

The evolution of the metabolic network over long timelines

Markus Ralser^{1,2}, Sreejith J. Varma¹ and Richard A. Notebaart³

Abstract

Metabolism is executed by an efficient, interconnected and ancient biochemical system, the metabolic network. Its evolutionary origins are, however, barely understood. We here discuss that because of niche adaptation, the evolutionary selection acting on the metabolic network structure distinguishes modern species and early life forms. Yet, its basic structure remained conserved over more than three billion years of diverging evolution. We speculate that this situation attributes key roles in metabolic network evolution to (i) the reaction properties of central metabolites, (ii) simple catalysts (e.g. metal ions, amino acids) whose importance remained unchanged during evolution, and (iii) the interconnectivity of the network that limits its expansion. The conservation of network structure hence implies that early life forms already used similar metabolic reaction topologies as modern species.

Addresses

¹ Department of Biochemistry, Charité Universitätsmedizin, Berlin, Germany

² The Molecular Biology of Metabolism Laboratory, The Francis Crick Institute, London, UK

³ Food Microbiology, Wageningen University & Research, Wageningen, the Netherlands

Corresponding author: Notebaart, Richard A. (richard.notebaart@wur.nl)

Current Opinion in Systems Biology 2021, 28:100402

This review comes from a themed issue on **Metabolic Networks (2022)**

Edited by **Sarah-Maria Fendt & Markus Ralser**

For a complete overview see the [Issue](#) and the [Editorial](#)

Available online 22 October 2021

<https://doi.org/10.1016/j.coisb.2021.100402>

2452-3100/© 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Keywords

Evolutionary forces, Underground metabolism, Early and modern evolution, Metabolic network structure, Geochemistry, Systems biology, Adaptive evolution, Stepping-stone evolution, Niche adaptation, Genetics, Biochemistry.

Introduction

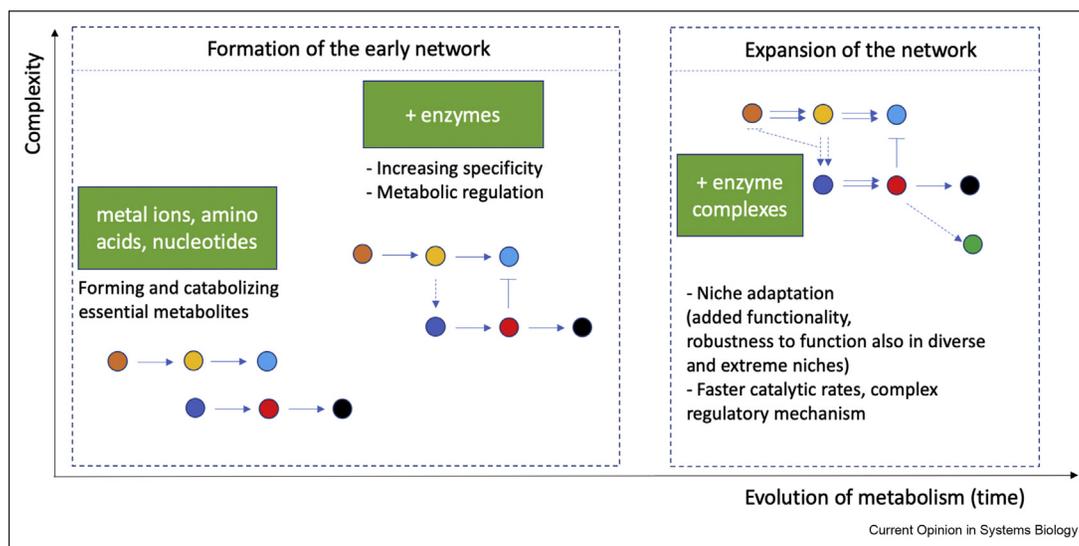
Despite its huge diversity, all forms of life are essentially built from a structurally limited set of metabolites. Amino acids, nucleotides, fatty acids, and many co-enzymes are used universally. There is more variability in the biomolecules that build cell membranes and cell walls, but also these are made from a conserved set of precursor metabolites that originate in what is known as central metabolism, i.e. glycolysis (the Embden-Meyerhof-Parnas-Pathway or related glucose oxidation pathways), the pentose phosphate pathway, and the tricarboxylic acid cycle. Importantly, not only are the metabolites conserved between the kingdoms of life but also most of the metabolic pathways that synthesize them. Metabolic pathways often possess a striking similarity in their topological structure, that is, their biochemical reaction sequences. Interestingly, there are examples that show that the reaction sequences and metabolites can be older than the enzymes that catalyze them. For instance, Eubacteria and Archaea use a glycolytic pathway, that is, made of similar intermediates but without sharing sequence conservation in their metabolic enzymes [1]. Either the two kingdoms have, hence, evolved the same pathway independently or one or both lineages have over time replaced older enzymes with modern ones. In either case, the long evolutionary timeline of at least three billion years of divergent evolutionary pressure did not fundamentally alter many of the core metabolic reaction topologies. Here, we debate key evolutionary and biochemical constraints acting on cellular metabolism that differ, or remain, constant over these extreme evolutionary periods (Figure 1). We conclude that early lifeforms must have been using metabolism-like biochemical reaction sequences from very early on.

