Modern views of ancient metabolic networks

Joshua E. Goldford[1] and Daniel Segrè[1,2,3,4]

Abstract

Metabolism is a molecular, cellular, ecological and planetary phenomenon, whose fundamental principles are likely at the heart of what makes living matter different from inanimate one. Systems biology approaches developed for the quantitative analysis of metabolism at multiple scales can help understand metabolism's ancient history. In this review, we highlight work that uses network-level approaches to shed light on key innovations in ancient life, including the emergence of proto-metabolic networks, collective autocatalysis and bioenergetics coupling. Recent experiments and computational analyses have revealed new aspects of this ancient history, paving the way for the use of large datasets to further improve our understanding of life's principles and abiogenesis.

Addresses

[1] Bioinformatics Program and Biological Design Center, Boston University, Boston, MA 02215, United States
[2] Department of Biology, Boston University, Boston MA 02215, United States
[3] Department of Biomedical Engineering, Boston University, Boston MA 02215, United States
[4] Department of Physics, Boston University, Boston MA 02215, United States

Corresponding author: Segrè, Daniel, Bioinformatics Program, Boston University, Boston, MA 02215, United States. (dsegre@bu.edu)

Keywords

Metabolism, Evolution, Network expansion, Origin of life.

Introduction

The metabolic network of a cell transforms free energy and environmentally available molecules into more cells, moving electrons step by step along gradients in a complex energetic landscape [1,2]. The ability of a cell to efficiently and simultaneously manage hundreds of metabolic processes so as to accurately balance the production of its internal components constitutes a very complex resource allocation problem. In fact, it is only through recent systems biology research that we have begun to quantitatively assess this resource allocation problem at the whole-cell level [3,4]. A common perspective in the analysis of cellular self-reproduction is the notion that the genome, with its crucial information-storage role, is the central molecule of the cell, and that everything else can be collectively regarded as the machinery whose role is to produce a copy of the DNA. It is therefore not surprising that, as we struggle with the fascinating question of how life started on a lifeless planet, it is tempting to look for how a single information-containing molecule could arise spontaneously from prebiotic compounds. However, in spite of the appeal of thinking of DNA (or its historically older predecessor, RNA) as the central molecule that is being replicated in the cell, no molecule in the cell really self-replicates: the cell is a network of chemical transformations capable of collective autocatalytic self-reproduction. Collective autocatalysis is the capacity for an ensemble of chemicals to enhance or catalyze the synthesis or import of its own components, enabling a positive feedback mechanism that can lead to their sustained amplification. Combining this systems-level view of a cell with the argument of what is usually called the "metabolism first" view of the origin of life, one could propose that the ability of a chemical network to produce more of itself (or to grow autocatalytically) is and has always been a key hallmark of life [5−9]. An interesting modern version of this very same principle is embedded in one of the most popular systems biology approaches for the study of whole cell metabolism: this approach, based on reaction network stoichiometry and efficient constraint-based optimization algorithms, is commonly known as flux balance analysis (FBA) [3]. FBA solves mathematically the resource allocation problem that every living cell needs to solve in order to transform available nutrients into the macromolecular building blocks that are necessary for maintenance and reproduction. When an FBA calculation estimates the maximal growth capacity of a cell, it essentially computes the set of reaction network fluxes that enable optimally efficient autocatalytic self-reproduction. While in cellular life this process is finely regulated and controlled, ancient life must have gone through many different stages of similar, but much less organized collectively autocatalytic processes. Thus, one of the key problems of the origin of life is the question of how an initially random path in the space of possible chemical transformations driven far from thermodynamic equilibrium could have ended up being dynamically "trapped" in a collectively autocatalytic state.

The focus on cellular self-reproduction as the fundamental level at which life and its origin should be

understood is however too narrow. An exciting recent development in systems biology of metabolism is the rise of methods to extend FBA models from the genome scale to the ecosystem level [10–12]. In addition to solving the resource allocation problem of metabolism for individual organisms in a given environment, these approaches take into account the fact that metabolites can be exchanged across species, giving rise to metabolically-driven ecological networks [11]. These advances suggest that metabolism may be best understood as an ecosystem-level phenomenon (Figure 1b), where the collective biochemical capabilities of multiple co-existing organisms may reflect — better than any individual metabolic network — an optimal capacity of life to utilize resources present in a given environment [9,13]. The ecosystem-level nature of metabolism is another feature of present-day life whose roots likely date back to the early stages of life on our planet. For example, the chemical networks that gradually gave rise to reproducing protocells may have wandered for quite some time in a broader chemical space, effectively generating molecular ecosystems before the rise of spatially and chemically well-defined cellular structures.

At an even larger scale, metabolism can be viewed as operating not just at the level of individual cells or ecosystems, but as a planetary phenomenon, in which cellular processes collectively affect (and are affected by) the flow of molecules at geological scales (Figure 1c). The strong coupling between the metabolic processes of ecosystems and planetary-scale geochemistry [14,15] suggest that biosphere-level metabolism should be viewed as one of the natural scales for the study of life's history. A paramount challenge in the study of life's history is thus bridging the gap between material and energy fluxes at the biosphere scale, and detailed molecular mechanisms responsible for the

properties of life at the cellular and subcellular level [16]. Bridging this gap could greatly benefit from the use of integrative models similar to the ones used in systems biology research and data science. In this perspective, we will discuss some recent system-level approaches that have provided new important insight into life's ancient history at multiple scales, highlighting that metabolism and its multiscale nature — from the single reaction to the biosphere — are taking a center stage role in this endeavor.

