PHILOSOPHICAL TRANSACTIONS B

royalsocietypublishing.org/journal/rstb

Opinion piece

Cite this article: Carbonell P, Peretó J. 2025 Before LUCA: unearthing the chemical roots of metabolism. *Phil. Trans. R. Soc. B* **380**: 20240292. https://doi.org/10.1098/rstb.2024.0292

Received: 7 October 2024
Accepted: 19 November 2024

One contribution of 15 to a theme issue 'Origins of life: the possible and the actual'.

Subject Areas:
evolution, synthetic biology, biochemistry, bioinformatics

Keywords:
LUCA, prebiotic chemistry, protometabolism, enzyme promiscuity, systems chemistry, generative metabolism

Authors for correspondence:
Pablo Carbonell
e-mail: pablo.carbonell@csic.es
Juli Peretó
e-mail: juli.pereto@uv.es

Before LUCA: unearthing the chemical roots of metabolism

Pablo Carbonell[1] and Juli Peretó[1,2]

[1]Institute for Integrative Systems Biology I2SysBio (Universitat de València-CSIC), Paterna 46980, Spain
[2]Department of Biochemistry and Molecular Biology, Universitat de València, 46100 Burjassot, Spain

PC, 0000-0002-0993-5625; JP, 0000-0002-5756-1517

The transition from geochemistry to biochemistry marks a fundamental event in the emergence of life. Recent advancements in computational methods and systems chemistry offer new insights into how primitive biochemical systems evolved. While early Earth chemistry laid the foundation for protometabolic processes, linking these to the more complex metabolic capabilities of last universal common ancestor (LUCA) remains a challenge. Here, we propose potential solutions by mapping reaction rules from LUCA's metabolism to prebiotic reactions, enabling the exploration of possible connections between abiotic chemistry and early life. A comprehensive modelling framework is proposed to test these hypotheses, incorporating factors such as thermodynamics, planetary conditions and the availability of key reactants. Additionally, the role of enzyme promiscuity and metabolic evolvability is discussed as a driving force in the transition from early chemistry to life. By leveraging generative models, this research paves the way for future synthetic biology applications, potentially expanding our understanding of metabolism and its evolutionary origins.

This article is part of the theme issue 'Origins of life: the possible and the actual'.

1. Introduction

The Universe has synthesized all possible chemical elements from available subatomic particles, leading to a universal periodic table. If the Universe were to begin again under the same conditions, the same particles and forces, the same set of elements would emerge [1]. This reflects the Universe's ergodic nature, where all possible combinations of elementary particles have been realized. However, when it comes to chemistry—the combination of atoms into more or less stable molecules—factors like historical contingency, environmental conditions, structural constraints and the inherent stability of molecular structures narrow the range of possibilities. While theoretically there exists a vast array of possible molecules, only a limited set are feasible under real-world conditions. The realm of chemical possibilities is vast but largely unexplored owing to material and temporal limitations. These enormous combinatorial possibilities, termed 'hyper-astronomical' by Kauffman [2], underscore how physicochemical constraints and historical events limit the space of possible solutions, ultimately shaping what is real and what will ever exist at the molecular scale.

From a chemical perspective, one could argue that the Universe is not ergodic. As we consider more complex structures, such as biomacromolecules, proteins and nucleic acids, additional restrictions emerge, further limiting the space of possible structural solutions. As Alberch suggested, there is a 'logic of monsters'—not all anatomical forms that we can imagine are feasible [3]. This same principle applies to molecules. Studies of meteorites and abiotic organic syntheses conducted under prebiotic conditions suggest that early Earth

THE ROYAL SOCIETY PUBLISHING

© 2025 The Author(s). Published by the Royal Society. All rights reserved.

Downloaded from http://royalsocietypublishing.org/rstb/article-pdf/doi/10.1098/rstb.2024.0292/2821375/rstb.2024.0292.pdf
by University of Valencia (Universidad de Valencia user
on 13 January 2026

hosted a moderately complex chemodiversity, including the most common biological building blocks. During this period of chemical evolution, physicochemical constraints—especially thermodynamics and kinetics in specific environments—dictated which molecules were synthesized, persisted, and interacted with each other, shaping the abiotic chemical landscape.

The chemical conditions of early Earth laid the groundwork for the organization of the most primitive biochemical systems, including autocatalytic cycles that produced essential building blocks, self-reproducing lipid vesicles and genetic templates capable of self-replication. These initial processes likely relied on simple catalysts (e.g. mineral surfaces and small organometallic compounds). The emergence of catalysts that could be optimized by natural selection, such as ribozymes, marked a crucial phase transition, enabling more efficient and innovative exploration of chemical possibilities. However, it is essential to recognize that the synthesis of actual molecules represents only a minuscule fraction of the vast chemical possibility space. Given the hyper-astronomical number of potential molecules, even in our non-ergodic world—where most complex molecules will never be realized—life has managed, over nearly 4 billion years (a relatively brief period for such an immense survey), to innovate and explore a remarkable diversity of functions, forms and physiologies. In this article, we consider how metabolic innovation began during the origin of life and how synthetic biology might push the boundaries of naturally evolved metabolism.

2. The chemical beginnings of life

The molecular diversity in extraterrestrial samples is extensive, spanning a wide range of organic molecules from simple hydrocarbons, like methane, to large molecules, such as amino acids, nucleobases, polycyclic aromatic hydrocarbons and long-chain hydrocarbons. A non-targeted ultrahigh-resolution molecular analysis of the solvent-accessible organic fraction of the Murchison meteorite (a carbonaceous chondrite) allowed the identification of tens of thousands of different molecular compositions and several millions of different chemical compounds, the product of abiotic chemistry involving the six major elements involved in life, namely, C, H, O, N, S and P [4]. A significant fraction of this chemodiversity was present on the early Earth owing to the high frequency of asteroid, comet and interstellar dust particles' impacts, especially during the Archaean. Most authors assume that this chemical diversity, together with the organic matter produced by planetary geochemistry, which has been the object of study of classical prebiotic chemistry [5], was necessary for the origin of primitive cells, which were made up of a small subset of all these molecules. However, the presence in cosmochemical matter of the building blocks we observe in biochemistry does not necessarily imply that the most primitive metabolism was nourished by them.