The early evolution of metabolism

We do not know in which environment the early forms of the metabolic network emerged, but it is a reasonable assumption that the formation of an early metabolic network was a key step in the origins of life [2]. The latest moment when cells needed a comprehensive metabolic network that can produce the life-essential metabolites, including amino acids, nucleotides, lipids, and carbohydrates, was the point when life started to colonize diverse ecological niches. We can hence assume

“This review is dedicated to Dan Tawfik, 1955–2021”.

Figure 1



The evolution of the metabolic network structure as expressed by complexity and time. The formation of the early network is driven by simple chemical constraints involving the reaction properties of the early intermediates, and metal ions, amino acids and later more complex constraints involving the first enzymes, acting as catalysts. The expansion of the network is further driven by increasing ecological competition stimulating the evolution of enzymes, enzyme complexes and regulation of metabolic fluxes. Increased competition and network expansion distinguish the evolution of metabolism in early and modern-day metabolic networks. Metabolites (chemical compounds) are represented by ●, metabolic reactions (metabolite turnover) by → and regulation by (inhibition) ⊥.

that the metabolic network is at least 3.7 billion years old, as old as the earliest fossil records. One school of thought suggests that the metabolic network could have originated in specialized extreme environments, such as hydrothermal vents, as they offer a constant supply of reducing electrons and matter to feed the emerging primitive forms of life [3]. The other possibility is that life emerged in ambient environments, such as what is amenable to mesophiles, which by far has given rise to the most sophisticated lifeforms on earth, including us humans. It speaks for the latter that cells require a high degree of sophisticated and specialized enzyme structures to grow and survive in extreme environments. Moreover, the metabolic network contains several central metabolites, like 3-phosphoglycerate, that decompose at a high temperature within seconds [4]. However, no matter what the origins scenario is, the earliest life forms had to adapt their metabolism to meet the environmental constraints to survive and be able to have colonized the different niches.

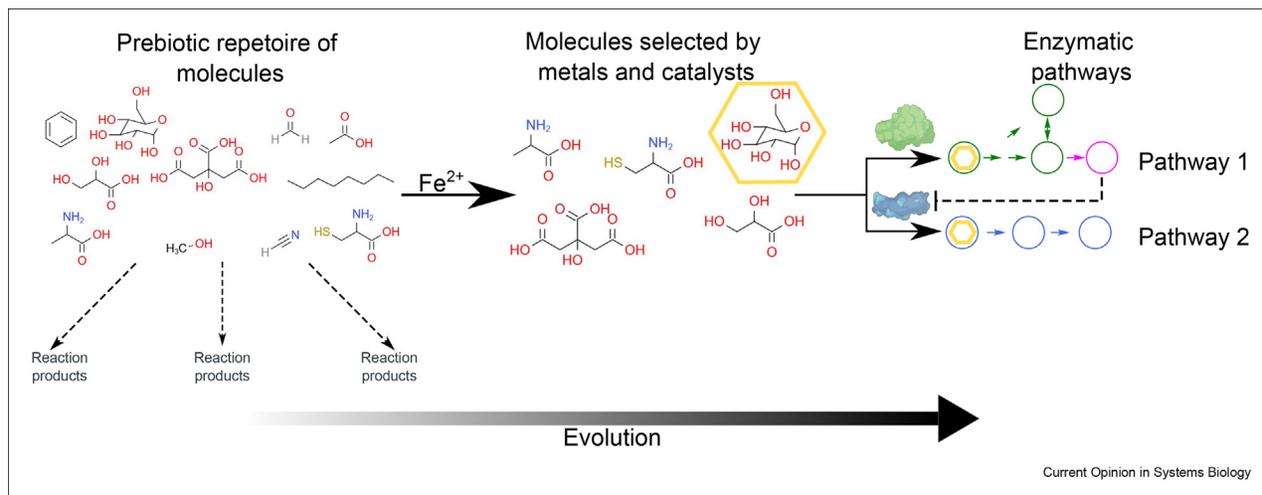
The reactions that form the metabolic network have been selected out of the space of the thermodynamically plausible reactions (Figure 2). However, thermodynamics alone does not suffice to explain the metabolic network structure. The situation that a reaction is thermodynamically possible does not mean the reaction is spontaneous. Moreover, metabolism exploits only a small fraction of the thermodynamically possible space. For example, many of the Maillard reactions between

amino acids and glycosides that could be catalyzed to occur under ordinary conditions, are not part of the metabolic network [5]. Indeed, evolution does not change the thermodynamic properties of the metabolites it converts, but evolution can select an enzyme that lowers the activation energy of a specific interconversion, and in this way, defines the structure of the metabolic network (Figure 2).

