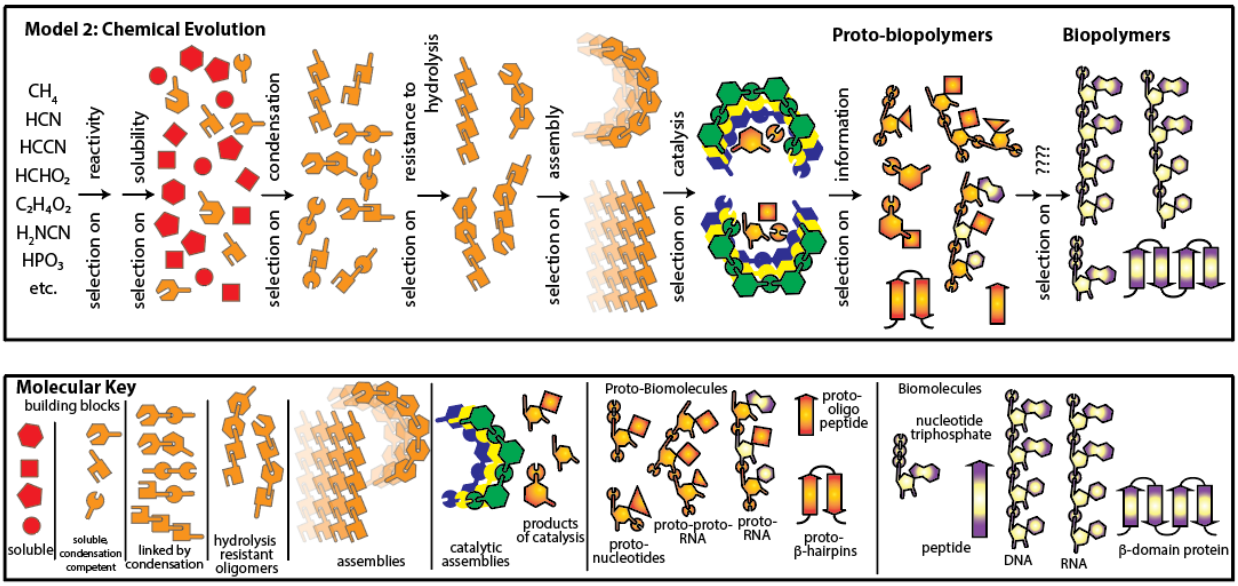

Figure 2. Chemical evolution, in which energy is harvested from environmental cycling and molecules are sculpted by unremitting selection. The basis of selection is fluid, as indicated. This model predicts that biopolymers are composed, at least in part, from building blocks that are not available by direct chemical synthesis. The bottom panel is a key explaining the molecular symbolism. This schematic omits some mechanisms of selection such as compartmentalization. Wet-dry, freeze-thaw or pressure cycling are possible drivers of chemical evolution.

In this model, complex mixtures of small molecules were sculpted and transformed during continuous chemical selection to yield biopolymers (47-49). The extant building blocks of biopolymers are fundamentally different from organic molecules accessible by direct synthesis, just as the skeleton of an extant elephant is fundamentally different from its ancestral fish skeleton. In this model, the origins of life can be understood only by a fusion of evolutionary theory and chemical sciences.

Water: The centrality of water in biochemistry, both as a medium and in chemical reactions, can also help us understand the chemical origins of life. Water is the most frequent and abundant chemical reagent in biology (66). Water is fully integrated into processes of biological bond making and bond breaking. Water chemically combines with, withdraws from, or intercedes less directly in all biochemical transformations. Between a third and a half of known biochemical reactions involve direct chemical consumption or production of water, and all universal biopolymers and most metabolites are produced by condensation-dehydration reactions. No other substance known to science is as abundant, with the capacity to play active roles as both a physical unique solvent, and as a hyperactive chemical reagent (67, 68). Cycling water activity can cause near-equilibrium reactions to oscillate in direction and to ratchet in energy and complexity (69). In our view, origins of life models should focus on scenarios where the

https://doi.org/10.26434/chemrxiv-2023-1jrcq-v2 **ORCID:** https://orcid.org/0000-0002-7215-4194 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

chemistry of water on a temperate spinning planet drives the evolution of complex chemical systems and subsequently biochemistry.