Protometabolism before enzymes

A top-down reconstruction of ancient metabolic networks can be achieved based on the inferred history of gene families, using traditional phylogenomic techniques [17–20]. Leveraging information on the newly mapped genomic diversity of modern life [21], Martin and colleagues recently proposed a comprehensive phylogenetic reconstruction of the metabolic capabilities of the last universal common ancestor (LUCA), suggesting that LUCA was an autotrophic, thermophilic, N_2-fixing anaerobic prokaryote, living in hydrothermal vents and equipped with life's most complex molecular machines (e.g. ATP synthase) [22]. Although the details of LUCA's specific repertoire of metabolic enzymes are still subject of debate [23,24], these results corroborate the notion that LUCA was very complex, highlighting a massive gap in knowledge with regard to the transition from prebiotic geochemical processes to the biochemical complexity of LUCA and its progeny.

A major challenge in the study of the origin of metabolic networks is to gain insight on the structure of metabolic networks before LUCA and before the rise of genetic coding (Figure 2). At the core of this challenge is the question of whether and how metabolic reactions, which depend on genome-encoded enzymes in modern cells,

Figure 1
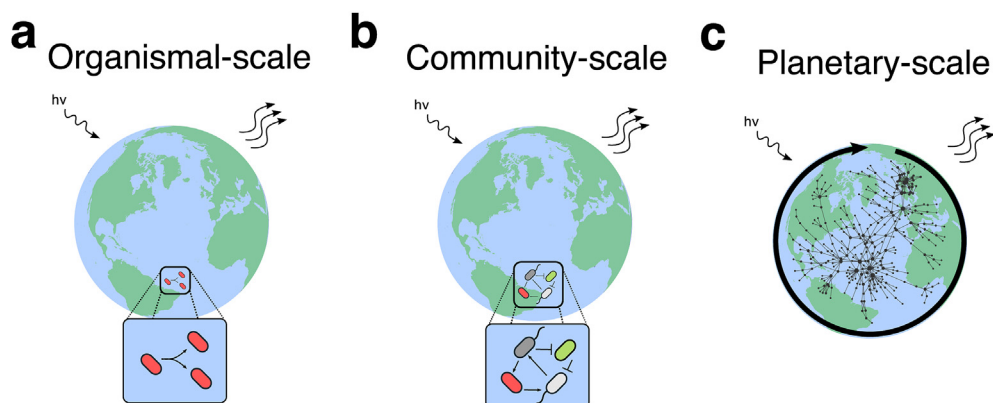
Metabolism at different scales. (a) Metabolic networks can be modeled at the organismal level, where environmentally-supplied resources, under an energy flow, are collectively transformed into biomass of the self-reproducing and evolving organisms. (b) At higher scales, metabolic networks can be viewed as an ecosystem-level phenomenon, where biochemical processes include metabolic exchange and competition between species. (c) Metabolism can be also considered a planetary scale phenomenon, whereby the free energy flow maintains global biogeochemical cycles.

