

Supplementary Information for Metabolic homeostasis and growth in abiotic cells

Amir Akbari, Bernhard O. Palsson*

¹*Department of Bioengineering, University of California San Diego, La Jolla, CA 92093*

²*Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800
Lyngby, Denmark*

This PDF file includes:

Supplementary Text

Figs. S1–S6

Tables S1–S3

*Corresponding author.

Email address: palsson@ucsd.edu (Bernhard O. Palsson)

Model Description

We begin by describing the protocell model that we introduced in the main article to study the emergence of first metabolic cycles. The model we used in the study is similar to the one we previously considered to estimate the membrane potentials that can be generated in primitive cells with iron-sulfide membranes [1]. A primary purpose of this model is to simulate alkaline hydrothermal vent (AHV) conditions on the Early earth. Specifically, the protocell in this model is intended to represent AHV micropores and answer whether they could have harbored early carbon-fixing cycles at the origin of metabolism.

Our model is based on metabolic theories of life’s origins—those that are associated with the metabolism-first hypothesis [2–4]. These theories postulate that life originated from an ancient core metabolic network on the early Earth long before the advent of enzymes or genes. Although what this core network might have been like is unknown, phylogenetic studies have identified some of its characteristics. These studies indicate an anoxic, phosphate-free, and autotrophic network at the core of metabolism [5, 6]. We now have experimental evidence, suggesting that key parts of the core network, including the acetyl-CoA pathway [7–9], tricarboxylic acid cycle [10–13], glycolysis and pentose phosphate pathway [14, 15], amino-acid and pyrimidine-nucleobase biosynthesis pathways [16, 17], and gluconeogenesis [18] are thermodynamically favored, and they can proceed nonenzymatically. However, whether all or a subset of these nonenzymatic pathways could have spontaneously emerged and formed a self-sustaining metabolic network under plausible prebiotic conditions is yet to be demonstrated.

AHVs in the Hadean ocean would have been a suitable environment for first metabolic pathways to emerge. These environments could have provided steep proton and redox gradients to drive carbon metabolism [19, 20]. Moreover, they would have been rich in CO₂, reducing gases (*e.g.*, H₂ and H₂S), and transition metal minerals with catalytic and reducing properties [19]. Besides CO₂, simple C2 carbon sources may also have been present in AHVs. Acetate and glyoxylate are organic C2 carbon sources that could have played an important role at the origin of metabolism. They can be synthesized nonenzymatically from CO₂ and HCN—prebiotically plausible C1 carbon sources—through the Wood-Ljungdahl pathway [7, 8, 21] and oligomerization reactions [22]. Glyoxylate can also be synthesized from CO—an intermediate of the Wood-Ljungdahl pathway—under hydrothermal conditions using transition metal catalysts without aqueous CN[−] through the intermediacy of glycolate [23, 24]. These are simple linear pathways that could have operated without enzymes in unconfined environments, such as oceanic crusts (Fig. S1A). However, because of the concentration problem, complex carbon-fixing cycles likely required a confined environment to emerge, stabilize, self-amplify, and evolve. The concentration problem arises as a result of the restrictions associated with nonenzymatic origins of metabolism. Before enzymes, inorganic catalysts (*e.g.*, transition metals) could have promoted early metabolic reactions. However, because of the poor efficiency of inorganic catalysts, a stable flux through early metabolic pathways could not have been sustained. As a result, metabolic fluxes could have hardly counterbalanced diffusive losses in unconfined environments to concentrate the intermediates of these pathways beyond a minimum physical limit [1, 25, 26]. Although the concentration problem affects both linear and cyclic pathways, it would have limited the feasibility of cyclic pathways more [27]. Hence, it is plausible to assume that first carbon-fixing cycles emerged in confined regions of AHVs in the primitive ocean (Fig. S1A).