Selection: In the model proposed here, selection is intrinsic to evolution, both chemical and biological. Selection in chemical evolution, like selection in biology, is unremitting and relentless, yet dynamic and fluid. During chemical evolution, molecules were selected on varying combinations of (i) solubility in water, (ii) ability to link by condensation-dehydration during environmental cycling, (iii) chemical transitions into kinetically trapped (persistent) condensates, such as ester-amide exchange, (iv) resistance to hydrolysis by molecular assembly, and (v) autocatalysis. In this model, intense and mutable selection sparked the genesis of biopolymers.

Chickens and Eggs: Linkage is a norm in evolutionary processes. Advances ripple across and through systems and organisms. The nucleus is linked to the mitochondrion (71), the tibia is linked to the fibula (72), and the wasp is linked to the fig (73). Chemical evolution implies that RNA is linked with other biopolymers and with metabolism, and that all arose in concert. Evolutionary models discount proverbial chicken and egg dilemmas because selection is linked, and changes are coupled across broad fronts. Diverse feedstocks, proto-building blocks, and proto-biopolymers were inter-connected with each other and with primitive metabolism by the chemistry of water and other mechanisms.

Creativity: Evolutionary creativity (70) is seen in microbial metabolism, tetrapod limbs and primate brains. We agree with Jacob that creative molecular phase of chemical evolution preceded the ongoing creative sequence-based phase of Darwinian evolution (4). In this model, chemical evolution invented biological molecules, which are therefore inaccessible via direct chemical synthesis. Thus, the pathways from chemical feedstocks to biopolymers in models of chemical evolution differ fundamentally from in models of direct synthesis. Molecules created by chemical evolution, before the advent of Darwinian evolution, blur the distinction between prebiotic and biotic.

Multiple Models: Models of chemical evolution are new and are advancing rapidly (47, 62, 63). Wet-dry, freeze-thaw, and pressure cycling are possible drivers of chemical evolution (74) beyond Earth. Hud and coworkers described a model of chemical progression of proto-RNA to RNA, from simple to complex (49). Changes in chemical composition consistent with chemical evolution have been reported during wet-dry cycling (54, 59, 75-78). Unceasing chemical changes and exploration of new chemical spaces has been experimentally authenticated in prolonged wet-dry cycling (69). Baum and coworkers have explored chemical ecosystems (79). Huck has investigated effects of environmental changes on organized reaction systems (80). Mutually catalytic systems (79) and reproducing catalytic micelles (81) have been investigated. The importance of various parameters for chemical evolution (82), including

https://doi.org/10.26434/chemrxiv-2023-1jrcq-v2 **ORCID:** https://orcid.org/0000-0002-7215-4194 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

complexity and systems chemistry (83) have been discussed. Oscillatory networks of organic reactions are sustained by compositional heterogeneity, but not by homogeneity (84). Dynamic combinatorial chemistry has been used to discover a variety of functional molecules (85, 86). Auto-catalytic synthesis has been used for selecting functional molecules (87, 88). The majority of work on chemical evolution has taken place over the last 10 years; progress is accelerating.

Predictions of Origins by Chemical Evolution.

Models in which biopolymers are products of co-evolution with each other and with metabolism enable many predictions. The model proposed here predicts that all biopolymers are characterized by: (i) a unified chemistry of polymerization; (ii) thermodynamically unstable, kinetically trapped linkages, (iii) highly sophisticated proficiencies of assembly; (iv) homochirality, as selected by and required for assembly, (v) protection from degradation by assembly; (vi) integration at synthetic, structural, functional and metabolic levels; and (vii) divergence of biosynthetic pathways from inherent chemical reactivities. Chemical evolution predicts the incredible assemblies and long lifetimes of polysaccharide (cellulose persists for over 5,000 years (89)), of polypeptide (collagen persists for over 40,000 years (90)) and of RNA (with Goldilocks zones of persistence (91)). The unity of biopolymer assembly is illustrated in **Figure 3**.