Figure 2
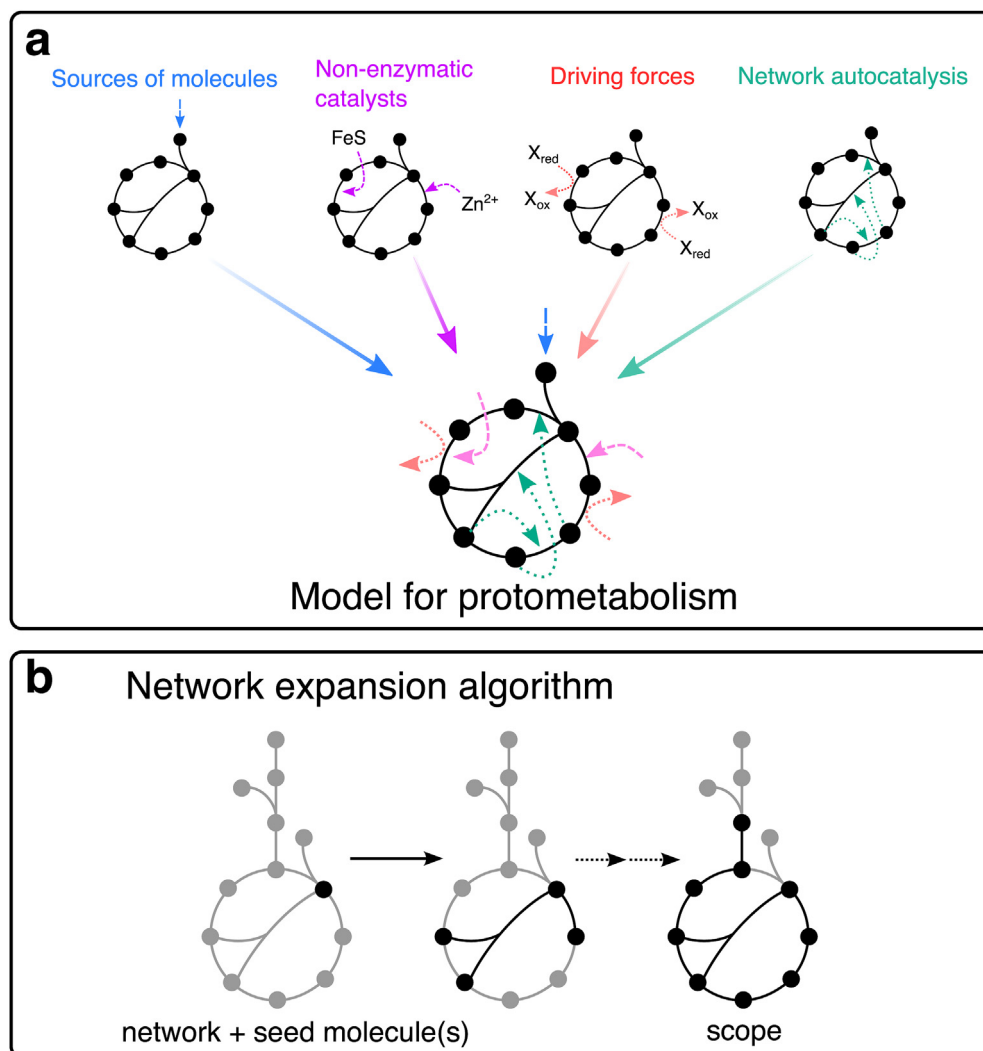
Towards a model of ancient metabolism. (a) Models of protometabolism can be constructed using a wide range of data including geochemically-supported data of environments and atmospheres in the early Archean Eon, knowledge of non-enzymatic chemical reactions, plausible driving forces keeping protometabolism out of equilibrium and a mechanism for sustained growth (e.g. network autocatalysis). (b) The structure of plausible networks can be investigated using the network expansion algorithm which models the integrative expansion of metabolic networks from a set of seed compounds and allowable chemical reactions.

could have been carried out without such enzymes, resulting in a classical origin of life "chicken-and-egg" problem. One possible way out of this conundrum is the possibility that some of these metabolic reactions were initially catalyzed by less sophisticated and less specific catalysts, such as small organic molecules, metal ions, minerals, short RNA polymers, prebiotic amino acids or peptides. These small molecules could have persisted throughout evolution, gradually becoming incorporated into protein enzymes as catalytic cores or cofactors [25,26]. Adding on to a large body of evidence on individual metabolic reactions being catalyzed by small molecules [27,28], recent experimental work has demonstrated that several key pathways found in modern day metabolic networks can be catalyzed non-

enzymatically [29–34]. For instance, Ralser and colleagues [29–32,35] have shown the feasibility of non-enzymatic networks that resemble modern day biochemical pathways, including the TCA cycle, glycolysis and gluconeogenesis. In addition, Moran and colleagues have demonstrated that metals can selectively catalyze and drive portions of non-enzymatic reductive TCA cycle (rTCA) [33]. These experimental results support the hypothesis that the catalytic cores of some modern enzymes may represent evolved variants of simple geochemically available prebiotic catalysts like transition metals, iron-sulfur clusters or organic cofactors [36]. Despite these important advances, the known instances of non-enzymatic catalysis are still the tip of the iceberg relative to the large

number of possible catalyst-reaction pairs. Future high-throughput experiments could greatly expand the scope of possible prebiotic chemical networks and test the limits of non-enzymatic catalysis, shedding important light on the complexity of prebiotic chemistry obtainable before the availability of proteinaceous enzymes.

From non-enzymatic catalysis to collective autocatalysis

The above examples illustrate the fact that chemical reactions, and whole pathways, typically viewed in biological context as feasible only in the presence of protein enzymes, could take place under much more primitive conditions, through the catalytic action of small molecules, minerals or even non-covalent supramolecular assemblies [37]. As mentioned above, however, a major leap in the history of life must have involved the rise of a collectively autocatalytic chemical system. The feasibility of pre-enzymatic chemistry suggests a possible path for the rise of such collective autocatalysis: if the molecules produced by these reactions are themselves good catalysts, or if these reactions contribute to solubilize inorganic catalysts from rocks, there is a chance that a subset of reactions and molecules will effectively display a dynamic behavior that is equivalent to that of a single autocatalytic, exponentially growing entity [5–8].

Insight into how these autocatalytic sets may have operated has come from both theoretical and experimental work. Recent theoretical work has uncovered generic constraints of autocatalytic networks [38], and offered plausible biophysical mechanisms leading to sustained autocatalysis of biopolymer ensembles [39,40]. Experimentally, Whitesides and colleagues have constructed an autocatalytic chemical network based on simple, biologically relevant organic compounds [41]. In particular, by using a continuous flow of nutrients into and out of their reaction vessel, they showed that simple mixtures of thiols and thioesters could display a wide range of dynamical properties, such as bistability, oscillations and autocatalysis. Notably, this work demonstrated that dynamical properties observed in biological networks can emerge from simple mixtures of prebiotically plausible chemicals held out of equilibrium. As described recently by Vetsigian and Baum, the time is ripe for experimental explorations of how collectively autocatalytic cycles could spontaneously arise from mixtures of small molecules and mineral surfaces [42].