Up to 40% of modern enzymes bind a metal ion for their catalytic function [6,7]. As metal ions were prebiotically available, it is hence a reasonable assumption that also early forms of metabolism did exploit metal ions as catalysts or cofactors. Geochemistry has elaborated environmental constraints that apply to both the extreme and mundane prebiotic niches. This would have been the earliest and most significant among the constraints that have shaped early metabolic structure, that is, chemicals available in the environment, as these constrained the evolution of the first catalysts and their substrates.

Environmental conditions on primitive earth comprised principally CO₂, N₂, H₂O, CO [8]. This dysoxic environment acted as an inert atmosphere, favoring high concentrations of the water-soluble Fe(II)- the by far most abundant transition metal in Archean Sediment [9,10]. Iron in various forms has been reported to drive many reactions abiotically reminiscent of metabolic reactions, including those found within glycolysis, the

Figure 2



Constraints acting on the evolution of metabolic networks. The space of thermodynamically plausible reactions is large, but only a small fraction is exploited in the expansion of the metabolic network. Besides selection on available thermodynamically plausible reactions, metabolites (indicated with circles) produced in metabolic reactions could also inhibit other reactions/pathways to occur by inhibiting enzyme catalysis (indicated by \perp), hence become another constraint in the evolution of the metabolic network.

Krebs cycle, and the pentose phosphate pathway [4,11,12]. Some of these reactions are pivotal to the topological organization of the metabolic network, indicating that Fe(II) could have been one of metabolism's earliest tools in its catalyst inventory. Indeed, the omnipresence of ferrous iron would have made it difficult for early life forms to escape iron-driven reactions. Even in the oxygenated environment of today, chelators are required to deprive environmental water from life-essential iron concentrations. Hence it is plausible that early forms of metabolism were exposed to the reaction spectrum this metal catalyst provides.

It is only reasonable to assume that many environments of the primordial earth might have contained also a spectrum of molecules that are not taking part in essential reactions of the modern metabolic network. HCN or formaldehyde, for instance, show interesting reaction properties and are, hence, discussed in the context of the origins of life [8]. Both HCN and formaldehyde molecules, however, indiscriminately react with several metabolic intermediates. They hence disrupt the metabolic network structure, and are toxic to metabolism also in the most sophisticated cells. Most likely, these molecules were hence toxic also to simpler forms of cells that had far fewer defense mechanisms to their disposal compared to modern cells [13,14]. We conclude that the fact that a metabolite was prebiotically available does not necessarily mean it was involved in metabolic evolution. Like in modern metabolism, a key constraint is that a molecule's reaction properties need to be compatible with the basic functional principles of metabolism, which centre around its topological organisation.

Early metabolic reactions could have either occurred via redox and rearrangement reactions of different inorganic species with a carbon source or driven by processes, such as mineral serpentinization or photolysis of water, generating the necessary potentials to reduce carbon [15–18]. Because the still recent discovery of iron-driven non-enzymatic reaction networks resemble metabolic network topologies [4], there is increasing data support for a hypothesis in which the earliest life forms have contained chemical reactions that resemble many extant biological pathways. Such would indicate that the basic structure of the metabolic network is fundamentally driven by chemical rules that were already in place before life formed, that is constraint by the availability of the aforementioned metal catalysts and the reaction properties of the key metabolites [19–21]. Importantly, a theory that roots the origin of the metabolic network structure in environmental non-enzymatic reaction sequences, does not exclude the possibility that at least some metabolites formed in extreme environments, that can give rise to several biomolecules [22–24], got incorporated into the developing metabolic network as well. This scenario entails that these molecules reached sufficient concentration environments so mundane that the metabolic network that contains many low abundant and reactive intermediates could function and evolve.

Genetics, which selects enzymes eventually, could only emerge and thrive once a diverse set of biomolecules, sufficient to initiate a Darwinian process, could be formed. The set of metabolites essential for Darwinian selection include amino acids for protein biosynthesis,

nucleotides including ATP for forming nucleic acids and the cyclic energy transfer that drives both transcription and translation, and lipids or at least fatty acids to build a cell membrane. We can derive that early selection pressure for the metabolic network was (i) to form these molecules at a sufficient quantity and specificity, and (ii) to be able to degrade these molecules to maintain homeostasis, gain efficiency, and escape thermodynamic equilibrium. It adds that metabolism functions far from equilibrium with its chemical environment, which highlights the need for compartmentalization. Indeed, early metabolism could not form in a chemical environment that only builds up metabolites but does not provide the ability to metabolize them.