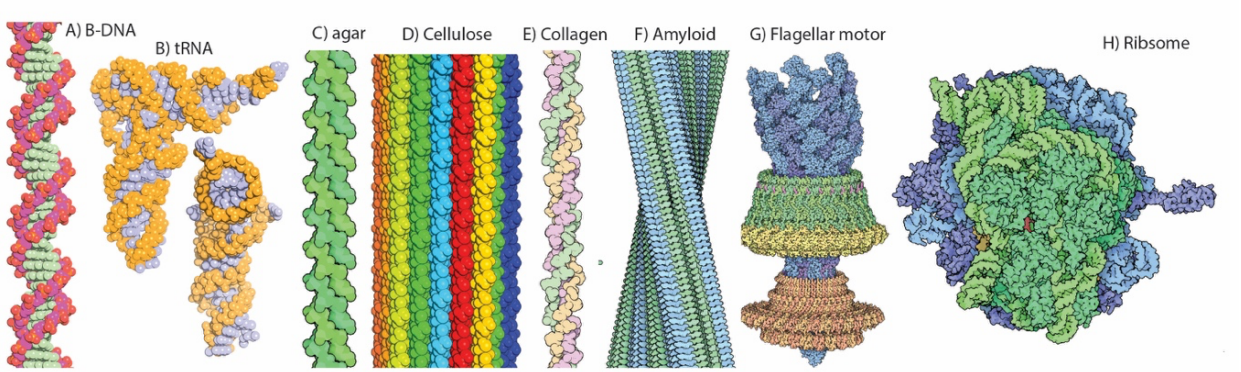

Figure 3. A unity of biopolymer synthesis, folding and protection predicted by chemical evolution. Each biopolymer is synthesized by condensation-dehydration chemistry, has sophisticated proficiency in folding and assembly, and is resistant to hydrolysis when folded and/or assembled. A) DNA, a double-helical polydeoxyribonucleotide. B) tRNA, a complex polyribonucleotide stabilized in part by double helices and in part by more complex interactions. C) Agar, a double-helical polysaccharide. D) Crystalline cellulose, a multistranded polysaccharide assembly. E) Collagen, a triple-helical polypeptide. F) An amyloid fiber composed of a helical assembly of β-sheet polypeptide. G) A flagellar motor, which is a pseudo-symmetric assembly of five distinct polypeptide chains. H) The ribosome, a large non-symmetric co-assembly of around 50 polypeptide chains and over 3,000 deoxyribonucleotides. Some of these images were produced by David S. Goodsell and the RCSB PDB.

https://doi.org/10.26434/chemrxiv-2023-1jrcq-v2 **ORCID:** https://orcid.org/0000-0002-7215-4194 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

Chemical evolutionary models envision creativity and innovation before the emergence of biopolymers; in chemical analogy with creativity and innovation in, for example, the biological conversion of fish fins to elephant legs (92) and the conversion of jellyfish nerve nets to human brains (93).

The predictions of chemical evolution appear to be consistent with observations of contemporary biology. Among these are the amazing folded structures, assemblies and co-assemblies the characterize biopolymers (polynucleotide, polypeptide, and polysaccharide, **Figure 3**) and the sophisticated assembly by biopolymers of elaborate structures such as DNA, tRNA, cellulose, and collagen. The profound integration of biological subsystems is indicated by the multiple roles for compounds such as ATP, which are central to and inextricable from both biopolymers and metabolic systems (94).Biopolymers are made by unified chemistry (phosphate-mediated condensation-dehydration) and are protected from hydrolysis by folding and assembly (48, 91). Integration is seen in the co-synthesis of biopolymers. RNA makes protein in the ribosome and protein makes RNA in polymerases (94). Integration is also seen in building block biosynthesis - five amino acids are consumed in the biosynthesis of one guanine (20, 94). Biopolymers and their building blocks are deeply integrated with each other and with metabolic systems, suggesting co-emergence of information and metabolism.

Weaknesses of Origins by Chemical Evolution.

The current model of chemical co-evolution has significant weaknesses. In comparison with direct synthesis models, very little effort has been invested in understanding chemical evolution. Therefore, the mechanisms of chemical evolution are not well constrained. The production of long complex polymers by chemical evolution has not been demonstrated. There are few laboratory examples of evolutionary formation of complex oligomers or polymers from small molecules by chemical evolution. Nor are the molecular mechanisms or duration of chemical evolution fully understood. Moreover, there are no realistic molecular models for the critical phase in which chemical evolution transitions to biological evolution.