Navigating possible paths from primordial to present-day networks

Studies in evolutionary biology suggest that biological systems evolve by partially building on prior innovations. If this principle extends back to the origin of metabolic networks, then it is reasonable to hypothesize that early proto-metabolic networks were based on previously accessible chemistry. Such logic leads to the conjecture that the structure of metabolic networks encodes the evolutionary history of metabolism, and that the chemistry of core metabolism is similar to the initial abiotic chemical networks that lead to life's emergence [43,44]. This conjecture is supported, as discussed above, by experimental work demonstrating that a significant portion of core metabolism is accessible without the use of proteinaceous enzymes. While these concepts have been heavily utilized in origin of life research, recent efforts have transformed this conceptual paradigm into an algorithmic and quantitative framework using metabolic network modeling [19,45]. A modeling approach recently used to explore the plausible evolutionary history of very early stages of biochemistry is the network expansion algorithm (Figure 2b), which iteratively simulates the growth of new metabolites and reactions starting from an initial seed set [46–48]. We used the network expansion algorithm to construct a model for ancient prebiotic metabolism, specifically addressing the question of whether any portion of current biochemistry could have possibly emerged in the absence of phosphate (and thus prior to transcription/translation) [45]. Models of prebiotic networks were constructed starting from minimal sets of compounds thought to have been readily available on early Earth. Notably, even if these initial compounds did not include any phosphate-containing molecule, a surprisingly large expanded network could ensue, covering several pathways that are part of central metabolism today, and of previously proposed models of biogenesis [44]. This finding is consistent with the possibility that thioesters, sulfur-based energy rich chemical moieties, could have predated phosphates as energy carriers in the cell, providing the required thermodynamic driving force. Interestingly, recent work has experimentally demonstrated the possibility that a thioester-based chemistry could fuel autocatalytic networks [41]. Future approaches could extend the use of network expansion models by incorporating additional constraints on metabolic network growth, such as the removal of likely toxic intermediates. Although further experimental and theoretical work is required to fully address the scope, implications, and fundamental limitations of an early phosphate-free biosphere, the use of the network expansion algorithm to explore plausible routes of abiogenesis represents an interesting research direction.

Although the majority of chemical reactions important in early living systems may still be encoded in modern day living systems, there is also a possibility that key reactions and compounds initially critical for living systems were lost throughout the course of evolution. Even more broadly, it is plausible that a much bigger space of chemically possible reactions could have given rise to an organized metabolism [49,50]. As shown in recent elegant experimental work, molecules important for life as we know it may in principle be producible through

reactions and pathways that are not part of current biochemistry [51,52]. On the theoretical side, recent advances in computational chemistry [53] have enabled the construction of chemical network models beyond the scope of modern living systems, paving the way for future broader analyses of possible transient chemistries along the history of life, and of putative alternative outcomes that could have materialized but did not [49].

Beyond these realistic chemical spaces, biochemical organization has been studied extensively using simplified toy models based on artificial chemical rules, such as the so called "string chemistries" [54]. Similar approaches were the foundation of some of the early work on collectively autocatalytic networks [8,55]. More recently, a very simple string chemistry, simulated and analyzed using systems biology approaches (including FBA [3]) yielded a family of optimally efficient pathways, some of which resemble functionally and topologically the rTCA cycle network [56]. An artificial chemistry which incorporated catalytic polymers with a toy folding process was recently shown to be helpful towards explaining the emergence of polymer-based structures within a compositional inheritance world [39]. In addition to serving as a basis to explore possible scenarios for the emergence of metabolism, abstract chemistry models can be very helpful in the exploration of statistical physics-based models of non-equilibrium chemical systems [57].

Overcoming energy barriers, then and now
Whether realistic or abstract, ancient or modern, any metabolism can operate only if kept far from thermodynamic equilibrium by an external free energy source. Thus, to achieve a working theory for the origin of metabolism, one should identify not only sources of materials, but also sources of free energy consistent with geochemical data. Effectively, even if early life may have extensively used abiotic organic material heterotrophically, this question largely hinges on our understanding of what free energy source could have fueled the production of electron donors capable of reducing abundant gases like CO_2 and N_2 into the reduced forms readily used by biological systems. Two potential sources include chemical energy from hydrothermal vents, and photochemical energy from solar (especially UV) radiation [58]. The former scenario is consistent with recent phylogenomic studies [22], where chemical energy in the form of molecular hydrogen is used to fix carbon dioxide using a variant of the Wood-Ljungdahl pathway in LUCA. However, it is unclear whether this scenario would be compatible with thioesters as a key component for free energy transduction, given that these molecules have been recently shown to be highly unstable in simulated hydrothermal systems [59]. Interestingly, UV light can support the synthesis of organic molecules [60,61] and iron-sulfur clusters [62]

as well as drive the reductive steps in the rTCA cycle [63]. Future work exploring the potential roles of various energy sources to fuel non-enzymatic prebiotic networks will be important in determining plausible models for ancient metabolism.