It was considered for a long time that also an ‘RNA world’ could give rise to a metabolic network, in the sense that early metabolic reactions could have been catalyzed by ribozymes. However, there are meanwhile substantial question marks about an RNA-based origin of metabolic catalysis. First, despite sequencing thousands of genomes, there is still no evidence for a considerable amount of ribozymes catalyzing central metabolic reactions indicating that ribozymes are not efficient in catalyzing the reactions found in central metabolism. And second, an Ockham’s razor applies, because many metabolism-like chemical reactions can be catalyzed by the (prebiotically available) amino acids and metal ions, there was no selective advantage for a much more complex ribozyme to evolve for catalyzing similar reactions. And indeed, in modern cells, a vast majority of metabolic catalysis can be attributed to the reaction properties of amino acids, nucleotides, and metal ions while at the same time polymeric ribonucleotides are absent from central metabolism.

Biological complexity gives rise to new constraints in the evolution of metabolism

A key step in the evolution of modern metabolism is, however, new constraints that appeared with the emergence of biological complexity. With all the restrictions on substrate availability, the origin of life had several hundred million years without the degree of ecological competition modern cells are exposed to. This means, one of the key constraints of modern metabolism was less important to early life, the speed of the biochemical reactions and the speed of their uptake. Early forms of the metabolic network could have very well contained much slower reactions than modern networks if they provided or converted the metabolites at a sufficient specificity or one that provided sufficient selective advantage. The prime example of an enzymatic reaction where the selective advantage outweighs a slow reaction rate and low product yields is Ribulose-1,5-bisphosphat-carboxylase/-oxygenase (RuBisCO). Despite catalyzing one of the slowest biochemical reactions discovered so far, Rubisco is the most abundant enzyme on Earth. It provides an autocatalytic route for

fixing carbon dioxide to carbohydrates, bypassing gluconeogenesis for sugar synthesis. It is plausible that many enzymatic reactions started from slow reaction rates that did provide low product yields initially. However, it is equally plausible that evolvable enzyme structures that facilitate faster reaction rates became increasingly necessary with mounting chemical and, eventual, ecological competition (Figure 1).

Computational simulation studies have shown that organizational transition from a chemical reaction system to one that is controlled by a polymeric genetic material is plausible, albeit a highly simplified ‘toy version’ of quantum chemistry was used. Similar approaches have demonstrated that a small set of starting components, such as one might find on the early earth, could populate large areas of the modern-day biochemical metabolic network [25–27]. But *in vitro*, the scope of prebiotic reactions networks reported so far is limited to just a few products representing only a fraction of plausible molecules that can be generated. A key aspect missing is the control over reactivity. Without such control flux is channeled into non-recoverable products or trapped as metastable intermediates [28]. As many of the reactions generating these intermediates exist by virtue of laws in chemistry, a major bottleneck would have been the rise of mutually compatible, stable catalytic structures to manipulate different proto-metabolic reactions so as to sustain them. This brings us to the next constraint — the need for complex enzyme structures.

With the prebiotic availability of amino acids and metal ions, it seems plausible that the early metabolic enzymes were made from the very same components as modern enzymes, that is, amino acids, and for reaction properties that cannot be catalyzed by the amino acid side chains themselves, nucleotides, small organic molecules, and metal ions were sequestered as co-factors. These integrated into the non-enzymatic networks and provided an advantage by accelerating them with increasing competition (Figure 1). In a recent communication, Noda-Garcia et al. [29] have introduced an integrated metabolite-enzyme coevolution model that attempts to explain how associations between emerging catalytic polymers and the existing proto-metabolic network arose. A non-enzymatic reaction or a side product formed due to promiscuous reactivity of an existing enzyme provides a context for evolutionary innovation. Under these circumstances, a step-wise evolution of enzymes and pathways is possible, in contrast to requiring multiple enzymes, to develop simultaneously for a pathway to emerge. Such an expansion of the core network requires high levels of compatibility of the newly evolved reactions or products with the existing network requiring sufficient substrate allocation to produce the metabolite of interest in quantities sufficient to elicit an advantageous

phenotype, as well as stability against the numerous reactive species that are already part of the metabolic network. Indeed, interactions between metabolites formed as part of the metabolic network and their potential to inhibit unrelated enzymatic reactions because of structural similarity with the individual substrates have been studied recently [30]. It has been found that inhibitory interactions by unrelated metabolites affect most enzymes, and compartmentalization mitigates this problem in eukaryotic cell's metabolism. In other words, unspecific inhibitory interactions limit the complexity, and hence, the expansion of the metabolic network (Figure 2). Therefore, despite thermodynamics allowing a larger chemical space that could be the basis for a much larger set of metabolic pathways, metabolic networks are biochemically constrained in their expansion due to their intrinsic functional properties.

Niche adaptation and the continued evolution of the modern metabolic network

The process of adjusting cellular networks, such as metabolism, in response to environmental changes is known as niche adaptation and is stimulated by increasing ecological competition. Selective advantages such as the ability to use nutrients and to achieve faster metabolic rates drive this process. Thus, niche adaptation can serve as a framework to understand how networks expand and is used herein as a model to illustrate how the evolution of modern metabolism is constrained. Recent advances in systems biology contributed to the predictability of the underlying evolutionary trajectories of niche adaptation. Here, we consider two trajectories of metabolic network expansion. One of which involves enzyme promiscuity also referred to as underground metabolism and the other horizontal gene transfer and pre-adaptation serve as a stepping stone to adapt to novel environments.