Another weakness of this model is that the definition of fitness in chemical evolution models remains partially unresolved and appears more elastic than in Darwinian evolution. Fitness may refer, at some stages of chemical evolution, to the ability of fragile molecular systems/assemblies to persist under hydrolytic stress. At other stages fitness might refer to the ability to tune persistence by folding (91). Although models suggest that the molecular losers of chemical evolution are racemates of esters, thioesters, depsipeptides and thiodepsipeptides (25, 58, 59), other extinct intermediates are not characterized (49).

https://doi.org/10.26434/chemrxiv-2023-1jrcq-v2 **ORCID:** https://orcid.org/0000-0002-7215-4194 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

Summary

We have described two general models for the origins of biopolymers. In one model, biopolymers arose from intrinsic reactivities of prebiotic chemicals on the ancient Earth. This model can lead naturally to an RNA World; chemistry invents RNA then RNA invents evolution. In the second model, prolonged chemical evolution breaks the direct connection of prebiotic chemistry to biology. This model leads to a complex world of many players; evolution invented RNA as one component of an intensely integrated system of symbiotic biopolymers (44). Each of these models has strengths and weaknesses. We do not know which model, or a hybrid of them, is closer to reality. Many studies conducted to understand the origins of life have been guided by the direct synthesis model. Far less has been done to establish experimental methods focused on the evolutionary model. Understanding and controlling chemical evolution offers exciting possibilities. We know one thing for certain - future work directed toward understanding the origin of life will reveal fascinating new chemical phenomena.

Acknowledgements

This work was supported by NASA Grant 80NSSC18K1139 to LDW and the Azrieli Early Career Faculty Grant to MFP. We thank Ramanarayanan Krishnamurthy, Steve Benner, and Ulrich Muller for helpful discussions.

REFERENCES

1. Woese CR (2002) On the evolution of cells. *Proc Natl Acad Sci USA* **99**: 8742-8747.
2. El Baidouri F, Venditti C, Suzuki S, Meade A, & Humphries S (2020) Phenotypic reconstruction of the last universal common ancestor reveals a complex cell. *bioRxiv*.
3. Koonin EV (2009) Darwinian evolution in the light of genomics. *Nucleic Acids Res* **37**: 1011-1034.
4. Jacob F (1977) Evolution and tinkering. *Science* **196**: 1161-1166.
5. Schmitt-Kopplin P, Gabelica Z, Gougeon RD, Fekete A, Kanawati B, Harir M, Gebefuegi I, Eckel G, & Hertkorn N (2010) High molecular diversity of extraterrestrial organic matter in murchison meteorite revealed 40 years after its fall. *Proc Natl Acad Sci USA* **107**: 2763-2768.
6. Zahnle KJ, Lupu R, Catling DC, & Wogan N (2020) Creation and evolution of impact-generated reduced atmospheres of early earth. *Planet Sci J* **1**: 1-11.
7. Genda H, Brasser R, & Mojzsis S (2017) The terrestrial late veneer from core disruption of a lunar-sized impactor. *Earth Planet Sci Lett* **480**: 25-32.
8. Sutherland JD (2016) The origin of life—out of the blue. *Angew Chem Int Ed* **55**: 104-121.
9. Patel BH, Percivalle C, Ritson DJ, Duffy CD, & Sutherland JD (2015) Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism. *Nat Chem* **7**: 301-307.
10. Kim H-J & Benner SA (2021) Abiotic synthesis of nucleoside 5′-triphosphates with nickel borate and cyclic trimetaphosphate (ctmp). *Astrobiology* **21**: 298-306.

https://doi.org/10.26434/chemrxiv-2023-1jrcq-v2 **ORCID:** https://orcid.org/0000-0002-7215-4194 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