Even if a source of free energy is available, a major open question in the evolution of bioenergetics is the rise of coupling between driving forces and driven reactions. Through this coupling, currently enabled by large proteins, reactions that dissipate free energy (e.g. thioester or phosphodiester bond breaking) drive reactions that require a free energy input. Such couplings have recently been proposed to universally operate as a "Brownian ratchet", in which enzyme complexes rely on the stepwise, gated mechanism of highly coordinated multidomain enzymes [2]. Martin and colleagues proposed that electron bifurcation, the most recently discovered energy conserving process [64,65], may have been the first mechanism through which ancient metabolic networks coupled free energy sources to drive endergonic reactions. Electron bifurcation is a mechanism that enables coupling between available, mid-potential electron donors (e.g. H_2) and acceptors (e.g. CO_2) to generate low-potential electron donors. This mechanism is for example capable of producing reduced ferrodoxin, an energy source common in diverse biochemical pathways like photosynthesis and methanogenesis [66]. In general, identifying the scope of non-enzymatic analogs for such free energy coupling processes remains an open challenge, and efforts to this end will undoubtedly shed light on the earliest phases of bioenergetic evolution.

Towards data-driven origin of life research
As the above examples clearly illustrate, origin of life research is a multidisciplinary endeavor, requiring consideration of multiple, increasingly large datasets (chemical, geological, biological, physical) for both experimental and computational analyses [67]. Currently available databases that may be useful for the study of ancient life range from collections of genetic and phenotypic diversity of microbial species and communities [68], to "knowledge-base" resources available for exploring metabolites, reactions and biochemical pathways [69]. As origin of life research may require data from broader categories of molecules and reactions beyond present-day biochemistry [51], databases of known organic and inorganic chemicals [70] and reactions [50] will constitute important components of future attempts to reconstruct the first biochemical processes. Other categories of data relevant to the ancient history of metabolism are available in more specialized databases [71,72]. Future efforts could assemble other databases useful for the computational analysis of prebiotic chemistry, including a database of documented prebiotic chemistry experiments. Furthermore, and most importantly, a standardization of

experimental and computational results would enable comparisons across different efforts, allowing researchers to build more systematically on previous work. Integrating data from various sources, ranging from prebiotic chemistry experiments to inferred early Earth geochemical data, could allow for the construction of large-scale models of ancient metabolic states at unprecedented levels of resolution.

Future work aimed at understanding early life will increasingly benefit from ongoing synthetic biology efforts towards the implementation of minimal living systems, and from quantitative approaches developed for systems biology of metabolism [73]. It would be highly beneficial for origin of life research to embrace theory and modeling as essential tools for transforming data and hypotheses into testable, nontrivial predictions, i.e. predictions whose outcome may not be known *a priori*, and whose validation or falsification may be clearly achievable, even if technologies may be years away from feasibility. Conversely, the study of early metabolism has a chance to provide new tools and ideas for how to move systems biology approaches beyond the current paradigms. For example, the exploration of putative early metabolic pathways not known in present-day organisms bears some similarities with the huge and challenging efforts of annotating metabolic enzyme functions in newly sequenced genomes and metagenomes [74]. Furthermore, biosphere-level analyses of ancient metabolism [45,48] could inspire new approaches for studying the collective biochemistry of microbial ecosystems.

Acknowledgements

References

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

*   of special interest
* * of outstanding interest

1. Falkowski PG, Fenchel T, Delong EF: **The microbial engines that drive Earth's biogeochemical cycles**. *Science* 2008, **320**: 1034−1039.

2. Branscomb E, Biancalani T, Goldenfeld N, Russell M: **Escapement mechanisms and the conversion of disequilibria; the engines of creation**. *Phys Rep* 2017, **677**:1−60.
* * Br*anscomb* et al. provide critical insight on the attributes of all of life's molecular machines that harvest disequilibria, demonstrating that "brownian ratchet" mechanisms are widespread and critical for the fidelity of reactions that coupling endergonic reactions with dissipative exergonic reactions

3. Orth JD, Thiele I, Palsson BO: **What is flux balance analysis?** *Nat Biotech* 2010, **28**:245−248.

4. Basan M, Hui S, Okano H, Zhang Z, Shen Y, Williamson JR, Hwa T: **Overflow metabolism in Escherichia coli results from efficient proteome allocation**. *Nature* 2015, **528**:99−104.

5. Segré D, Ben-Eli D, Lancet D: **Compositional genomes: prebiotic information transfer in mutually catalytic noncovalent assemblies**. *Proc Natl Acad Sci* 2000, **97**:4112−4117.

6. Kauffman SA: **Autocatalytic sets of proteins**. *J Theor Biol* 1986, **119**:1−24.

7. Dyson FJ: **A model for the origin of life**. *J Mol Evol* 1982, **18**: 344−350.

8. Bagley R, Farmer J, Fontana W: **Evolution of a metabolism**. In *Artificial life II*. Edited by Langton CG, Taylor C, Farmer JD, Rasmussen S, Reading, MA: Addison−Wesley; 1991:141−158.