Take the case of nucleobases. Using state-of-the-art analytical techniques, Oba *et al.* [6] identified canonical purine and pyrimidine nucleobases (i.e. the set of five used by life in RNA and DNA) and a diversity of non-canonical nitrogenous bases present in different meteorites. As remarked by Krishnamurthy *et al.* [7], these observations underscore the universal nature of the abiotic chemistry responsible for the formation of heterocyclic compounds, but they do not, on their own, explain the chemical, geological or evolutionary processes that would ultimately produce the specific compounds and reaction networks that led to life. This discussion can be extended to other classes of biomolecules present in extraterrestrial organic matter, such as amino acids, sugars or fatty acids. The underlying debate, however, is whether the origin of life involved a selection of some of these components amidst a great chemical diversity or, rather, abiotic chemistry provided chemical evolution with a much less diverse set of molecules. The results of a systems approach to prebiotic chemistry [8] seem to indicate that this second possibility is more plausible than has been assumed until recently. As Sutherland and coworkers have shown, nucleotides and amino acids can both be made by cyanosulfidic protometabolism [9]. Remarkably, this chemical network does not produce an intractable mixture of products but rather a selective set of chemical species representative of many biomolecular components [10]. Additionally, a connection of the cyanosulfidic chemistry to a very specific geochemical scenario on early Earth has been suggested [11]. Another promising protometabolic map, explored by the Krishnamurthy and Springsteen groups, is inspired by the 'glyoxylate scenario' first suggested by Eschenmoser [12] and based on the interplay of HCN and aldehyde chemistries. Initially, they found the HKG (4-hydroxy-2-ketoglutarate)–MAL (malonate) bi-cycle, connected through a common intermediate, oxaloacetate [13]. The occurrence of some di- and tricarboxylic acids present in the TCA (tricarboxylic acid) or Krebs cycle that are involved in reactions analogous to the ones used by this central metabolic pathway could indicate that these cycles of abiotic, non-enzymatic reactions might represent a chemical vicinity of the TCA cycle that could have been present before the acquisition of complex catalysts. Furthermore, in the presence of simple nitrogenous compounds (ammonia, urea) some intermediaries of the protometabolic analogues of the TCA cycle led to the formation of amino acids from the corresponding alpha-ketoacids, as well as to the non-canonical nucleobase orotic acid [14].

Compared with the notable experiments by Moran and coworkers on non-enzymatic transformations between TCA cycle intermediates (for a review, see [15]), the HKG–MAL bi-cycle operates at moderate temperatures and without the need for metal ions, which, while promoting synthesis, can also lead to the degradation of intermediates [16]. What makes these protometabolic cycles particularly exciting is that they were discovered by adopting an agnostic approach, deliberately ignoring established metabolic pathways and focusing instead on what is chemically possible under prebiotic conditions [17]. It is tempting to suggest that this kind of prebiotic exploration on the primitive Earth provided a foundation upon which catalysts—such as ribozymes, simple or complex peptides or their combinations—could have emerged during the earliest stages of metabolic evolution. Thus, as proposed by the semi-enzymatic model, a combination of spontaneous reactions and reactions promoted by simple catalysts—including the pre-enzymatic action of coenzymes [18]—would configure these primitive chemical networks [19], whereas the onset of natural selection [20], as ribozymes began to co-opt reactions present in the protometabolic network, marked a crucial phase transition from pure chemical determinism to evolution by natural selection. This transition was significant not only because catalysts became subject to optimization—although the outcome is not always the best possible solution, as suboptimal or imperfect solutions may still suffice for survival and reproduction [21]—but also because their inherently flexible structures enabled functional promiscuity, allowing the exploration of new chemical reactions and products within the network.

Downloaded from http://royalsocietypublishing.org/rstb/article-pdf/doi/10.1098/rstb.2024.0292/2821375/rstb.2024.0292.pdf
by University of Valencia (Universidad de Valencia user
on 13 January 2026

True autocatalytic cycles involving different intermediates within these protometabolic landscapes have not yet been discovered. However, it has been proposed that glyoxylate-derived chemical networks (the so-called glyoxylose reaction by Krishnamurthy & Liotta [22]) have the potential to generate autocatalytic cycles. Several authors have examined the significance of autocatalysis in metabolism and its evolutionary origins [23]. Others, like Pross & Pascal [24,25], have proposed a theoretical model, suggesting minimal autocatalytic units composed of kinetically stable–thermodynamically activated (KSTA) compounds that achieve a state of global kinetic stability, allowing them to persist over time. The frozen reactivity of the KSTA molecules is made real by catalysts (e.g. minerals, organometallic compounds, ribozymes or enzymes), making useful the chemical energy that is stored in them [26], and we can recognize that metabolism is built on this fundamental property that biochemists also name as 'metastability'. We propose that metastable compounds or KSTAs are older than life and were present in prebiotic times, albeit in small proportions amidst abiotic chemodiversity. The autocatalytic entities made of KSTA elements are thought to originate within protometabolism. Their association with compartments—specifically, vesicles self-assembled from amphiphilic molecules generated by the network and capable of self-reproduction—and self-replicative genetic templates—also derived from protometabolic monomers—could allow the rise of the most primitive cells. In any case, at this point we do not intend to opt for any of the classic alternatives (metabolism first versus genetics first versus compartments first), as these are debates that, in our opinion, seem outdated and scientifically unfruitful (on the false dilemma of metabolism versus genetics first, see Negrón-Mendoza *et al.* [27]).

Krishnamurthy and his colleagues have proposed a series of interconnected protometabolic cycles based on glyoxylate chemistry, mirroring the structure of modern metabolism [14]. Their model starts with an inner layer of reactions involving carbon, hydrogen and oxygen compounds, followed by a second layer that incorporates nitrogen, leading to the formation of amino acids. From these, a third layer emerges, producing nucleobases. This framework represents a promising draft for a protometabolic map, which, when fully developed, could reveal various control mechanisms within these interconnected pathways essential for the establishment of protocells with a minimal metabolic system [28].

In summary, a widely accepted conceptual framework proposes that the cosmo- and geochemical environments of the early Earth would constitute a vast chemical diversity as the cradle of life, a large complex network of chemical reactions and species that would later have to undergo a pruning process by early catalysts to give rise to the (less complex) structure of the first minimal metabolisms [29]. This idea is now confronted by a wide range of empirical data from systems chemistry that suggest, on the contrary, an emergence under prebiotic conditions of a rather limited number of chemical species, a 'sparse' protometabolic network, as suggested by Smith & Morowitz [30], albeit not as a result of selection by primitive catalysts but as a constitutive property of an early protometabolism relatively simple *ab initio* compared with the universe of the chemically possible.

3. Bridging the gap between LUCA and protometabolism

In the evocative final passage of his seminal 1924 booklet, Oparin recounts the work of two scientific teams studying the origin of life, as if drilling a tunnel from the two ends: one from chemistry to biology and, on the opposite side, one from biology to chemistry [31]. One hundred years later, tracing the early evolution of life still relies on these two key strategies: reconstructing the evolutionary history encoded in genomes through the prokaryotic tree of life (the top-down approach), and empirical exploration of the network of prebiotic chemical reactions (the bottom-up approach). Both biological and chemical perspectives offer valuable insights into the evolution of early life. Chemistry helps to shed light on the early stages of the history of life, while more recent evolutionary events can be studied through the evolution of genomes and enzymes. Today, emerging insights from systems chemistry, combined with advanced computational methods and large chemical datasets, bring us closer than ever to modelling how cells arose from an initially abiotic world driven by deterministic chemical processes.