Besides catalytic activities for which enzymes are evolved, referred to as the native activity, enzymes frequently display the ability to catalyze the turnover of non-native substrates with low catalytic efficiency known as underground reactions [31–33]. Typically, these products are also formed in non-enzymatic reactions that involve metabolites; these can be either spontaneous, or again catalyzed by metabolism-typical metal ions or other biological small molecules [5]. The underground activities provide the (genetic) basis to add to the catalytic properties of environmental molecules and increase the repertoire of possible biochemical reactions to innovate novel metabolic pathways by pathing together existing pathways [34]. Thus, the part of metabolism involving underground reactions that patch together existing pathways is known as underground metabolism. The novel pathways become selective traits once such novel pathways enable the utilization of new nutrients [35]. Recently, an underground metabolic network of

Escherichia coli has been reconstructed, and by using computational and experimental systems biology, it has been demonstrated that underground metabolic reactions are not only prevalent but also facilitate niche adaptation [36]. The first reconstructed *E. coli* metabolic network, including underground metabolism, consists of 260 experimentally verified underground reactions of which half completely fit into the native metabolic network, that is, both substrate and product of the metabolic reaction are also present in the native network. It has been predicted by genome-scale metabolic modeling that underground metabolism enables or enhances growth in fifty environments [36]. Interestingly, these predictions agree with high-throughput gene over-expression experiments illustrating the predictive power of the metabolic model. Since gene over-expression does not prove the evolutionary capability of underground metabolism, adaptive laboratory evolution has recently been conducted [37]. The study revealed that *E. coli* can adapt quickly to the predicted novel environments by gradually changing the environment from a preferred carbon source for growth to a novel carbon source (in approximately twenty generations). In line with the observed short-term adaptation, genome sequencing revealed that only a few novel causal mutations lead to adaptation and that these mutations affect either substrate binding or regulation of gene expression of the enzymes with the predicted underground activity [37]. Even though causal mutations directly affected the underground activity in this study, they cannot be generalized [38]. Besides facilitating growth on new substrates, underground metabolism also provides the selective advantage to cope with detrimental mutations in existing pathways [39,40]. Taken together, these recent studies do not only demonstrate that underground metabolism plays an important role in niche adaptation, which, in turn, could trigger further niche expansion [41], but also show that adaptation is predictable with systems biology.

Niche adaptation does not only involve strategies to use novel nutrients needed for growth but it also involves the evolution of anti-stress mechanisms to cope with changing environments [42]. Underground metabolism seems to play a role in such adaptations suggested by a recent study on stress responses in yeast [43]. Yeast cells import much higher lysine concentrations than needed for growth, and lysine is converted to cadaverine by an underground activity of the first enzyme of the polyamine pathway (Spe1p). In this cellular state, NADPH is not needed for the synthesis of lysine, and as a result, the cell can maintain much higher levels of reduced glutathione, along with a higher flux in required NADPH forming reactions. This leads to a decrease in the concentration of reactive oxygen species and increased oxidant tolerance. We anticipate that adaptive laboratory evolution experiments in combination with ~omics experiments will increasingly shed light on the question of whether the observed anti-stress

mechanism in high lysine concentrations is an adaptive trait by pinpointing causal (selective) mutations that affect regulatory and metabolic states of native and underground metabolism.

Another evolutionary trajectory of niche adaptation is known as the stepping-stone model of metabolic network expansion [44,45]. The model explains complex adaptations and relies on horizontal gene transfer and pre-adaptations in dynamic changes environments. Let us say two enzymes (enzymes 1 and 2) are required for growth in environment A but only one of the two enzymes are needed to enable growth in environment B (enzyme 2). Hence, adaptation to environment A would only require the addition (transfer) of one enzyme (enzyme 1) once pre-adapted to environment B. Thus, this stepping-stone evolutionary trajectory allows adaptation to environment A, which would or not likely be accessible at a realistic evolutionary time-scale without first adapting to environment B. Indeed, a recent systems biology approach predicts via integration of genome-scale metabolic modeling and horizontal gene transfer that such stepping-stone trajectories can happen [44]. Importantly, phylogenetic reconstruction of gene gain (transfer) patterns on the basis of hundreds of sequenced genomes, as well as laboratory experiments, are consistent with these predictions. Nonetheless, it is worth mentioning the above-mentioned constraint of biochemical compatibility between the newly acquired reactions and the existing metabolic system also applies in this context [30], that is, newly introduced enzymes can only provide an advantage if they are not inhibited by the existing metabolome, or if they do not disadvantage the cell by synthesizing inhibitors to existing metabolic reactions.