Given the importance of confinement in the emergence and evolution of complex

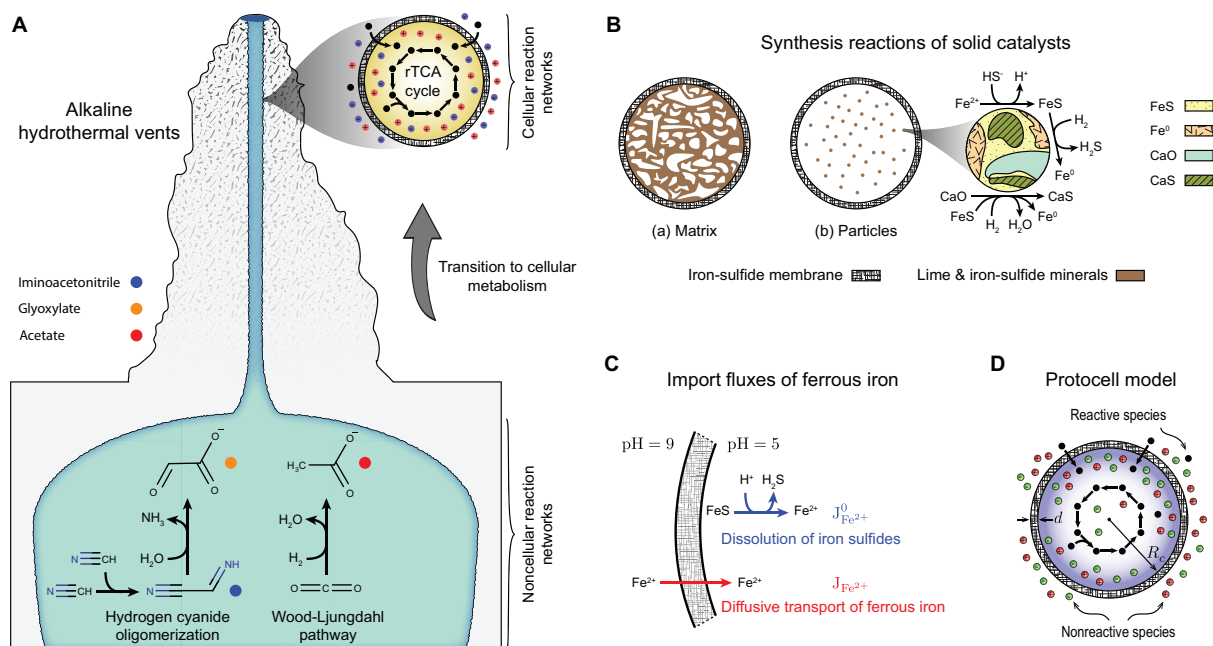


Figure S1: Cellular origins of metabolism. (A) AHVs could have harbored earliest cellular structures capable of sustaining carbon-fixing cycles on the early Earth. C2 organic compounds, such as acetate and glyoxylate, could have been synthesized in deep seabeds from inorganic C1 carbon sources, such as CO₂ and HCN. Acetate and glyoxylate could have been produced nonenzymatically through the Wood-Ljungdahl pathway [7, 8, 21] and HCN-oligomerization reactions [22] under prebiotic conditions. Because these pathways are linear and simple, they could have spontaneously arisen in oceanic crusts in unconfined environments. However, autocatalytic carbon-fixing pathways would likely have required a controlled environment to operate and self-amplify, necessitating a transition to cellular metabolism. Thin-walled vent micropores made of iron sulfides could have furnished a suitable environment for first carbon-fixing cycles to emerge, providing a barrier to facilitate primitive forms of energy coupling and contain the organic products of the foregoing cycles. (B) The protometabolic network in Fig. 1C is driven by inorganic catalysts and reducing agents, such as FeS and Fe⁰. Under hydrothermal conditions, mixtures of FeS and Fe⁰ can be synthesized through the interaction of Fe²⁺ and HS⁻ [28]. As a result of these interactions in the protocell model, assemblages of FeS-Fe⁰ precipitate on lime minerals, which would have been present in hydrothermal-vent environments. The resulting mineral mixtures are distributed inside the protocell either as a continuous matrix (a) or dispersed particles (b). (C) Ferrous iron that participates in metabolic reactions and mineral synthesis enters the protocell from outside by passive diffusion. The dissolution of the iron-sulfide membrane on the acidic side also yields ferrous iron, which contributes to its import flux. The dissolution rate of iron-sulfide membrane and the membrane transport rate of ferrous iron are denoted $J_{\text{Fe}^{2+}}^0$ and $J_{\text{Fe}^{2+}}$, respectively. (D) Protocell model, simulating the operation of first carbon-fixing cycles in AHVs micropores.