One of the most challenging knowledge gaps in studying the early evolution of life is bridging the connection between the protometabolism inferred from prebiotic systems chemistry and the ancestral LUCA metabolism reconstructed through phylogenetic analysis. Lazcano & Miller [19] compared this gap to a cloud that obscures these connections, making it difficult to determine the extent of chemical predetermination in the origin of metabolic pathways and the 'congruence' between protometabolism and metabolism as proposed by de Duve [32]. A potential strategy to address this could involve projecting backward the chemical potentialities of metabolism, considering that older enzymes might have been less specific or more promiscuous. This approach could illuminate the fuzzier chemistries that preceded the more defined map of metabolism, which in turn could guide empirical exploration in systems chemistry.

(a) Setting the clock back in enzyme evolution across the tree of life towards LUCA

One of the primary goals of the top-down approach is the reconstruction of LUCA. Since the pioneering work by Lazcano *et al.* [33] and the subsequent expansion of genomic data, at least 10 studies have attempted such a reconstruction [34]. The use of varying datasets, alongside continuously improving computational and statistical methodologies, and the consideration of phenomena such as horizontal gene transfer (HGT) and gene loss, have led to a wide range of results and differing views on LUCA's genetic and metabolic features. Nevertheless, the general consensus depicts LUCA as a cell type with complexity comparable to that of modern prokaryotes. A common finding across all studies is the conservation of the translation machinery [34]. However, there is less agreement regarding LUCA's metabolism, particularly whether it was autotrophic. Crapitto *et al.* [35] have proposed a consensus on LUCA's functional repertoire based on eight different studies of its genome or proteome reconstruction. This consensus suggests LUCA had an almost complete central metabolism, including the core pathways for amino acid and nucleotide metabolism, but details surrounding lipid metabolism and specialized pathways, such as terpene biosynthesis, remain unresolved. On the other

Downloaded from http://royalsocietypublishing.org/rstb/article-pdf/doi/10.1098/rstb.2024.0292/2821375/rstb.2024.0292.pdf
by University of Valencia (Universidad de Valencia user
on 13 January 2026

hand, dating LUCA is a complex task that involves calibrating molecular clocks using HGT inferred from phylogenetic trees that can be matched with geochemical or fossil signatures [36].

One recent study by Moody *et al.* [37] based on phylogenetic reconciliation reconstructed a core set of approximately 400 clusters of genes (COGs) that were likely to be present in LUCA. The result is a picture of a prokaryotic organism that used a primitive Wood–Ljungdahl pathway for acetogenic growth and carbon fixation, with gluconeogenesis and glycolysis capabilities and an early Cas system as an ancestral immune system. The set of 400 LUCA COGs contains a primaeval set of enzymes from which, based on the assumption of the evolutionary model, the extant metabolism of life on Earth could have evolved. The evolution of enzymes is often associated with their promiscuity [38], as seen in many modern enzymes that can accept multiple substrates or catalyse diverse reactions. These versatile functionalities were likely present in ancestral enzymes before evolutionary divergence led to the specialization of distinct functions. Gaining a better understanding of the latent catalytic capabilities of LUCA enzymes might shed new light on how biochemical pathways shaping extant metabolism emerged through evolution.

(b) Mapping the ancestral clusters of genes in LUCA to their generative clusters of reactions

A strategy for modelling the promiscuous capabilities of the core set of enzymes in LUCA is based on surveying the reaction rules associated with each enzyme [39], specifically through a representation of the chemical transformation around the reaction centre [40]. Reaction rules are generative (box 1) and often used to characterize the potential promiscuity of the enzyme, as they can provide an enumeration of the different substrates that might be potentially processed by the reactions that are latently present in the enzyme. The generalization capabilities of the rule can be controlled in the model by means of the diameter, which is a parameter that represents the distance of the reaction rule to the centre. Generative metabolism can, thus, provide informed hints about the size of the total biosynthetic space that can be reached by present enzymes. This estimate thus provides a view of the actual limits and capabilities of organisms, notably as we try to evolve or even synthesize them (box 2).

Box 1: The actual and the possible metabolisms

(a) *Generative metabolism: decoding evolvability*

A key feature of metabolic networks that might help unravel their early history is their evolvability and the evolutionary advantages this property conferred on organisms, enabling them to thrive in otherwise adverse environments [41]. Evolvability might be present at several levels, but perhaps the most important is the one associated with the plasticity of enzymes due to their versatility and promiscuity, which means that alternative pathways are latently present and conferring metabolism with a robust adaptive system to environmental changes [42]. Enzyme promiscuity thus is seen as a powerful mechanism for understanding metabolism. If we could master it, promiscuity might also serve as a critical driver in synthetic biology, mainly in the directed evolution of enzymes [43,44]. There are many examples of studies providing mechanistic explanations about enzyme promiscuity based on structural analysis of substrate interactions [45]. The field has most recently exploded thanks to the readily available protein structures from AlphaFold [46]. However, even if these studies can pose strong hypotheses about a particular enzymatic affinity or catalytic efficiency, they have proved elusive for generalization through a clear modelling pipeline. To overcome such a limitation, a more powerful approach has recently been developed based on generative models [47,48]. Generative models for enzyme promiscuity can predict, sometimes with a striking level of accuracy, alternative substrates or reactions that can be potentially catalysed by an enzyme. The generative approach has lately become widely adopted by the community thanks to the widespread availability of trained large language models [49–51]. Nonetheless, generative biochemical networks have a rich tradition that predates large language models and stemmed from bioretrosynthesis, and eventually from retrosynthesis approaches, based on reaction rules in organic synthesis that earned Corey a Nobel Prize in 1990 [52].

Generative metabolism based on reaction rules is rooted in the key finding that enzyme promiscuity is a feature that can be more clearly explained through the chemical side of a biochemical reaction rather than by looking at the enzyme sequence [39]. Enzyme sequences can be altered to accept a different substrate through residue substitution. However, there is no obvious way to guess which substrates, even within a class of compounds sharing common chemical features, would be accepted by an enzyme. As studies have shown, such ability can be more advantageously described by using a reaction rule-based approach, where a chemical description around the atomic vicinity of the reaction rule is encoded and thereafter used to generate novel or new-to-nature predicted reactions [40]. The rationale is that the reaction rule can then be processed by some generative algorithm in search of the substrate's chemical patterns that will allow them to pass through the reaction rule. How far such potential substrate acceptance as identified by reaction rules would have a counterpart through experimental screening has been assessed by Johnson *et al.*, showing a promising alignment between reaction rules and promiscuity mechanisms [53]. These studies suggest that the predictive accuracy of generative approaches will improve as we expand and refine the characterization of the repertoire of metabolic reactions.