9. Smith E, Morowitz HJ: *The origin and nature of life on Earth*. 1st ed. Cambridge, United Kingdom: Cambridge University Press; 2016, https://doi.org/10.1017/CBO9781316348772.

10. Klitgord N, Segrè D: **Environments that induce synthetic microbial ecosystems**. *PLoS Comput Biol* 2010, https://doi.org/10.1371/journal.pcbi.1001002.

11. Harcombe WR, Riehl WJ, Dukovski I, Granger BR, Lang AH, Bonilla G, Kar A, Leiby N, Mehta P, Marx CJ, Segrè D: **Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics**. *Cell Rep* 2014, **7**:1104−1115.

12. Freilich S, Zarecki R, Eilam O, Segal ES, Henry CS, Kupiec M, Gophna U, Sharan R, Ruppin E: **Competitive and cooperative metabolic interactions in bacterial communities**. *Nat Commun* 2011, **2**:589.

13. Braakman R, Follows MJ, Chisholm SW: **Metabolic evolution and the self-organization of ecosystems**. *Proc Natl Acad Sci* 2017, **114**:E3091−E3100.

14. Jelen BI, Giovannelli D, Falkowski PG: **The role of microbial electron transfer in the coevolution of the biosphere and geosphere**. *Annu Rev Microbiol* 2016, **70**. annurev-micro-102215−095521.

15. Moore EK, Jelen BI, Giovannelli D, Raanan H, Falkowski PG:
*   **Metal availability and the expanding network of microbial metabolisms in the Archaean eon**. *Nat Geosci* 2017, **10**: 629−636.
Moore et al. integrate geochemical data and enzymatic information across the tree of life to demonstrate the strong coupling between the availability of metal ions in the Archean ocean with the metal usage in modern enzymes, highlighting metabolic pathway innovations that reshaped the major biogeochemical electrical circuits on geological time scales.

16. Martin W, Russell MJ: **On the origin of biochemistry at an alkaline hydrothermal vent**. *Philos Trans R Soc London B Biol Sci* 2007, **362**:1887−1926.

17. Srinivasan V, Morowitz HJ: **The canonical network of autotrophic intermediary metabolism: minimal metabolome of a reductive chemoautotroph**. *Biol Bull* April 2009, **216**:126−130.

18. Braakman R, Smith E: **Metabolic evolution of a deep-branching hyperthermophilic chemoautotrophic bacterium**. *PLoS One* 2014, https://doi.org/10.1371/journal.pone.0087950.

19. Braakman R, Smith E: **The emergence and early evolution of biological carbon-fixation**. *PLoS Comput Biol* 2012, https://doi.org/10.1371/journal.pcbi.1002455.

20. Kacar B, Hanson-Smith V, Adam ZR, Boekelheide N: **Constraining the timing of the great oxidation event within the rubisco phylogenetic tree**. *Geobiology* 2017, **15**:628−640.

21. Williams TA, Foster PG, Cox CJ, Embley TM: **An archaeal origin of eukaryotes supports only two primary domains of life**. *Nature* 2013, **504**:231−236.

22. Weiss MC, Sousa FL, Mrnjavac N, Neukirchen S, Roettger M,
* * Nelson-sathi S, Martin WF: **The physiology and habitat of the last universal common ancestor**. *Nat Microbiol* 2016, **1**: 1−8.
The authors use a phylogenomic approach to estimate the metabolic repertoire of LUCA, suggesting that LUCA contained at-least 355 protein families. This study suggests that LUCA was complex and highly evolved, including genes involved in carbon fixation (WL-pathway) and nitrogen fixation, and indicates that LUCA lived in thermophilic environments with sources of iron, sulfur and hydrogen gas, consistent with proposed ancient hydrothermal vent systems.

23. Gogarten JP, Deamer D: **Is LUCA a thermophilic progenote?** *Nat Microbiol* 2016, **1**:16229.

24. Weiss MC, Neukirchen S, Roettger M, Mrnjavac N, Nelson-Sathi S, Martin WF, Sousa FL: **Reply to "Is LUCA a thermophilic progenote?"**. *Nat Microbiol* 2016, **1**:16230.

25. Wachtershauser G: **Evolution of the first metabolic cycles**. *Proc Natl Acad Sci* 1990, **87**:200–204.

26. Wieczorek R, Adamala K, Gasperi T, Policelli F, Stano P: **Small and random peptides: an unexplored reservoir of potentially functional primitive organocatalysts. The case of seryl-histidine**. *Life* 2017, **7**:19.

27. Ruiz-Mirazo K, Briones C, de la Escosura A: **Prebiotic systems chemistry: new perspectives for the origins of life**. *Chem Rev* 2014, **114**:285–366.

28. Laurino P, Tawfik DS: **Spontaneous emergence of S-Adenosylmethionine and the evolution of methylation**. *Angew Chemie - Int Ed* 2017, **56**:343–345.

29. Keller MA, Zylstra A, Castro C, Turchyn AV, Griffin JL, Ralser M: **Conditional iron and pH-dependent activity of a non-enzymatic glycolysis and pentose phosphate pathway**. *Sci Adv* 2016, **2**, e1501235.

30. 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.