Important questions regarding the patchwork model of underground metabolism and the stepping-stone model still remain despite major advances in our understanding [45,46]. First, which of the two models explain most of the metabolic network expansion and second, are these two models linked and can they be unified into an overall model? Interestingly, a recent study suggests a link between horizontal gene transfer and underground metabolism in compensating for the deletion of L-threonine dehydratase in *E. coli* [47]. With advances in the development of underground metabolic networks, phylogenetic approaches, ALE, ~omics and genome engineering techniques [48–52], it is expected that future studies will provide answers to these open questions.

Why has the metabolic network maintained its basic structure over billions of years?

Despite niche pressure changing over time, the metabolic network seems to have maintained its basic structure over the extreme timeframe of its evolution. There are several factors that contribute to this situation. First,

as the Russian Chemist A.G. Golubev commented, ‘nothing of chemistry disappears in biology’ [53]. If the biochemical topology came into place not only because of the progressive evolution of ever-better enzymes but also because of the reaction property of the metabolites in the presence of the metal ions and biological small molecules that always surround them evolution will revert to the same topology in different environments. This is specifically true if species adapt to more extreme or resource-scarce environments, where reaction specificity becomes an increasing constraint. Second, despite the huge diversity of enzyme structures, the basic molecules that conduct catalysis remained unchanged over billions of years of evolution. These include amino acid side chains, that act in most modern enzymes’ catalytic pockets, nucleotides, a few universal organic molecules that act as cofactors (PALP for instance), as well as just a handful of life-essential metal ions, including iron, zinc, magnesium, and copper. Together, these explain the vast majority of biochemical reactions. Finally, there is also evidence of network effects. The complexity of enzyme–metabolite interactions constrains the expansion of the metabolic network. On average, each cellular enzyme is specifically inhibited by at least five other cellular metabolites [30]. Each enzyme has evolved sufficient activity to be active despite the presence of these inhibitors. Each novel metabolite added to the network will inhibit additional enzymes, creating a strong selection pressure to keep the chemical diversity, as well as metabolite concentrations, as low as possible. This situation limits the chance of new pathways to evolve and to expand, specifically, as these can only provide a selective advantage if they either allow a new adaptation or if they outperform the existing metabolite pathway. The latter explains why metabolic pathways have very limited redundancies.

Conflict of interest statement

Nothing declared.

Acknowledgements

Work by MR was supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (FC001134), the UK Medical Research Council (FC001134), and the Wellcome Trust (FC001134), the Eisenhower Foundation, the Deutsche Forschungsgemeinschaft (DFG), the Bundesministerium für Bildung und Forschung, the European Research Council and the Wellcome Trust. Work by RAN at Wageningen University & Research (WUR) is supported by internal funding from WUR.

References

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
1. Dandekar T, Schuster S, Snel B, Huynen M, Bork P: **Pathway alignment: application to the comparative analysis of glycolytic enzymes.** *Biochem J* 1999, **343**(Pt 1):115–124.
 2. Wächtershäuser G: **Evolution of the first metabolic cycles.** *Proc Natl Acad Sci USA* 1990, **87**:200–204.
 3. Martin W, Baross J, Kelley D, Russell MJ: **Hydrothermal vents and the origin of life.** *Nat Rev Microbiol* 2008, **6**:805–814.