(b) *Charting the extended metabolic space*

The main observation from evolvability studies is that metabolism representations based on reaction rules are generative enough to potentially provide substantial coverage, in a validated and consistent form, of what can potentially be

Downloaded from http://royalsocietypublishing.org/rstb/article-pdf/doi/10.1098/rstb.2024.0292/2821375/rstb.2024.0292.pdf
by University of Valencia (Universidad de Valencia user
on 13 January 2026

generated by extant enzymes and metabolic networks [54]. This statement should trigger an interest in deepening our knowledge about life by charting through this technique the unknown in the metabolic space. The extended metabolic space should be understood as a latent or underground metabolism [55] that might already be present in organisms, as there exists increasing evidence that much of the metabolic dark matter that is observed, for instance in metabolomics samples, could eventually be mapped into that extended metabolic space [56]. Therefore, a major effort is underway in order to map more precisely and reconcile the unknown chemodiversity in metabolomics dark matter with the increasingly accurate identification based on generative metabolism [57].

As more high-throughput screening methods become available, we might expect an increasing accuracy in the validation and prediction of the extended metabolic space assisted by generative models. By traversing metabolic networks, the method allows gap-filling and alternative pathway suggestions, with strong implications for understanding cell processes. This has lately been shown in bioprospecting and metagenomics analysis, where generative metabolism linked with enzyme sequence assignment [58] has notably been successful in finding novel enzyme activities [59]. Generative metabolism should eventually also lead to determining the enclosing contour or closure of the metabolic space, i.e. what is the available repertoire of metabolism that can be reached in nature or through evolutionary constraints [60]. A solution to the question of the size of the reachable metabolic space compared with the extant one has been sought for a long time without any clear answer [61].

Generative metabolism based on reaction rules allows us to provide an estimate *in silico* of the extent of the metabolic space, i.e. the extended metabolic space. The approach is based on assuming a primordial set of enzymes represented by their reaction rules. This primordial set can be seen when reaction rules are made general by lowering the diameter, which accounts for the size of the atomic vicinity around the reaction centre that is used in the reaction rule. Starting from this set of generative reactions and the considered set of metabolites, an iterative process is proposed where new potential metabolites are generated through the rules [62]. The process can be repeated until no more metabolites are generated, which typically occurs after a finite number of iterations.

In that way, the size of the extended metabolic space can be estimated. However, it is important to keep in mind that what is relevant from these computational experiments is the scale of expansion that might be expected in the enlarged metabolic space starting from the initial metabolic dataset, rather than the absolute number of compounds in the extended space, which might depend on the initial hypotheses and the definition of the dataset. Some recent estimates suggest that the size of the metabolic dark matter might be close to 50-fold larger than the observed metabolic space. For instance, in a recent study based on 489 reaction rules and starting with a database of 1.5 million bioactive compounds, a total number of 77 million compounds were generated, which provides a sizeable estimate of the extended metabolic space [56].

(c) *Exploring possible metabolic worlds*

Exploring the metabolic space through generative metabolism can also be connected with the evolutionary mechanisms that led to extant metabolisms. Therefore, generative metabolism not only is valuable for investigating metabolic dark matter but also can be used to shed light on how metabolism progressively evolved from a protometabolism [28]. As a matter of fact, a generative model can be built to show how, starting from a primaeval combination of basic nutrients and highly promiscuous enzymes harbouring the basic metabolic functions for sustaining life, a variety of possible metabolic worlds can be postulated, or even possible evolutionary trails leading to modern metabolic networks can be tracked through time. More precisely, it has been shown that a basic repertoire of enzymes carrying the potential activities, as encoded in the extended metabolic space, can be evolved *in silico* by generating events for gaining activity based on the propensity of the enzymes towards evolvability, which will gradually lead to networks that resemble modern metabolism [63,64]. These results might suggest that the promiscuity of enzymes somehow encodes the information of the evolutionary path leading to what would eventually become life on Earth, suggesting that enzyme sequences were cast through evolution towards the one particular form of life that nowadays accommodates the biodiversity found across the tree of life.

Given the non-ergodic nature of the potential metabolic functional space (i.e. everything chemically possible), how could biological evolution, within its temporal and spatial margins, have innovated and generated de novo functions? There has been much discussion about how genetic space relates to phenotypic space and how an initially limited number of elements of that genetic space can navigate through the functional map. Manrubia has argued for the surprising ease with which function can emerge during molecular evolution [65]. First, consider the effect of neutral mutations or how different genotypes map onto the same function, a phenomenon known as evolutionary convergence. To this functional redundancy—also represented by catalytic promiscuity and multifunctionality explored through gene duplication and divergence [21,66]—we must add the intrinsic plasticity of phenotypes, the modular nature of biological structures, the cooperative associations between genomic sequences or the combinatorics resulting from phenomena such as HGT and genetic recombination. In short, the exploration of a functional space through available and limited molecular variants is a very eloquent manifestation of the creative power of evolution as a tinkerer [67].

Therefore, a question arises about how other possible metabolic worlds could have evolved under different environmental conditions such as those found in outer worlds. This question might be approached by recreating those conditions as far as the ancestral enzymatic reactions that would have emerged within these conditions are known. Towards that direction, we can start by reproducing harsh conditions for life on Earth that might be present elsewhere and test how far life could have

Downloaded from http://royalsocietypublishing.org/rstb/article-pdf/doi/10.1098/rstb.2024.0292/2821375/rstb.2024.0292.pdf
by University of Valencia (Universidad de Valencia user
on 13 January 2026

emerged in those conditions if evolutionary propensities were similar to those found at home. The main question at stake is whether our findings on how life evolved to extant metabolic networks can be considered as a general rule sustaining life, even in alien environments.

Box 2: The path towards synthetic metabolism (transmetabolism)

As we gain more knowledge and predictive power about the latent capabilities of enzymes, bio-based production of alternative products to those coming from fossil resources becomes more viable in the context of the circular bioeconomy. To that end, synthetic metabolism, or transmetabolism [68], allows metabolism to be altered by engineering novel pathways in chassis organisms. However, a major bottleneck in producing valuable products is the low performance of such cell factories because of the complexity associated with the optimization of the processes. A major feat in order to alleviate such a shortcoming has been the development of directed evolution of enzymes so that they can reach performances that are beyond those found in extant natural metabolic networks, which often are not highly optimized [69] so that their catalytic efficiency does not draw too many resources from other essential processes in cells. Through biotechnology, we can nowadays engineer highly streamlined strains with altered metabolism and bespoke-designed enzymes displaying high performance up to the maximum theoretical catalytic efficiency [70]. Ancestral protein reconstructions can resurrect metabolic activities that existed in ancient organisms, leading to novel biotechnological applications [71]. Synthetic metabolism is currently a major technology that will shape our society and eventually enable Mars exploration and beyond. Adding to that vision, generative metabolism provides the tools for the directed evolution of enzymes, as the method can pinpoint which capabilities might be present in parent enzyme sequences, which can be afterwards optimized by directed evolution [72]. This can be mainly achieved through high-throughput screening, especially at the single-cell level, by the development of biosensor-based screening [73,74].