metabolic networks, transition to cellular metabolism from very early stages of chemical self-organization might have been inevitable. In the context of the metabolism-first hypothesis, abiotic cells that contained first metabolic cycles must have been made of purely inorganic materials, and all metabolic reactions in these cells must have been promoted by inorganic catalysts without relying on enzymes. In the absence of enzymes, the operation of abiotic cells would have been restricted in several ways. For example, the import of energy or carbon sources and export of products should have occurred unselectively through passive diffusion. Carbon fixation should have been driven by primitive energy-coupling means without electron bifurcation or proton pumps. Moreover, reactant concentrations and metabolic fluxes would have been extremely low due to the poor efficiency of inorganic catalysts, limiting the feasibility of primitive pathways, especially those with multiple thermodynamic bottlenecks.

AHV micropores could have served as the earliest cellular structures in which first carbon-fixing cycles emerged. However, whether a self-sustaining carbon-fixing cycle can operate on purely inorganic precursors (*i.e.*, energy sources, reducing compounds, and catalysts) alone in such inorganic cells is unclear. To answer this question, we developed a protocell model of AHV micropores. Our goal is to demonstrate the feasibility of such protocells despite the highly restrictive physico-chemical constraints of the early Earth.

Since fully autotrophic pathways would have been challenging to operate without enzymes, we study the feasibility of partially autotrophic carbon fixation by including C2 carbon sources, such as acetate and glyoxylate, in the food set. As previously stated, these could have been synthesized from inorganic C1 carbon sources in oceanic crusts of the primitive Earth under prebiotic conditions. In particular, because the partial pressure of CO₂ in hydrothermal systems of the early Earth would have been significantly higher than those in modern oceans [19, 20, 29], acetate could have been synthesized in high concentrations through the Wood-Ljungdahl pathway. Acetate and glyoxylate, which were present in AHV environments, then would have been transported into primitive cells along with CO₂ by passive diffusion (see Fig. 1B).

We consider an idealized protocell model of AHV micropores. The protocell is a sphere of radius R_c with a semi-permeable membrane of thickness d (Fig. S1D). The membrane is primarily made of iron sulfides possibly mixed with other minerals. Under hydrothermal conditions, similar cellular structures made of iron monosulphide can form when ferrous iron is mixed with alkaline fluids that are rich in HS⁻ [30]. We assume that a simple hypothetical thiol HS-R (*e.g.*, ethane thiol [19]) existed in primitive AHVs that drove the thioesterification steps of first metabolic networks. The interior of the protocell is filled with acidic ocean water (pH = 5), while its outer membrane is exposed to alkaline hydrothermal fluids (pH = 9) (see Fig. 1A,B). Because of this pH gradient, a positive membrane potential $\Delta\psi$ develops across the membrane [1]. Hereafter, we refer to the acidic interior of the protocell as the intracellular environment (denoted ‘In’ in Fig. 2A) and alkaline hydrothermal fluids outside the protocell as the extracellular environment (denoted ‘Out’ in Fig. 2A). Note that the pH gradient in our model exists only across the membrane. We assume that all the species in the model (including H⁺) are uniformly distributed inside the protocell; thus, there are no pH or concentration gradients within the intracellular environment. Carbon sources, energy sources, and all the organic intermediates of first metabolic cycles in our model are negatively charged. Thus, positive membrane potentials facilitate the import of carbon and energy sources, while alleviating the dissipation of the organic intermediates. The protocell imports carbon sources (acetate, glyoxylate, and CO₂), reducing gases (H₂ and H₂S), and thiols