4. Keller MA, Turchyn AV, Ralser M: **Non-enzymatic glycolysis and pentose phosphate pathway-like reactions in a plausible Archean ocean.** *Mol Syst Biol* 2014, **10**:725.
5. Keller MA, Piedrafita G, Ralser M: **The widespread role of non-enzymatic reactions in cellular metabolism.** *Curr Opin Biotechnol* 2015, **34**:153–161.
6. Andreini C, Bertini I, Rosato A: **Metalloproteomes: a bio-informatic approach.** *Acc Chem Res* 2009, **42**:1471–1479.
7. Andreini C, Bertini I, Cavallaro G, Holliday GL, Thornton JM: **Metal ions in biological catalysis: from enzyme databases to general principles.** *J Biol Inorg Chem* 2008, **13**:1205–1218.
8. Zahnle K, Schaefer L, Fegley B: **Earth's earliest atmospheres.** *Cold Spring Harb Perspect Biol* 2010, **2**:a004895.
9. Rouxel OJ, Bekker A, Edwards KJ: **Iron isotope constraints on the Archean and Paleoproterozoic ocean redox state.** *Science* 2005, **307**:1088–1091.
10. **The bioinorganic chemistry of the ancient ocean: the co-evolution of cyanobacterial metal requirements and biogeochemical cycles at the Archean–Proterozoic boundary?** *Inorg Chim Acta* 2003, **356**:308–318.
11. Keller MA, Kampjut D, Harrison SA, Ralser M: **Sulfate radicals enable a non-enzymatic Krebs cycle precursor.** *Nat Ecol Evol* 2017, **1**:83–89.
12. Muchowska KB, Varma SJ, Chevallot-Beroux E, Lethuillier-Karl L, Li G, Moran J: **Metals promote sequences of the reverse Krebs cycle.** *Nat Ecol Evol* 2017, **1**:1716–1721.
13. Knowles CJ: **Cyanide utilization and degradation by microorganisms.** *Ciba Found Symp* 1988, **140**:3–15.
14. Marx CJ, Van Dien SJ, Lidstrom ME: **Flux analysis uncovers key role of functional redundancy in formaldehyde metabolism.** *PLoS Biol* 2005, **3**:e16.
15. Stubbs RT, Yadav M, Krishnamurthy R, Springsteen G: **A plausible metal-free ancestral analogue of the Krebs cycle composed entirely of α -ketoacids.** *Nat Chem* 2020, **12**:1016–1022.
16. Barge LM, Flores E, Baum MM, VanderVelde DG, Russell MJ: **Redox and pH gradients drive amino acid synthesis in iron oxyhydroxide mineral systems.** *Proc Natl Acad Sci USA* 2019, **116**:4828–4833.
17. Sleep NH, Bird DK, Pope EC: **Serpentinite and the dawn of life.** *Philos Trans R Soc Lond B Biol Sci* 2011, **366**:2857–2869.
18. *Prebiotic photochemistry: from Urey-Miller-like experiments to recent findings.* Royal Society of Chemistry; 2021 [no volume].
19. Ralser M: **An appeal to magic? The discovery of a non-enzymatic metabolism and its role in the origins of life.** *Biochem J* 2018, **475**:2577–2592.
20. Sousa FL, Martin WF: **Biochemical fossils of the ancient transition from geoenergetics to bioenergetics in prokaryotic one carbon compound metabolism.** *Biochim Biophys Acta* 2014, **1837**:964–981.
21. Muchowska KB, Varma SJ, Moran J: **Nonenzymatic metabolic reactions and life's origins.** *Chem Rev* 2020, **15**:7708–7744.
22. Shen C, Yang L, Miller SL, Oro J: **Prebiotic synthesis of histidine.** *J Mol Evol* 1990, **31**:167–174.
23. Powner MW, Gerland B, Sutherland JD: **Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions.** *Nature* 2009, **459**:239–242.
24. Saladino R, Crestini C, Pino S, Costanzo G, Di Mauro E: **Formamide and the origin of life.** *Phys Life Rev* 2012, **9**:84–104.
25. Handorf T, Ebenhöf O, Heinrich R: **Expanding metabolic networks: scopes of compounds, robustness, and evolution.** *J Mol Evol* 2005, **61**:498–512.
26. Goldford JE, Hartman H, Smith TF, Segrè D: **Remnants of an ancient metabolism without phosphate.** *Cell* 2017, **168**:1126–1134. e9.
27. Goldford JE, Segrè D: **Modern views of ancient metabolic networks.** *Curr Opin Struct Biol* 2018, **8**:117–124.
28. Orgel LE: **The implausibility of metabolic cycles on the prebiotic Earth.** *PLoS Biol* 2008, **6**:e18.
29. Noda-Garcia L, Liebermeister W, Tawfik DS: **Metabolite-enzyme coevolution: from single enzymes to metabolic pathways and networks.** *Annu Rev Biochem* 2018, **87**:187–216.
- Introduction of an integrated metabolite-enzyme coevolution model that attempts to explain how new metabolic pathways arose through new metabolites and enzymes and the order or their addition.
30. Alam MT, Olin-Sandoval V, Stincone A, Keller MA, Zelezniak A, Luisi BF, Ralser M: **The self-inhibitory nature of metabolic networks and its alleviation through compartmentalization.** *Nat Commun* 2017, **8**:16018–16113.
31. Lerma-Ortiz C, Jeffryes JG, Cooper AJL, Niehaus TD, Thamm AMK, Frelin O, Aunins T, Fiehn O, de Crécy-Lagard V, Henry CS, *et al.*: **“Nothing of chemistry disappears in biology”: the Top 30 damage-prone endogenous metabolites.** *Biochem Soc Trans* 2016, **44**:961–971.
32. Khersonsky O, Tawfik DS: **Enzyme promiscuity: a mechanistic and evolutionary perspective.** *Annu Rev Biochem* 2010, **79**:471–505.
33. D’Ari R, Casadesús J: **Underground metabolism.** *Bioessays* 1998, **20**:181–186.
34. Notebaart RA, Kintses B, Feist AM, Papp B: **Underground metabolism: network-level perspective and biotechnological potential.** *Curr Opin Biotechnol* 2018, **49**:108–114.
35. Jensen RA: **Enzyme recruitment in evolution of new function.** *Annu Rev Microbiol* 1976, **30**:409–425.
36. Notebaart RA, Szappanos B, Kintses B, Pál F, Györkei Á, Bogos B, Lázár V, Spohn R, Csörgő B, Wagner A, *et al.*: **Network-level architecture and the evolutionary potential of underground metabolism.** *Proc Natl Acad Sci USA* 2014, **111**:11762–11767.
37. Guzmán GI, Sandberg TE, LaCroix RA, Nyerges Á, Papp H, de Raad M, King ZA, Hefner Y, Northen TR, Notebaart RA, *et al.*: **Enzyme promiscuity shapes adaptation to novel growth substrates.** *Mol Syst Biol* 2019, **15**, e8462.
- Study providing empirical evidence that underground metabolism plays a role in adaptive evolution to novel nutrient environments.
38. Morgenthaler AB, Kinney WR, Ebmeier CC, Walsh CM, Snyder DJ, Cooper VS, Old WM, Copley SD: **Mutations that improve efficiency of a weak-link enzyme are rare compared to adaptive mutations elsewhere in the genome.** *Elife* 2019, **8**.
- Study showing empirical evidence that adaptive mutations to improve the efficiency of an essential promiscuous function rarely happens directly within the promiscuous function itself, but mostly elsewhere in the genome.
39. Cotton CA, Bernhardsgrütter I, He H, Burgener S, Schulz L, Paccia N, Dronsella B, Erban A, Toman S, Dempfle M, *et al.*: **Underground isoleucine biosynthesis pathways in E. coli.** *Elife* 2020, **9**.
40. Oberhardt MA, Zarecki R, Reshef L, Xia F, Duran-Frigola M, Schreiber R, Henry CS, Ben-Tal N, Dwyer DJ, Gophna U, *et al.*: **Systems-wide prediction of enzyme promiscuity reveals a new underground alternative route for pyridoxal 5'-phosphate production in E. coli.** *PLoS Comput Biol* 2016, **12**, e1004705.
41. Barve A, Wagner A: **A latent capacity for evolutionary innovation through exaptation in metabolic systems.** *Nature* 2013, **500**:203–206.
42. Bommer GT, Van Schaftingen E, Veiga-da-Cunha M: **Metabolite repair enzymes control metabolic damage in glycolysis.** *Trends Biochem Sci* 2020, **45**:228–243.
- Discussion of how metabolite repair enzymes remove side-products from glycolysis.
43. Olin-Sandoval V, Yu JSL, Miller-Fleming L, Alam MT, Kamrad S, Correia-Melo C, Haas R, Segal J, Peña Navarro DA, Herrera-