Understanding how organisms evolved by looking at the evolvability of their metabolic networks and their potential enzymatic capabilities is allowing us nowadays to unveil a large repertoire of natural and new-to-nature products that can be biologically produced. As such, biomanufacturing capabilities are boosted by the establishment of biofoundry facilities [75], allowing automation and high-throughput acceleration of the Design–Build–Test–Learn (DBTL) cycle [76] and eventually facilitating a wealthy metabolic new world to be made possible.

Interestingly, the elementary representations encoded by the rules are shared by several reactions that will progressively cluster together as we lower the diameter [77]. In that way, it would be possible to map the COGs present in LUCA to a set of clusters of orthologous reactions (CORs) that constitute a minimal generative set providing full coverage of the COGs in LUCA. In each of these mappings, which represents one possible explanation for the origin of extant networks, every LUCA COG in the sequence space would be associated with a set of LUCA's CORs in the chemical space, unlocking the connection between both domains. The reaction rules underlying the CORs representation are the ones that in turn might be related to enzyme evolution in such a way that continuity between backward and forward evolution is preserved in the mapping.

The modelling of enzyme evolution through CORs can be compared with various hypotheses on metabolic evolution (see [19] and references therein). For instance, the backward or retrograde model posits that new reactions were added as successive metabolic precursors became increasingly scarce in the primordial environment. In this scenario, older enzymes are expected to operate at the later stages of modern metabolic pathways. Conversely, the forward model suggests that new reactions were incorporated as precursors were already available, implying that older enzymes would be located upstream in these pathways. Both models, originally proposed for the origin of biosynthetic pathways in the context of the heterotrophic hypothesis, assume that enzymes acquired promiscuity over the course of evolution as new reactions emerged, explaining the observed trend that older enzymes tend to be more promiscuous, having had sufficient time to evolve additional functions. The patchwork model proposes that ancient enzymes were highly promiscuous or exhibited low specificity, and gradually specialized through gene duplication and divergence events over evolutionary time. According to this hypothesis, which would be valid for both anabolic and catabolic pathways, enzymes with specific functions were co-opted at various points across the metabolic network. The semi-enzymatic model refines this concept, suggesting that early metabolism included not only enzymatic reactions but also non-enzymatic reactions, either occurring spontaneously or catalysed by simpler molecules (such as free coenzymes, organometallic compounds or metal ions) or by means of intramolecular catalysis [26]. According to these theories, enzyme promiscuity would have been a common feature of ancient enzymes and should therefore be widespread across the tree of life.

(c) Exploring linkages connecting the prebiotic and the protometabolic worlds

To complete this modelling framework for the history of life, the next step is to move beyond protometabolism by establishing its connection to the metabolism present in LUCA. Several models have been proposed to explain how life emerged from prebiotic chemistry [17]. A recent study primarily focused on a set of reactions involving volatile organic compounds likely present on early Earth and potentially metabolized by primitive cells before LUCA [78]. Our proposal points to more foundational aspects of protometabolism. Current research adopts a systemic approach [8,79], in dialogue with the geo- and cosmo-chemical conditions of early Earth [11], including the recognition that some essential protometabolic components probably were lost and are absent from

Downloaded from http://royalsocietypublishing.org/rstb/article-pdf/doi/10.1098/rstb.2024.0292/2821375/rstb.2024.0292.pdf
by University of Valencia (Universidad de Valencia user
on 13 January 2026

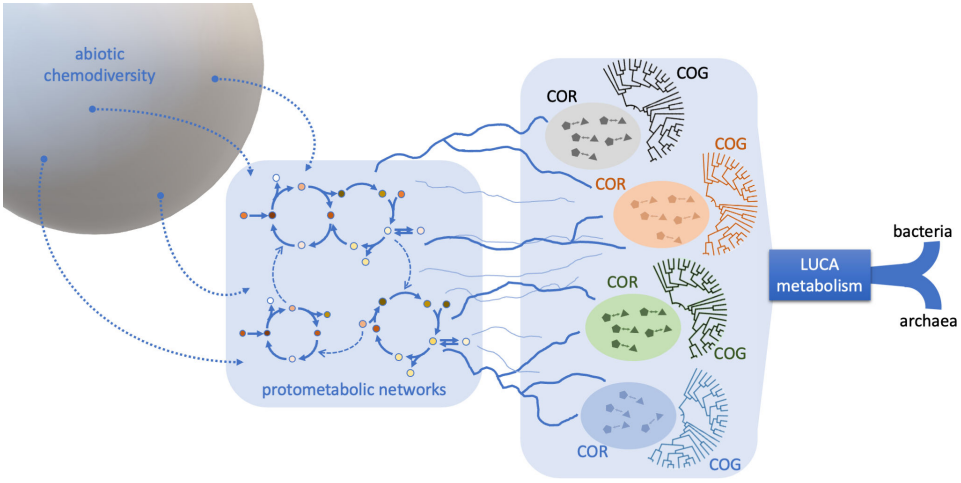

Figure 1. Connecting the last universal common ancestor's (LUCA's) metabolism with protometabolic networks. Top-down approaches infer models for LUCA metabolism consisting of primaeval sets of clusters of orthologous genes (COGs). Such clusters can be associated with clusters of orthologous rections (CORs) based on reaction rule representations. Bottom-up approaches, in turn, propose protometabolic reactions that can be linked to the prebiotic early chemodiversity. The transition establishing the linkage between protometabolic networks and CORs can be investigated through generative metabolism by establishing routes that connect CORs with ancient prebiotic reactions.

modern metabolism [80]. A key aspect is the role of autocatalytic cycles in creating a self-sustaining basis for life's building blocks [66], the participation of KSTA compounds [25,26] and their plausible prebiotic emergence as suggested by the glyoxylose scenario [22]. By exploring prebiotic systems chemistry and the candidate (activated) chemical bonds participating in the reactions of the proposed autocatalytic cycles, we can investigate the possible pathways involved in the transition to life by linking these reaction sets with elementary COR sets. These sets are based on the generative reaction rules from LUCA and have the potential to unfold modern metabolism.

Interestingly, reaction rules within each COR can be matched with potential prebiotic reactions to test whether the COR, along with its associated COG in LUCA (and eventually its enzyme catalyst), could have emerged from an ancestral reaction (figure 1). In this way, prebiotic reactions can be connected to COGs in LUCA. Once a minimal set of matches is found for all COGs, a hypothetical linkage is established, explaining the transition from abiotic chemistry to protometabolism and, ultimately, to life. This process can be repeated through exhaustive searches until all possible solutions have been enumerated. At that stage, under the given assumptions and evolutionary models, we would expect the set to contain at least one linkage that most plausibly describes the emergence of protometabolism from the abiotic world. Selecting the most plausible linkage will require defining likelihood criteria, potentially based on thermodynamic considerations, the environmental conditions of early Earth and the availability of reactants. Determining the constraints that shaped the critical transition from geochemistry to biochemistry requires examining both the geological history of the planet and experimental findings from research on the chemical origins of life.