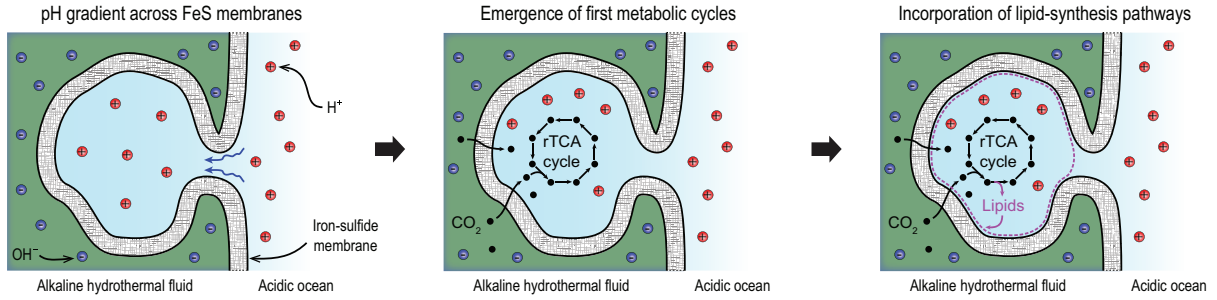


Figure S2: **Maintaining pH gradient across the membranes of earliest cellular structures.** As suggested in previous studies [31], thin-walled AHV cavities made of iron sulfides could have served as inorganic cells, harboring the first metabolic networks at the origin of metabolism. The inorganic membranes of these cellular structures could have separated alkaline hydrothermal fluids from acidic ocean waters. Early metabolic networks could have incorporated lipid-synthesis pathways at later stages once their precursors had been produced in sufficiently large concentrations. Single-chain amphiphiles produced by these pathways could have self-assembled into lipid membranes [32], replacing iron-sulfide membranes.

(HS-R) from outside. It has no specialized ion channels or transporters, so the only mechanism of membrane transport is unselective passive diffusion subject to an electric field.

Except carbon-fixing steps, all metabolic reactions are catalyzed by transition metals in aqueous (Fe^{2+} and Cr^{3+}) or solid (Fe^0 and FeS) phase. They are dispersed inside the protocell in an acidic environment. Beside iron sulfides, lime minerals would likely have been available in AHVs of the primitive ocean. Accordingly, we assume that the protocell is filled with lime minerals, which either form a continuous matrix (Fig. S1B(a)) or dispersed particles (Fig. S1B(a)). Assemblages of FeS-Fe^0 precipitate on these lime minerals upon interaction of Fe^{2+} and HS^- ions (Fig. S1B). Interestingly, these assemblages exhibit superior catalytic efficiencies to pure FeS in promoting prebiotically important reduction, reductive carboxylation, and reductive amination reactions [28]. In our protocell model, several steps of the protometabolic network that we examined in this work are catalyzed by these mineral mixtures (see “Kinetic Constraints”). On the other hand, the reductive carboxylation steps of the network are driven by protoferredoxins. We assume that the reduced and oxidized forms of these protoferredoxins, respectively denoted FD_{rd} and FD_{ox} , constitute a fraction of the iron-sulfide membrane. The energy released by the redox conversion $\text{FD}_{\text{rd}} \rightarrow \text{FD}_{\text{ox}}$ drives the reductive carboxylation steps on the inner surface of the membrane in an acidic environment. On the outer surface of the membrane, H_2 reduces FD_{ox} in an alkaline environment to regenerate FD_{rd} . Note that reductive carboxylation and FD_{rd} -regeneration reactions are more thermodynamically favorable in acidic and alkaline environments, respectively.

Acidic environments can dissolve iron sulfides. In our protocell model, the rate at which Fe^{2+} is released into the protocell as a result of the dissolution of the iron-sulfide membrane $J_{\text{Fe}^{2+}}^0$ constitute part of the import flux of Fe^{2+} . The other part $J_{\text{Fe}^{2+}}$ is due to its diffusive transport from the extracellular environment (Fig. S1C). The ferrous iron entering the protocell through these import fluxes interact with HS^- ions to form the foregoing FeS-Fe^0 assemblages. Other metal ions involved in metabolic reactions (*e.g.*, Cr^{3+}) are also imported from the extracellular environment similarly to Fe^{2+} .

Note that the idealized protocell model shown in Fig. S1D represents both closed and open AHV micropores. It is intended to capture important physico-chemical characteristics of hydrothermal systems that played a role in the emergence of first metabolic networks on the early Earth. In the main article, we provided a schematic representation of closed micropores (Fig. 1A, inset) as abiotic cells that could have harbored first carbon-fixing cycles at the origin of metabolism. However, open cavities in AHVs, such as that shown in Fig. S2, have also been considered in previous studies on the origins of metabolism [31]. These cavities could have maintained a pH gradient across their inorganic membranes to drive carbon fixation. Once stabilized, early carbon-fixing cycles could have produced the precursors of fatty acids, incorporating lipid-synthesis pathways [1, 32]. Lipid membranes could have spontaneously emerged in such cavities from the self-assembly of single-chain amphiphiles produced by these pathways [32, 33] and replaced iron-sulfide membranes at later stages (see Fig. S2).