11. Benner SA, Bell EA, Biondi E, Brasser R, Carell T, Kim HJ, Mojzsis SJ, Omran A, Pasek MA, & Trail D (2020) When did life likely emerge on earth in an RNA-first process? *ChemSystemsChem* **2**: e1900035.
12. Orgel LE (2004) Prebiotic chemistry and the origin of the RNA world. *Crit Rev Biochem Mol Biol* **39**: 99-123.
13. Powner MW, Gerland B, & Sutherland JD (2009) Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* **459**: 239.
14. Oró J (1960) Synthesis of adenine from ammonium cyanide. *Biochem Biophys Res Commun* **2**: 407-412.
15. Benner SA (2018) Science addressing questions not easily addressed by "the" scientific method. The origins of life1. *Transformative Roles of Science in Society: From Emerging Basic Science Toward Solutions for People's Wellbeing* 123.
16. Kitadai N & Maruyama S (2018) Origins of building blocks of life: A review. *Geoscience Frontiers* **9**: 1117-1153.
17. Benner SA, Kim H-J, & Biondi E (2019) Prebiotic chemistry that could not not have happened. *Life* **9**: 84.
18. Sanchez RA & Orgel LE (1970) Studies in prebiotic synthesis: V. Synthesis and photoanomerization of pyrimidine nucleosides. *J Mol Biol* **47**: 531-543.
19. Ferris JP (2005) Mineral catalysis and prebiotic synthesis: Montmorillonite-catalyzed formation of RNA. *Elements* **1**: 145-149.
20. Yadav M, Kumar R, & Krishnamurthy R (2020) Chemistry of abiotic nucleotide synthesis. *Chem Rev* **120**: 4766-4805.
21. Szostak JW (2017) The narrow road to the deep past: In search of the chemistry of the origin of life. *Angew Chem Int Ed* **56**: 11037-11043.
22. Miller SL (1953) A production of amino acids under possible primitive earth conditions. *Science* **117**: 528-529.
23. Meierhenrich UJ, Munoz Caro GM, Bredehoft JH, Jessberger EK, & Thiemann WH (2004) Identification of diamino acids in the murchison meteorite. *Proc Natl Acad Sci USA* **101**: 9182-9186.
24. Forsythe JG, Yu SS, Mamajanov I, Grover MA, Krishnamurthy R, Fernandez FM, & Hud NV (2015) Ester-mediated amide bond formation driven by wet-dry cycles: A possible path to polypeptides on the prebiotic earth. *Angew Chem Int Ed* **54**: 9871-9875.
25. Frenkel-Pinter M, Fernandez FM, Leman LJ, Williams LD, Hud NV, & Guzman A (2022) Thioesters provide a robust path to prebiotic peptides. *Nature Communications* **13**: 1-8.
26. Islam S, Bučar D-K, & Powner MW (2017) Prebiotic selection and assembly of proteinogenic amino acids and natural nucleotides from complex mixtures. *Nat Chem* **9**: 584-589.
27. Oba Y, Takano Y, Furukawa Y, Koga T, Glavin DP, Dworkin JP, & Naraoka H (2022) Identifying the wide diversity of extraterrestrial purine and pyrimidine nucleobases in carbonaceous meteorites. *Nat Commun* **13**: 2008.
28. Furukawa Y, Chikaraishi Y, Ohkouchi N, Ogawa NO, Glavin DP, Dworkin JP, Abe C, & Nakamura T (2019) Extraterrestrial ribose and other sugars in primitive meteorites. *Proc Natl Acad Sci USA* **116**: 24440-24445.
29. Shapiro R (2007) A simpler origin for life. *Sci Am* **296**: 46-53.
30. Bada JL & Lazcano A (2003) Prebiotic soup--revisiting the miller experiment. *Science* **300**: 745-746.
31. Johnson AP, Cleaves HJ, Dworkin JP, Glavin DP, Lazcano A, & Bada JL (2008) The Miller volcanic spark discharge experiment. *Science* **322**: 404.
32. Dowler MJ, Fuller WD, Orgel LE, & Sanchez RA (1970) Prebiotic synthesis of propiolaldehyde and nicotinamide. *Science* **169**: 1320-1321.

https://doi.org/10.26434/chemrxiv-2023-1jrcq-v2 **ORCID:** https://orcid.org/0000-0002-7215-4194 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