31. Messner CB, Driscoll PC, Piedrafita G, De Volder MFL, Ralser M: **Nonenzymatic gluconeogenesis-like formation of fructose 1,6-bisphosphate in ice**. *Proc Natl Acad Sci* 2017, **114**: 7403–7407.

32. Keller MA, Kampjut D, Harrison SA, Ralser M: **Sulfate radicals**
    * * **enable a non-enzymatic Krebs cycle precursor**. *Nat Ecol Evol* 2017, **1**:83.
    This study used high-throughput metabolomics to screen for potential catalysts that can interconvert compounds in the TCA cycle, finding that iron sulfides and sulfate radicals could serve as catalysts for reactions that span the TCA cycle, glyoxylate shunt and the succinate semialdehyde pathway.

33. 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.

34. Springsteen G, Yerabolu JR, Nelson J, Rhea CJ,
    * Krishnamurthy R: **Linked cycles of oxidative decarboxylation of glyoxylate as protometabolic analogs of the citric acid cycle**. *Nat Commun* 2018, **9**:91.
    This paper demonstrates that glyoxylate and hydrogen peroxide can fuel reaction cycles capable of producing organic molecules (including malate and oxaloacetate) typically observed in the TCA cycle, suggesting that protometabolism may have utilized pathways different than the current ones.

35. Keller MA, Piedrafita G, Ralser M: **The widespread role of non-enzymatic reactions in cellular metabolism**. *Curr Opin Biotechnol* 2015, **34**:153–161.

36. King GA: **Evolution of the coenzymes**. *Biosystems* 1980, **13**: 23–45.

37. Segré D, Ben-Eli D, Deamer DW, Lancet D: **The lipid world**. *Orig Life Evol Biosph* 2001, **31**:119–145.

38. Barenholz U, Davidi D, Reznik E, Bar-On Y, Antonovsky N, Noor E, Milo R: **Design principles of autocatalytic cycles constrain enzyme kinetics and force low substrate saturation at flux branch points**. *Elife* 2017, **6**:1–32.

39. Guseva E, Zuckermann RN, Dill KA: **Foldamer hypothesis for the growth and sequence differentiation of prebiotic polymers**. *Proc Natl Acad Sci* 2017, **114**:E7460–E7468.

40. Kinsler G, Sinai S, Lee NK, Nowak MA: **Prebiotic selection for motifs in a model of template-free elongation of polymers within compartments**. *PLoS One* 2017, **12**, e0180208.

41. Semenov SN, Kraft LJ, Ainla A, Zhao M, Baghbanzadeh M,
    * * Campbell VE, Kang K, Fox JM, Whitesides GM: **Autocatalytic, bistable, oscillatory networks of biologically relevant organic reactions**. *Nature* 2016, **537**:656–660.
    This paper demonstrats that a chemical network of small molecules and thioesters is capable of autocatalyic and oscillatory dynamics, suggesting that prebiotic chemical networks may have been capable of

42. Baum DA, Vetsigian K: **An experimental framework for generating evolvable chemical systems in the laboratory**. *Orig Life Evol Biosph* 2016, https://doi.org/10.1007/s11084-016-9526-x.

43. Hartman H: **Speculations on the origin and evolution of metabolism**. *J Mol Evol* 1975, **4**:359–370.

44. Morowitz HJ, Kostelnik JD, Yang J, Cody GD: **The origin of intermediary metabolism**. *Proc Natl Acad Sci U S A* 2000, **97**: 7704–7708.

45. Goldford JE, Hartman H, Smith TF, Segrè D: **Remnants of an**
    * * **ancient metabolism without phosphate**. *Cell* 2017, **168**. 1126–1134.e9.
    Using network-based computational methods, this paper shows that a large portion of core metabolism is reachable without the availability of phosphate. Analyses of energetic constraints in a metabolism without phosphate suggest that thioester coupling would have been important for the network to overcome thermodynamic bottlenecks.

46. Ebenhöh O, Handorf T, Heinrich R: **Structural analysis of expanding metabolic networks**. *Genome Inform* 2004, **15**: 35–45.

47. Handorf T, Ebenhöh O, Heinrich R: **Expanding metabolic networks: scopes of compounds, robustness, and evolution**. *J Mol Evol* 2005, **61**:498–512.

48. Raymond J, Segrè D: **The effect of oxygen on biochemical networks and the evolution of complex life**. *Science* 2006, **311**: 1764–1767.

49. Meringer M, Cleaves HJ: **Computational exploration of the chemical structure space of possible reverse tricarboxylic acid cycle constituents**. *Sci Rep* 2017, **7**:17540.

50. Smith E, Morowitz HJ: **Universality in intermediary metabolism**. *Proc Natl Acad Sci U S A* 2004, **101**:13168–13173.

51. Patel BH, Percivalle C, Ritson DJ, Duffy CD, Sutherland JD: **Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism**. *Nat Chem* 2015, **7**:301–307.

52. Sutherland JD: **Opinion: studies on the origin of life — the end of the beginning**. *Nat Rev Chem* 2017, **1**:12.

53. Andersen JL, Flamm C, Merkle D, Stadler PF: *A software pack-*
    * *age for chemically inspired graph transformation*. Cham: Springer; 2016:73–88, https://doi.org/10.1007/978-3-319-40530-8_5.
    This paper describes an open source software package to generate and analyze realistic chemical reaction networks based on graphs and rules representations of molecules and reactions, respectively.