First Carbon-Fixing Cycles at the Origin of Metabolism

According to metabolic theories of life's origins, life emerged as a consequence of the natural tendency of a self-organized geochemical reaction network to dissipate free energy [19, 34]. This network would have been thermodynamically favored and kinetically assisted by mineral catalysts. Thus, it could have occurred spontaneously and operated without enzymes. Such a protometabolic network would not have been fundamentally different than biochemical reactions at the core of biological metabolism in substrates, topology, catalysts, or energy-coupling mechanisms [35]. Given the highly interconnected structure of reaction networks leading up to extant biology and their nonlinear dynamics, a major change in these characteristics would likely have resulted in a catastrophic event and a sudden transition to a chaotic and nonfunctional system. There has been no evidence of such drastic transitions in the evolutionary history of metabolism at least as far back as 3.5 billion years ago, and they were unlikely to have had happened before [34]. Therefore, it is reasonable to search for metabolic origins that are congruent with modern metabolism.

All essential metabolic processes of life hinge on carbon fixation in some form. It is an important biosynthesis route for building complex molecules from simpler ones, and it would likely have been the primary function of the earliest metabolic pathways. Thus, a long-standing question in prebiotic chemistry has been whether there were primitive reaction networks at the origin of metabolism that could have fixed CO_2 —one of the most abundant inorganic carbon sources on the early Earth [36]. Using life as a guide to its origins, one may search for clues to the underlying chemistry of these reaction networks in the structure of CO_2 -fixing pathways in modern autotrophs. From the six known CO_2 -fixing pathways, the acetyl-CoA pathway and the reductive tricarboxylic acid (rTCA) cycle are ancient pathways that could provide insight into what the first CO_2 -fixing networks were like [3, 27, 36]. The acetyl-CoA pathway is present in both bacteria and archaea [31], while the full rTCA cycle is only found in bacteria [34]. They both could have played a significant role in the emergence of metabolic pathways. However, whether the complete rTCA cycle or its other forms could represent the underlying chemistry of the first carbon-fixing pathways has not been established. Nevertheless, a hybrid network comprising the acetyl-CoA pathway and the horseshoe rTCA cycle (*i.e.*, pyruvate to α -ketoglutarate) [34], a pyrite-pulled rTCA-cycle analogue [36], and a pyrophosphate driven biology-like rTCA cycle [37, 38] have been proposed as plausible models of prebiotic carbon fixation.