33. Bonfio C, Valer L, Scintilla S, Shah S, Evans DJ, Jin L, Szostak JW, Sasselov DD, Sutherland JD, & Mansy SS (2017) Uv-light-driven prebiotic synthesis of iron–sulfur clusters. *Nat Chem* **9**: 1229-1234.
34. Shen C, Yang L, Miller S, & Oró J (1990) Prebiotic synthesis of histidine. *J Mol Evol* **31**: 167-174.
35. Krishnamurthy R (2018) Experimentally investigating the origin of DNA/RNA on early earth. *Nat Commun* **9**: 1-3.
36. Benner SA (2018) Prebiotic plausibility and networks of paradox-resolving independent models. *Nat Commun* **9**: 1-3.
37. Toner JD & Catling DC (2019) Alkaline lake settings for concentrated prebiotic cyanide and the origin of life. *Geochim Cosmochim Acta* **260**: 124-132.
38. Mayer-Bacon C, Meringer M, Havel R, Aponte JC, & Freeland S (2022) A closer look at non-random patterns within chemistry space for a smaller, earlier amino acid alphabet. *J Mol Evol* **90**: 307-323.
39. Wong JT-F & Bronskill PM (1979) Inadequacy of prebiotic synthesis as origin of proteinous amino acids. *J Mol Evol* **13**: 115-125.
40. Harrison SA & Lane N (2018) Life as a guide to prebiotic nucleotide synthesis. *Nat Commun* **9**: 1-4.
41. Krishnamurthy R (2020) Systems chemistry in the chemical origins of life: The 18th camel paradigm. *J Syst Chem* **8**: 40-62.
42. Cafferty BJ, Fialho DM, Khanam J, Krishnamurthy R, & Hud NV (2016) Spontaneous formation and base pairing of plausible prebiotic nucleotides in water. *Nat Commun* **7**: 11328.
43. Van Der Valk T, Pečnerová P, Díez-Del-Molino D, Bergström A, Oppenheimer J, Hartmann S, Xenikoudakis G, Thomas JA, Dehasque M, & Sağlıcan E (2021) Million-year-old DNA sheds light on the genomic history of mammoths. *Nature* **591**: 265-269.
44. Lanier KA, Petrov AS, & Williams LD (2017) The central symbiosis of molecular biology: Molecules in mutualism. *J Mol Evol* **85**: 8-13.
45. Ec M (2010) Examining life's origins. *Thesis, University of California, San Diego*.
46. Wolf YI & Koonin EV (2007) On the origin of the translation system and the genetic code in the RNA world by means of natural selection, exaptation, and subfunctionalization. *Biol Direct* **2**: 1-25.
47. Szathmáry E & Smith JM (1995) The major evolutionary transitions. *Nature* **374**: 227-232.
48. Runnels CM, Lanier KA, Williams JK, Bowman JC, Petrov AS, Hud NV, & Williams LD (2018) Folding, assembly, and persistence: The essential nature and origins of biopolymers. *J Mol Evol* **86**: 598-610.
49. Hud NV, Cafferty BJ, Krishnamurthy R, & Williams LD (2013) The origin of RNA and "my grandfather's axe". *Chem Biol* **20**: 466-474.
50. Frenkel-Pinter M, Petrov AS, Matange K, Travisano M, Glass JB, & Williams LD (2022) Adaptation and exaptation: From small molecules to feathers. *J Mol Evol* **90**: 166-175.
51. Becker S, Schneider C, Okamura H, Crisp A, Amatov T, Dejmek M, & Carell T (2018) Wet-dry cycles enable the parallel origin of canonical and non-canonical nucleosides by continuous synthesis. *Nat Commun* **9**: 163.
52. Mamajanov I, Macdonald PJ, Ying J, Duncanson DM, Dowdy GR, Walker CA, Engelhart AE, Fernandez FM, Grover MA, & Hud NV (2014) Ester formation and hydrolysis during wet–dry cycles: Generation of far-from-equilibrium polymers in a model prebiotic reaction. *Macromolecules* **47**: 1334-1343.
53. Damer B & Deamer D (2020) The hot spring hypothesis for an origin of life. *Astrobiology* **20**: 429-452.
54. Ross DS & Deamer D (2016) Dry/wet cycling and the thermodynamics and kinetics of prebiotic polymer synthesis. *Life* **6**.
55. Ross D & Deamer D (2019) Prebiotic oligomer assembly: What was the energy source? *Astrobiology* **19**: 517-521.

https://doi.org/10.26434/chemrxiv-2023-1jrcq-v2 **ORCID:** https://orcid.org/0000-0002-7215-4194 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