4. Conclusion

As suggested by systems chemistry experiments, prebiotic conditions likely resulted in a relatively limited set of chemical species, challenging the previous assumption of a highly diverse early chemistry. This sparse protometabolic network forms a bottom-up foundation for later metabolic evolution as a scaffold for the emergence of complex catalysts. Similarly, according to the top-down approach, studying enzyme promiscuity or the ability of enzymes to catalyse multiple reactions is crucial to understanding early metabolic evolution. Here, we argue that generative metabolism models, based on reaction rules, offer a promising avenue for exploring enzyme evolution in the extended metabolic space, which includes both known and potential metabolic reactions. By understanding the full range of possible reactions, we can gain a deeper understanding of life's origins and its latent potential. Combining the bottom-up and the top-down approaches could bridge the gap between the prebiotic world and LUCA by tracing enzyme evolution back through the tree of life towards LUCA, and mapping ancestral genes in LUCA to their corresponding reaction clusters (CORs), eventually supporting the congruence scenario between prebiotic chemistry and metabolism proposed by de Duve [20]. Moreover, these studies allow examination of potential links between prebiotic reactions and CORs in LUCA to understand the transition from abiotic chemistry to life. Furthermore, the principles of generative metabolism can also be extended towards practical applications in synthetic biology, such as in the development of bio-based products and the engineering of novel metabolic pathways. By combining insights from systems chemistry, computational methods and evolutionary biology, we argue that a clearer understanding of life's origins and its potential is within reach.

Data accessibility. This article has no additional data.

Declaration of AI use. We have used AI-assisted technologies in creating this article.

Authors' contributions. P.C.: conceptualization, funding acquisition, investigation, methodology, visualization, writing—original draft, writing—review and editing; J.P.: conceptualization, funding acquisition, investigation, writing—original draft, writing—review and editing.

Both authors gave final approval for publication and agreed to be held accountable for the work performed herein.

Downloaded from http://royalsocietypublishing.org/rstb/article-pdf/doi/10.1098/rstb.2024.0292/2821375/rstb.2024.0292.pdf
by University of Valencia (Universidad de Valencia user
on 13 January 2026

Conflict of interest declaration. We declare we have no competing interests.

Funding. P.C. acknowledges financial support from Generalitat Valenciana through grant CIAICO/2021/159 (SmartBioFab) and MCIN/AEI/10.13039/501100011033 and European Union NextGenerationEU/ PRTR funding through grant TED2021-131049B-I00 (BioEcoDBTL). J.P. has not received specific funding for this study but has received support from the academic institutions to which he belongs.

Acknowledgements. J.P. acknowledges stimulating discussion with numerous colleagues, including the SFI working group on the origin of life organized by Ricard Solé and Chris Kempes.