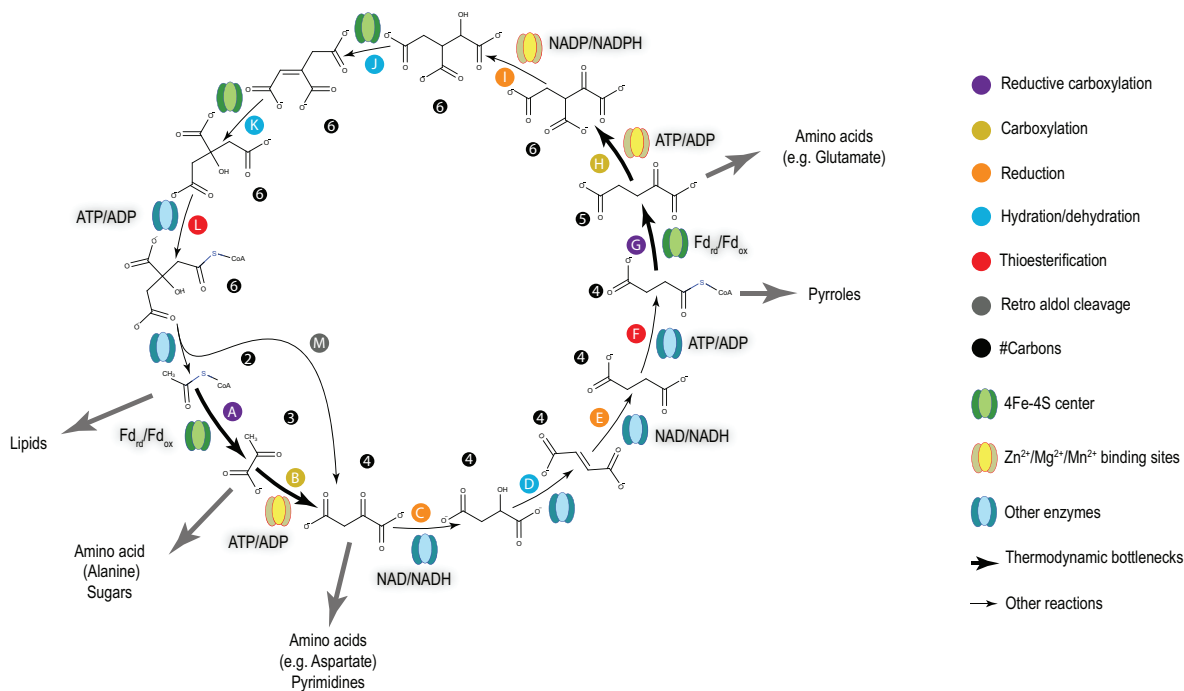


Figure S3: **The enzymatic rTCA cycle.** This ancient pathway contains the five universal intermediates from which all the building blocks of life, including lipids, amino acids, sugars, pyrimidines, and pyrroles are synthesized. The cofactors driving reduction or energy-demanding steps are highlighted in gray.

Although the antiquity of the complete rTCA cycle is not fully supported by the phylogeny of modern bacteria and archaea, it is still plausible to assume that some variant of it was the first CO_2 -fixing pathway, or at least played a significant role in one, for several reasons [19, 34, 36, 37]: (i) It is a truly autocatalytic cycle, so it can self-amplify and undergo evolutionary selection provided its carbon and energy sources are readily available in the environment. (ii) It is anoxic and fully autotrophic. Once it has been started and stabilized, it fixes four CO_2 molecules with every turn without needing other carbon sources. (iii) Except cofactors, all its intermediates are phosphate free. (iv) It contains the five universal intermediates from which all the building blocks of life (*i.e.*, amino acids, fatty acids, and nucleotides) are biologically synthesized. (v) So far, transition metals have been shown to promote seven out of thirteen steps of the cycle [10, 39]. (vi) Several enzymes that catalyze the reactions of the cycle have iron-sulfur clusters or transition-metal binding sites (see Fig. S3).

The enzymatic rTCA cycle may be an ancient biochemical reaction network. It could have been present in the Last Universal Common Ancestor, but part of it might have been lost in archaea. Nevertheless, the biological rTCA cycle is not an accurate representation of prebiotic CO_2 -fixing cycles: First, several reactions in the enzymatic rTCA cycle are driven by phosphate-containing cofactors, such as ATP and NADH. However, some of these steps can also be promoted nonenzymatically by transition metals [10, 11]. These metal-catalyzed reactions might have preceded their biological counterparts before the emergence of enzymes. Second, the thioesterification steps of the rTCA cycle are facilitated by ATP and coenzyme A. Both contain multiple phosphate groups and are products of a long evolutionary process [19], so they could not have played any role at the origin of metabolism. However, thioesterification can also be accomplished by simple,

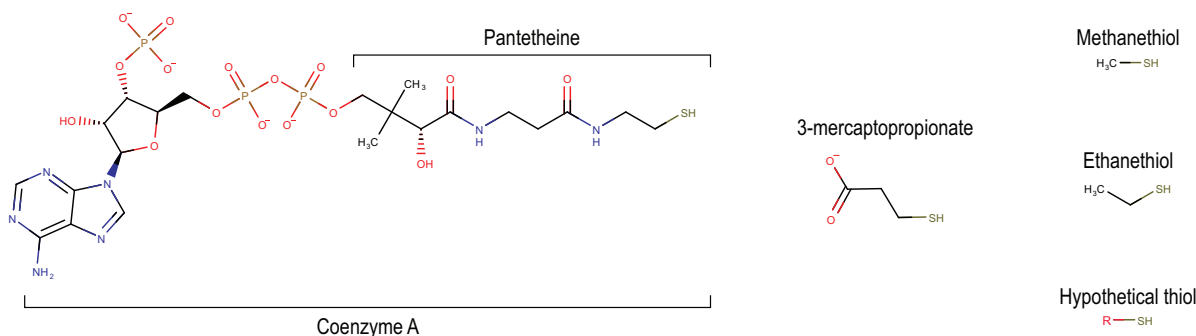


Figure S4: **Thiols participating in enzymatic and nonenzymatic thioesterification reactions.** The thioesterification steps of the protometabolic network in this work are driven by a simple, hypothetical thiol that could have been present in the hydrothermal vents of the Hadean ocean with R a hypothetical substituent.