56. Li Z, Li L, Mckenna KR, Schmidt M, Pollet P, Gelbaum L, Fernández FM, Krishnamurthy R, & Liotta CL (2019) The oligomerization of glucose under plausible prebiotic conditions. *Orig Life Evol Biosph* **49**: 225-240.
57. Frenkel-Pinter M, Sargon AB, Glass JB, Hud NV, & Williams LD (2021) Transition metals enhance prebiotic depsipeptide oligomerization reactions involving histidine. *RSC Advances* **11**: 3534-3538.
58. Frenkel-Pinter M, Haynes JW, Mohyeldin AM, Martin C, Sargon AB, Petrov AS, Krishnamurthy R, Hud NV, Williams LD, & Leman LJ (2020) Mutually stabilizing interactions between proto-peptides and RNA. *Nat Commun* **11**: 1-14.
59. Frenkel-Pinter M, Haynes JW, C M, Petrov AS, Burcar BT, Krishnamurthy R, Hud NV, Leman LJ, & Williams LD (2019) Selective incorporation of proteinaceous over nonproteinaceous cationic amino acids in model prebiotic oligomerization reactions. *Proc Natl Acad Sci USA* 201904849.
60. Wolf YI & Koonin EV (2007) On the origin of the translation system and the genetic code in the RNA world by means of natural selection, exaptation, and subfunctionalization. *Biol Direct* **2**: 1-25.
61. Gould SJ & Vrba ES (1982) Exaptation—a missing term in the science of form. *Paleobiology* **8**: 4-15.
62. Ruiz-Mirazo K, Briones C, & De La Escosura A (2017) Chemical roots of biological evolution: The origins of life as a process of development of autonomous functional systems. *Open biology* **7**: 170050.
63. Fried SD, Fujishima K, Makarov M, Cherepashuk I, & Hlouchova K (2022) Peptides before and during the nucleotide world: An origins story emphasizing cooperation between proteins and nucleic acids. *Journal of the Royal Society Interface* **19**: 20210641.
64. Baum DA & Vetsigian K (2017) An experimental framework for generating evolvable chemical systems in the laboratory. *Orig Life Evol Biosph* **47**: 481-497.
65. Makarov M, Sanchez Rocha AC, Krystufek R, Cherepashuk I, Dzmitruk V, Charnavets T, Faustino AM, Lebl M, Fujishima K, & Fried SD (2022) Early selection of the amino acid alphabet was adaptively shaped by biophysical constraints of foldability. *bioRxiv* 2022.2006. 2014.495995.
66. Frenkel-Pinter M, Rajaei V, Glass JB, Hud NV, & Williams LD (2021) Water and life: The medium is the message. *J Mol Evol* 1-10.
67. Ball P (2017) Water is an active matrix of life for cell and molecular biology. *Proc Natl Acad Sci USA* **114**: 13327-13335.
68. Brack A (1987) Selective emergence and survival of early polypeptides in water. *Orig Life Evol Biosph* **17**: 367-379.
69. Frenkel-Pinter M, Matange K, Rajaei V, Costner JT, Robertson A, Kim JS, Petrov AS, Bowman JC, & Williams LD (2023) Chemical evolution reimagined. *ChemRxiv* 10.26434/chemrxiv-22022-s26433cr26432-v26432.
70. Gould SJ (2002) *The structure of evolutionary theory* (Harvard University Press).
71. Roger AJ, Muñoz-Gómez SA, & Kamikawa R (2017) The origin and diversification of mitochondria. *Current Biology* **27**: R1177-R1192.
72. Botelho JF, Smith-Paredes D, Soto-Acuna S, O'connor J, Palma V, & Vargas AO (2016) Molecular development of fibular reduction in birds and its evolution from dinosaurs. *Evolution* **70**: 543-554.
73. Machado CA, Robbins N, Gilbert MTP, & Herre EA (2005) Critical review of host specificity and its coevolutionary implications in the fig/fig-wasp mutualism. *Proc Natl Acad Sci USA* **102**: 6558-6565.
74. Root-Bernstein R & Brown AW (2022) Novel apparatuses for incorporating natural selection processes into origins-of-life experiments to produce adaptively evolving chemical ecosystems. *Life* **12**: 1508.
75. Forsythe JG, Petrov AS, Millar WC, Yu SS, Krishnamurthy R, Grover MA, Hud NV, & Fernandez FM (2017) Surveying the sequence diversity of model prebiotic peptides by mass spectrometry. *Proc Natl Acad Sci USA* **114**: E7652-E7659.

https://doi.org/10.26434/chemrxiv-2023-1jrcq-v2 **ORCID:** https://orcid.org/0000-0002-7215-4194 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