References

1. Rees M. 2000 *Just six numbers. The deep forces that shape the Universe.* New York, NY: Basic Books.
2. Kauffman S. 1996 *At home in the Universe: the search for the laws of self-organization and complexity*. New York, NY: Oxford University Press.
3. Alberch P. 1989 The logic of monsters: evidence for internal constraint in development and evolution. *Geobios* **22**, 21–57. (doi:10.1016/S0016-6995(89)80006-3)
4. 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. (doi:10.1073/pnas.0912157107)
5. Peretó J. 2018 Prebiotic chemistry that led to life. In *Handbook of astrobiology* (ed. V Kolb), pp. 219–233. Boca Raton, FL: CRC Press. (doi:10.1201/b22230-18)
6. 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. (doi:10.1038/s41467-022-29612-x)
7. Krishnamurthy R, Goldman AD, Liberles DA, Rogers KL, Tor Y. 2022 Nucleobases in meteorites to nucleobases in RNA and DNA? *J. Mol. Evol*. **90**, 328–331. (doi:10.1007/s00239-022-10069-x)
8. Powner MW, Sutherland JD. 2011 Prebiotic chemistry: a new *modus operandi*. *Phil. Trans. R Soc. B* **366**, 2870–2877. (doi:10.1098/rstb.2011.0134)
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. (doi:10.1038/nchem.2202)
10. Islam S, Powner MW. 2017 Prebiotic systems chemistry: complexity overcoming clutter. *Chem* **2**, 470–501. (doi:10.1016/j.chempr.2017.03.001)
11. Sasselov DD, Grotzinger JP, Sutherland JD. 2020 The origin of life as a planetary phenomenon. *Sci. Adv*. **6**, eaax3419. (doi:10.1126/sciadv.aax3419)
12. Eschenmoser A. 2007 The search for the chemistry of life's origin. *Tetrahedron* **63**, 12821–12844. (doi:10.1016/j.tet.2007.10.012)
13. Springsteen G, Yerabolu JR, Nelson J, Rhea CJ, Krishnamurthy R. 2018 Linked cycles of oxidative decarboxylation of glyoxylate as protometabolic analogs of the citric acid cycle. *Nat. Commun*. **9**, 91. (doi:10.1038/s41467-017-02591-0)
14. Pulletikurti S, Yadav M, Springsteen G, Krishnamurthy R. 2022 Prebiotic synthesis of α-amino acids and orotate from α-ketoacids potentiates transition to extant metabolic pathways. *Nat. Chem*. **14**, 1142–1150. (doi:10.1038/s41557-022-00999-w)
15. Muchowska KB, Varma SJ, Moran J. 2020 Nonenzymatic metabolic reactions and life's origins. *Chem. Rev*. **120**, 7708–7744. (doi:10.1021/acs.chemrev.0c00191)
16. Stubbs RT, Yadav M, Krishnamurthy R, Springsteen G. 2020 A plausible metal-free ancestral analogue of the Krebs cycle composed entirely of α-ketoacids. *Nat. Chem*. **12**, 1016–1022. (doi:10.1038/s41557-020-00560-7)
17. Krishnamurthy R. 2018 Life's biological chemistry: a destiny or destination starting from prebiotic chemistry? *Chemistry* **24**, 16708–16715. (doi:10.1002/chem.201801847)
18. Dherbassy Q, Mayer RJ, Moran J. 2024 Coenzymes in a pre-enzymatic metabolism. *Sci. Adv*. **10**, eadr5357. (doi:10.1126/sciadv.adr5357)
19. Lazcano A, Miller SL. 1999 On the origin of metabolic pathways. *J. Mol. Evol.* **49**, 424–431. (doi:10.1007/PL00006565)
20. de Duve C. 2005 The onset of selection. *Nature* **433**, 581–582. (doi:10.1038/433581a)
21. Peretó J. 2024 A praise of imperfection. In *The first steps of life* (ed. E Di Mauro), pp. 79–98. Hoboken, NJ: John Wiley & Sons. (doi:10.1002/9781394264155.ch4)
22. Krishnamurthy R, Liotta CL. 2023 The potential of glyoxylate as a prebiotic source molecule and a reactant in protometabolic pathways—the glyoxylose reaction. *Chem* **9**, 784–797. (doi:10.1016/j.chempr.2023.03.007)
23. Xavier JC, Kauffman S. 2022 Small-molecule autocatalytic networks are universal metabolic fossils. *Phil. Trans. R. Soc. A* **380**, 20210244. (doi:10.1098/rsta.2021.0244)
24. Pross A, Pascal R. 2023 On the emergence of autonomous chemical systems through dissipation kinetics. *Life* **13**, 2171. (doi:10.3390/life13112171)
25. Pascal R. 2024 Evolutionary abilities of minimalistic physicochemical models of life processes. *Chemistry* **30**, e202401780. (doi:10.1002/chem.202401780)
26. Maguire OR. 2024 The many-chemicals problem of systems chemistry. *ChemSystemsChem* **6**, e202400027. (doi:10.1002/syst.202400027)
27. Negrón-Mendoza A, Hernández-Morales R, Lazcano A. 2025 Can the origin of biosynthetic routes be explained by a Frankenstein's monster-like spontaneous assembly of prebiotic reactants? *Phil. Trans. R. Soc. B* **380**, 20240289. (doi:10.1098/rstb.2024.0289)
28. Nogal N, Sanz-Sánchez M, Vela-Gallego S, Ruiz-Mirazo K, de la Escosura A. 2023 The protometabolic nature of prebiotic chemistry. *Chem. Soc. Rev*. **52**, 7359–7388. (doi:10.1039/d3cs00594a)
29. Novikov Y, Copley SD. 2013 Reactivity landscape of pyruvate under simulated hydrothermal vent conditions. *Proc Natl Acad. Sci. USA* **110**, 13283–13288. (doi:10.1073/pnas.1304923110)
30. Smith E, Morowitz HJ. 2010 The autotrophic origins paradigm and small-molecule organocatalysis. *Orig. Life Evol. Biosph.* **40**, 397–402. (doi:10.1007/s11084-010-9213-2)
31. Oparin AI. 1924 *Proiskhozhedenie zhizni.* Moscow, Russia: Mosckovskii Rabochii.
32. de Duve C. 2010 *Singularities*. Cambridge, UK: Cambridge University Press. (doi:10.1017/CBO9780511614736)
33. Lazcano A, Fox G, Oró J. 1992 Life before DNA: the origin and evolution of early Archean cells. In *The evolution of metabolic function* (ed. RP Mortlock), pp. 237–295. Boca Raton, FL: CRC Press.
34. Rivas M, Fox GE. 2024 On the nature of the last common ancestor: a story from its translation machinery. *J. Mol. Evol*. **92**, 593–604. (doi:10.1007/s00239-024-10199-4)
35. Crapitto AJ, Campbell A, Harris A, Goldman AD. 2022 A consensus view of the proteome of the last universal common ancestor. *Ecol. Evol*. **12**, e8930. (doi:10.1002/ece3.8930)
36. Goldman AD, Becerra A. 2024 A new view of the last universal common ancestor. *J. Mol. Evol*. **92**, 659–661. (doi:10.1007/s00239-024-10193-w)
37. Moody ERR *et al*. 2024 The nature of the last universal common ancestor and its impact on the early Earth system. *Nat. Ecol. Evol*. **8**, 1654–1666. (doi:10.1038/s41559-024-02461-1)
38. Khersonsky O, Roodveldt C, Tawfik DS. 2006 Enzyme promiscuity: evolutionary and mechanistic aspects. *Curr. Opin. Chem. Biol*. **10**, 498–508. (doi:10.1016/j.cbpa.2006.08.011)
39. Carbonell P, Lecointre G, Faulon JL. 2011 Origins of specificity and promiscuity in metabolic networks. *J. Biol. Chem*. **286**, 43994–44004. (doi:10.1074/jbc.m111.274050)
40. Duigou T, Lac M, Carbonell P, Faulon JL. 2019 RetroRules: a database of reaction rules for engineering biology. *Nucl. Acids Res.* **47**, D1229–D1235. (doi:10.1093/nar/gky940)
41. Rodrigues JFM, Wagner A. 2009 Evolutionary plasticity and innovations in complex metabolic reaction networks. *PLoS Comp. Biol.* **5**, e1000613. (doi:10.1371/journal.pcbi.1000613)

Downloaded from http://royalsocietypublishing.org/rstb/article-pdf/doi/10.1098/rstb.2024.0292/2821375/rstb.2024.0292.pdf
by University of Valencia (Universidad de Valencia user
on 13 January 2026