54. Banzhaf W, Yamamoto L: *Artificial chemistries*. MIT Press; 2015.

55. Kauffman SA: *The origins of order: self-organization and selection in evolution*. Oxford University Press; 1993.

56. Riehl WJ, Krapivsky PL, Redner S, Segrè D: **Signatures of arithmetic simplicity in metabolic network architecture**. *PLoS Comput Biol* 2010, https://doi.org/10.1371/journal.pcbi.1000725.

57. Horowitz JM, England JL: **Spontaneous fine-tuning to envi-**
    * **ronment in many-species chemical reaction networks**. *Proc Natl Acad Sci USA* 2017, **114**, 201700617.
    In this study, the authors simulated far from equilibirum chemical networks and found that steady-states after long simulation times are enriched with configurations capable of harvesting energy from externally supplied forces, suggesting that driven chemical networks far from equilibirum may retain "memories" of their driving forces.

58. Deamer D, Weber AL: **Bioenergetics and life's origins**. *Cold Spring Harb Perspect Biol* 2010, **2**, a004929.

59. Chandru K, Gilbert A, Butch C, Aono M, Cleaves HJ: **The abiotic chemistry of thiolated acetate derivatives and the origin of life**. *Sci Rep* 2016, **6**:29883.

60. Hartman H: **Conjectures and reveries**. *Photosynth Res* 1992, **33**:171–176.

61. Bar-nun a, Hartman H: **Synthesis of organic compounds from carbon monoxide and water by UV photolysis**. *Orig Life* 1978, **9**:93–101.

both "memory" and dynamics typically associated with complex enzyme-based biological networks.

62. Bonfio C, Valer L, Scintilla S, Shah S, Evans DJ, Jin L,
** Szostak JW, Sasselov DD, Sutherland JD, Mansy SS: **UV-light-driven prebiotic synthesis of iron−sulfur clusters**. *Nat Chem* 2017, https://doi.org/10.1038/nchem.2817.
In this work, Mansy and collegues demonstrate that iron-sulfur clusters can be formed from mixtures of Fe^{3+} and thiolates (e.g. glutathione) upon irradiation using UV light. The presence of sulfide leads to both [2Fe−2S] and [4F−4S] complexes, which are components of enzymes proposed to be part of ancient metabolism.

63. Zhou R, Basu K, Hartman H, Matocha CJ, Sears SK, Vali H,
* Guzman MI: **Catalyzed synthesis of zinc clays by prebiotic central metabolites**. *Sci Rep* 2017, **7**:533.
This paper demonstrates that the synthesis of clay minerals, which are plausible prebiotic catalysts, can be catalyzed by the TCA cycle intermediate succinate.

64. Sousa FL, Thiergart T, Landan G, Nelson-Sathi S, Pereira I a C, Allen JF, Lane N, Martin WF: **Early bioenergetic evolution**. *Philos Trans R Soc Lond B Biol Sci* 2013, **368**, 20130088.

65. Martin WF, Thauer RK: **Energy in ancient metabolism**. *Cell* 2017, **168**:953−955.

66. Martin WF: **Hydrogen, metals, bifurcating electrons, and proton gradients: the early evolution of biological energy conservation**. *FEBS Lett* 2012, **586**:485−493.

67. Braakman R, Smith E: **The compositional and evolutionary logic of metabolism**. *Phys Biol* 2013, **10**:11001.

68. Markowitz VM, Chen I-MA, Palaniappan K, Chu K, Szeto E, Grechkin Y, Ratner A, Jacob B, Huang J, Williams P, *et al.*: **IMG: the Integrated Microbial Genomes database and comparative analysis system**. *Nucleic Acids Res* 2012, **40**:D115−D122.

69. Kanehisa M, Goto S: **KEGG: kyoto encyclopedia of genes and genomes**. *Nucleic Acids Res* 2000, **28**:27−30.

70. Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker BA, *et al.*: **PubChem substance and compound databases**. *Nucleic Acids Res* 2016, **44**: D1202−D1213.

71. Goldman AD, Bernhard TM, Dolzhenko E, Landweber LF: **LUCApedia: a database for the study of ancient life**. *Nucleic Acids Res* 2013, **41**:D1079−D1082.

72. Flamholz A, Noor E, Bar-Even A, Milo R: **EQuilibrator - the biochemical thermodynamics calculator**. *Nucleic Acids Res* 2012, **40**:770−775.

73. Rasmussen S, Constantinescu A, Svaneborg C: **Generating minimal living systems from non-living materials and increasing their evolutionary abilities**. *Philos Trans R Soc Lond B Biol Sci* 2016, **371**, 20150440.

74. Chang Y-C, Hu Z, Rachlin J, Anton BP, Kasif S, Roberts RJ, Steffen M: **COMBREX-DB: an experiment centered database of protein function: knowledge, predictions and knowledge gaps**. *Nucleic Acids Res* 2016, **44**:D330−D335.