76. Frenkel-Pinter M, Bouza M, Fernández FM, Leman LJ, Williams LD, Hud NV, & Guzman-Martinez A (2022) Thioesters provide a plausible prebiotic path to proto-peptides. *Nat Commun* **13**: 1-8.
77. Yu SS, Krishnamurthy R, Fernandez FM, Hud NV, Schork FJ, & Grover MA (2016) Kinetics of prebiotic depsipeptide formation from the ester-amide exchange reaction. *Physical Chemistry Chemical Physics* **18**: 28441-28450.
78. Foster K, Hillman B, Rajaei V, Seng K, & Maurer S (2022) Evolution of realistic organic mixtures for the origins of life through wet–dry cycling. *Sci* **4**: 22.
79. Vincent L, Berg M, Krismer M, Saghafi ST, Cosby J, Sankari T, Vetsigian K, Cleaves HJ, & Baum DA (2019) Chemical ecosystem selection on mineral surfaces reveals long-term dynamics consistent with the spontaneous emergence of mutual catalysis. *Life* **9**: 80.
80. Robinson WE, Daines E, Van Duppen P, De Jong T, & Huck WTS (2022) Environmental conditions drive self-organization of reaction pathways in a prebiotic reaction network. *Nat Chem* **14**: 623-631.
81. Kahana A & Lancet D (2021) Self-reproducing catalytic micelles as nanoscopic protocell precursors. *Nature Reviews Chemistry* **5**: 870-878.
82. Nghe P, Hordijk W, Kauffman SA, Walker SI, Schmidt FJ, Kemble H, Yeates JaM, & Lehman N (2015) Prebiotic network evolution: Six key parameters. *Molecular BioSystems* **11**: 3206-3217.
83. Ruiz-Mirazo K, Briones C, & De La Escosura A (2014) Prebiotic systems chemistry: New perspectives for the origins of life. *Chem Rev* **114**: 285-366.
84. Cafferty BJ, Wong AS, Semenov SN, Belding L, Gmur S, Huck WT, & Whitesides GM (2019) Robustness, entrainment, and hybridization in dissipative molecular networks, and the origin of life. *J Am Chem Soc* **141**: 8289-8295.
85. Lam RT, Belenguer A, Roberts SL, Naumann C, Jarrosson T, Otto S, & Sanders JK (2005) Amplification of acetylcholine-binding catenanes from dynamic combinatorial libraries. *Science* **308**: 667-669.
86. Liu B, Pappas CG, Zangrando E, Demitri N, Chmielewski PJ, & Otto S (2018) Complex molecules that fold like proteins can emerge spontaneously. *J Am Chem Soc* **141**: 1685-1689.
87. Wołos A, Roszak R, Żądło-Dobrowolska A, Beker W, Mikulak-Klucznik B, Spólnik G, Dygas M, Szymkuć S, & Grzybowski BA (2020) Synthetic connectivity, emergence, and self-regeneration in the network of prebiotic chemistry. *Science* **369**.
88. Miao X, Paikar A, Lerner B, Diskin-Posner Y, Shmul G, & Semenov SN (2021) Kinetic selection in the out-of-equilibrium autocatalytic reaction networks that produce macrocyclic peptides. *Angew Chem* **133**: 20529-20538.
89. Reynaud C, Thoury M, Dazzi A, Latour G, Scheel M, Li J, Thomas A, Moulherat C, Didier A, & Bertrand L (2020) In-place molecular preservation of cellulose in 5,000-year-old archaeological textiles. *Proc Natl Acad Sci USA* **117**: 19670-19676.
90. Hattori S, Kiriyama-Tanaka T, Kusubata M, Taga Y, Ebihara T, Kumazawa Y, Imai K, Miura M, Mezaki Y, & Tikhonov A (2021) Preservation of collagen in the soft tissues of frozen mammoths. *PLoS One* **16**: e0258699.
91. Guth-Metzler R, Mohamed AM, Cowan ET, Henning A, Ito C, Frenkel-Pinter M, Wartell RM, Glass JB, & Williams LD (2023) Goldilocks and RNA: Where Mg^{2+} concentration is just right. *Nucleic Acids Res* **51**: 3529-3539.
92. Amemiya CT, Alföldi J, Lee AP, Fan S, Philippe H, Maccallum I, Braasch I, Manousaki T, Schneider I, & Rohner N (2013) The african coelacanth genome provides insights into tetrapod evolution. *Nature* **496**: 311-316.
93. Arendt D, Tosches MA, & Marlow H (2016) From nerve net to nerve ring, nerve cord and brain—evolution of the nervous system. *Nature Reviews Neuroscience* **17**: 61-72.
94. Nelson DL, Lehninger AL, & Cox MM (2021) *Principles of Biochemistry* (Macmillan).

https://doi.org/10.26434/chemrxiv-2023-1jrcq-v2 **ORCID:** https://orcid.org/0000-0002-7215-4194 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

https://doi.org/10.26434/chemrxiv-2023-1jrcq-v2 **ORCID:** https://orcid.org/0000-0002-7215-4194 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0