Dominguez L, *et al.*: **Lysine harvesting is an antioxidant strategy and triggers underground polyamine metabolism.** *Nature* 2019, **572**:249–253.

44. Szappanos B, Fritzsche J, Csörgő B, Lázár V, Lu X, Fekete G, Bálint B, Herczeg R, Nagy I, Notebaart RA, *et al.*: **Adaptive evolution of complex innovations through stepwise metabolic niche expansion.** *Nat Commun* 2016, **7**:11607.

Introduction of a model to explain complex adaptations. The model relies on horizontal gene transfer and pre-adaptations in dynamic changes environments.

45. Pál C, Papp B: **Evolution of complex adaptations in molecular systems.** *Nat Ecol Evol* 2017, **1**:1084–1092.
46. Glasner ME, Truong DP, Morse BC: **How enzyme promiscuity and horizontal gene transfer contribute to metabolic innovation.** *FEBS J* 2020, **287**:1323–1342.
47. Jerlström Hultqvist J, Warsi O, Söderholm A, Knopp M, Eckhard U, Vorontsov E, Selmer M, Andersson DI: **A bacteriophage enzyme induces bacterial metabolic perturbation that confers a novel promiscuous function.** *Nat Ecol Evol* 2018, **2**:1321–1330.

Study providing evidence of a link between horizontal gene transfer and underground metabolism in the adaptation to detrimental mutations.

48. Sandberg TE, Salazar MJ, Weng LL, Palsson BØ, Feist AM: **The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology.** *Metab Eng* 2019, **56**:1–16.
49. Amin SA, Chavez E, Porokhin V, Nair NU, Hassoun S: **Towards creating an extended metabolic model (EMM) for E. coli using enzyme promiscuity prediction and metabolomics data.** *Microb Cell Factories* 2019, **18**:109–112.
50. Sévin DC, Fuhrer T, Zamboni N, Sauer U: **Nontargeted in vitro metabolomics for high-throughput identification of novel enzymes in Escherichia coli.** *Nat Methods* 2017, **14**:187–194.
51. Davidi D, Milo R: **Lessons on enzyme kinetics from quantitative proteomics.** *Curr Opin Biotechnol* 2017, **46**:81–89.
52. Messner CB, Demichev V, Bloomfield N, Yu JSL, White M, Kreidl M, Egger A-S, Freiwald A, Ivosev G, Wasim F, *et al.*: **Ultrafast proteomics with scanning SWATH.** *Nat Biotechnol* 2021, **39**:846–854.
53. Golubev A: **How could the Gompertz-Makeham law evolve.** *J Theor Biol* 2009, **258**:1–17.