42. Aharoni A, Gaidukov L, Khersonsky O, Gould SMQ, Roodveldt C, Tawfik DS. 2005 The 'evolvability' of promiscuous protein functions. *Nat. Genet*. **37**, 73–76. (doi:10.1038/ng1482)
43. Bloom JD, Labthavikul ST, Otey CR, Arnold FH. 2006 Protein stability promotes evolvability. *Proc. Natl Acad. Sci. USA* **103**, 5869–5874. (doi:10.1073/pnas.0510098103)
44. Arnold FH. 2018 Directed evolution: bringing new chemistry to life. *Angew. Chem.* **57**, 4143–4148. (doi:10.1002/anie.201708408)
45. Buda K, Miton CM, Fan XC, Tokuriki N. 2023 Molecular determinants of protein evolvability. *Trends Biochem. Sci*. **48**, 751–760. (doi:10.1016/j.tibs.2023.05.009)
46. Jumper J *et al*. 2021 Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589. (doi:10.1038/s41586-021-03819-2)
47. Ingraham JB *et al*. 2023 Illuminating protein space with a programmable generative model. *Nature* **623**, 1070–1078. (doi:10.1038/s41586-023-06728-8)
48. Madani A *et al*. 2023 Large language models generate functional protein sequences across diverse families. *Nat. Biotechnol*. **41**, 1099–1106. (doi:10.1038/s41587-022-01618-2)
49. Malbranke C, Bikard D, Cocco S, Monasson R, Tubiana J. 2023 Machine learning for evolutionary-based and physics-inspired protein design: current and future synergies. *Curr. Opin. Struct. Biol*. **80**, 102571. (doi:10.1016/j.sbi.2023.102571)
50. Ferruz N, Schmidt S, Höcker B. 2022 ProtGPT2 is a deep unsupervised language model for protein design. *Nat. Commun*. **13**, 4348. (doi:10.1038/s41467-022-32007-7)
51. Barghout RA, Xu Z, Betala S, Mahadevan R. 2023 Advances in generative modeling methods and datasets to design novel enzymes for renewable chemicals and fuels. *Curr. Opin. Biotechnol.* **84**, 103007. (doi:10.1016/j.copbio.2023.103007)
52. Corey EJ. 1991 The logic of chemical synthesis: multistep synthesis of complex carbogenic molecules (Nobel Lecture). *Angew. Chem. Int. Edn Engl*. **30**, 455–465. (doi:10.1002/anie.199104553)
53. Johnson SR, Fu X, Viknander S, Goldin C, Monaco S, Zelezniak A, Yang KK. 2025 Computational scoring and experimental evaluation of enzymes generated by neural networks. *Nat. Biotechnol*. **43**, 396–405. (doi:10.1038/s41587-024-02214-2)
54. Carbonell P, Delépine B, Faulon JL. 2018 Extended metabolic space modeling. In *Methods in molecular biology synthetic metabolic pathways* (eds MK Jensen, JD Keasling), pp. 83–96. New York, NY: Springer. (doi:10.1007/978-1-4939-7295-1_6)
55. D'Ari R, Casadesús J. 1998 Underground metabolism. *Bioessays* **20**, 181–186. (doi:10.1002/(sici)1521-1878(199802)20:23.3.co;2-6)
56. Mohammadi Peyhani H, Hafner J, Sveshnikova A, Viterbo V, Hatzimanikatis V. 2022 Expanding biochemical knowledge and illuminating metabolic dark matter with ATLASx. *Nat. Commun.* **13**, 1560. (doi:10.1038/s41467-022-29238-z)
57. Coler EA, Chen W, Melnik AV, Morton JT, Aksenov AA. 2024 Metabolomics in the era of artificial intelligence. *Microbiota Host* **2**, e230017. (doi:10.1530/mah-23-0017)
58. Stoney RA, Hanko EKR, Carbonell P, Breitling R. 2023 SelenzymeRF: updated enzyme suggestion software for unbalanced biochemical reactions. *Comp. Struct. Biotechnol. J.* **21**, 5868–5876. (doi:10.1016/j.csbj.2023.11.039)
59. Presnell KV, Alper HS. 2019 Systems metabolic engineering meets machine learning: a new era for data-driven metabolic engineering. *Biotech. J.* **14**, 1800416. (doi:10.1002/biot.201800416)
60. Letelier JC, Cárdenas ML, Cornish-Bowden A. 2011 From *L'Homme Machine* to metabolic closure: steps towards understanding life. *J. Theor. Biol*. **286**, 100–113. (doi:10.1016/j.jtbi.2011.06.033)
61. Ravikrishnan A, Nasre M, Raman K. 2018 Enumerating all possible biosynthetic pathways in metabolic networks. *Scient. Rep*. **8**, 9932. (doi:10.1038/s41598-018-28007-7)
62. Faulon JL, Carbonell P. 2010 Reaction network generation. In *Handbook of chemoinformatics algorithms* (eds JL Faulon, A Bender), pp. 317–341. Boca Raton, FL: Chapman and Hall/CRC. (doi:10.1201/9781420082999-c11)
63. Caetano-Anollés G, Yafremava LS, Gee H, Caetano-Anollés D, Kim HS, Mittenthal JE. 2009 The origin and evolution of modern metabolism. *Int. J. Biochem. Cell Biol*. **41**, 285–297. (doi:10.1016/j.biocel.2008.08.022)
64. Basler G, Grimbs S, Ebenhöh O, Selbig J, Nikoloski Z. 2012 Evolutionary significance of metabolic network properties. *J. R. Soc. Interface* **9**, 2011. (doi:10.1098/rsif.2011.0652)
65. Manrubia S. 2022 The simple emergence of complex molecular function. *Phil. Trans. R. Soc. A* **380**, 20200422. (doi:10.1098/rsta.2020.0422)
66. Peretó J. 2012 Out of fuzzy chemistry: from prebiotic chemistry to metabolic networks. *Chem. Soc. Rev.* **41**, 5394–5403. (doi:10.1039/c2cs35054h)
67. Wagner A. 2019 *Life finds a way: what evolution teaches us about creativity*. New York, NY: Basic Books.
68. Peretó J. 2020 Transmetabolism: the non-conformist approach to biotechnology. *Microb. Biotechnol.* **14**, 1751–7915. (doi:10.1111/1751-7915.13691)
69. Bar-Even A, Milo R, Noor E, Tawfik DS. 2015 The moderately efficient enzyme: futile encounters and enzyme floppiness. *Biochemistry* **54**, 4969–4977. (doi:10.1021/acs.biochem.5b00621)
70. Goldsmith M, Tawfik DS. 2017 Enzyme engineering: reaching the maximal catalytic efficiency peak. *Curr. Opin. Struct. Biol*. **47**, 140–150. (doi:10.1016/j.sbi.2017.09.002)
71. Kaçar B. 2024 Reconstructing early microbial life. *Annu. Rev. Microbiol.* **78**, 23. (doi:10.1146/annurev-micro-041522-103400)
72. Wittmann BJ, Johnston KE, Wu Z, Arnold FH. 2021 Advances in machine learning for directed evolution. *Curr. Opin. Struct. Biol*. **69**, 11–18. (doi:10.1016/j.sbi.2021.01.008)
73. Hwang HG, Ye DY, Jung GY. 2023 Biosensor-guided discovery and engineering of metabolic enzymes. *Biotech. Adv.* **69**, 108251. (doi:10.1016/j.biotechadv.2023.108251)
74. Tellechea-Luzardo J, Stiebritz MT, Carbonell P. 2023 Transcription factor-based biosensors for screening and dynamic regulation. *Front. Bioeng. Biotechnol*. **11**, 1118702. (doi:10.3389/fbioe.2023.1118702)
75. Tellechea-Luzardo J, Otero-Muras I, Goñi-Moreno A, Carbonell P. 2022 Fast biofoundries: coping with the challenges of biomanufacturing. *Trends Biotechnol.* **40**, 831–842. (doi:10.1016/j.tibtech.2021.12.006)
76. Carbonell P. 2018 An automated Design-Build-Test-Learn pipeline for enhanced microbial production of fine chemicals. *Commun. Biol.* **1**, 66. (doi:10.1038/s42003-018-0076-9)
77. Otero-Muras I, Carbonell P. 2021 Automated engineering of synthetic metabolic pathways for efficient biomanufacturing. *Metab. Eng*. **63**, 61–80. (doi:10.1016/j.ymben.2020.11.012)
78. Ledford SM, Meredith LK. 2024 Volatile organic compound metabolism on early Earth. *J. Mol. Evol.* **92**, 605–617. (doi:10.1007/s00239-024-10184-x)
79. 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. (doi:10.1021/cr2004844)
80. Krishnamurthy R. 2020 Systems chemistry in the chemical origins of life: the 18th camel paradigm. *J. Syst. Chem*. **8**, 40–62.

Downloaded from http://royalsocietypublishing.org/rstb/article-pdf/doi/10.1098/rstb.2024.0292/2821375/rstb.2024.0292.pdf
by University of Valencia (Universidad de Valencia user
on 13 January 2026