inorganic thiols with acids or metal ions as catalysts [40, 41]. Third, the carbon-fixation steps of the rTCA cycle are highly endergonic. These thermodynamic bottlenecks must be overcome to achieve a self-sustaining complete-cycle flux state. Modern organisms (*e.g.*, thermophiles) can sustain a complete rTCA cycle thanks to their efficient and evolutionary optimized enzymes [42]. However, in the absence of enzymes, the alternative inorganic catalysts would not have been efficient enough to diminish the kinetic barriers associated with the thermodynamic bottlenecks. Consequently, maintaining a fully cyclic flux in early metabolic pathways without enzymes would have been extremely challenging. Hence, energy-demanding reactions should have been coupled to inorganic energy sources by prebiotic mechanisms.

Addressing the foregoing three problems, we derived a protometabolic network from the rTCA cycle (Fig. 1C) that is compatible with the AHV conditions and could have operated prebiotically. Firstly, we replaced all enzymatic reactions that can be promoted by transition metals with their nonenzymatic counterpart. Specifically, in our protometabolic network, hydration and dehydration steps are catalyzed by Fe²⁺ and Cr³⁺, while reduction steps are promoted by Fe⁰. As discussed in section “Model Description”, Fe²⁺ and Cr³⁺ are imported into the protocell from the extracellular environment, and Fe⁰ is synthesized through the interaction of Fe²⁺ with HS⁻.

Furthermore, coenzyme A is replaced by a hypothetical thiol HS-R in the thioesterification steps of the protometabolic network. Such simple, inorganic thiols could have been synthesized in the sulphide-rich AHVs of the primitive ocean. Previous studies on core reaction networks preceding phosphate-based metabolism considered simpler phosphate-free variants of coenzyme A, such as pantetheine, 3-mercaptopropionate, methanethiol, and ethanethiol as prebiotic thiols involved in thioester chemistry (see Fig. S4) [5, 19]. In this work, the chemical structure of the substituent R is unspecified. Accordingly, to estimate the transformed Gibbs energy of thioesterification reactions, the transformed Gibbs energy of formation of this hypothetical thiol and its corresponding thioesters are provided as given parameters of the model (see “Abiotic Thioesterification”). However, to reduce the number of model parameters, its diffusivity coefficient is approximated based on the van der Waals radius of ethanethiol (see “Mass Balance Constraints and Maxwell’s first law”).

Finally, we considered primitive energy-coupling mechanisms to overcome the energy barriers associated with carbon fixation. In the enzymatic rTCA cycle, there are

four carbon-carbon bond-forming reactions: two reductive carboxylation steps driven by ferredoxin and two neutral carboxylation steps coupled to ATP hydrolysis (see Fig. S3). Among these, only the reductive carboxylation of acetate to pyruvate has been experimentally shown to occur without enzymes using iron-containing minerals (*e.g.*, greigite, magnetite, and awaruite) as catalysts and hydrogen as the reducing agent [8]. Therefore, it is plausible to assume that, in early metabolic cycles, both reductive carboxylation steps were promoted by primitive redox couples, possibly comprising iron, nickel, and sulfur with a similar crystal structure to Fe-S clusters in modern ferredoxin [19]. However, the neutral carboxylation steps have not been reproduced nonenzymatically in laboratory experiments so far, and they were unlikely to have occurred under prebiotic conditions without evolutionarily optimized enzymes or cofactors, such as ATP.

In the remainder of this section, we elaborate further on the four carbon-fixation steps discussed above. First, consider the reductive carboxylation steps. In our protocell model, the two reductive carboxylation steps are driven by protoferredoxins. In extant biology, these steps are accomplished by ferredoxins—enzymes containing Fe_4S_4 thiocubane centers that are ligated to proteins by four cysteine residues. Strong metal-ligand bonds stabilize the crystal structure of Fe-S clusters in these enzymes [43]. The crystal structure in turn determines the reduction potential and catalytic capacity of Fe-S clusters, although other factors might also play a role [44]. Moreover, the ligands mitigate oxidative degradation of Fe-S clusters in the presence of oxidizing agents, such as Brønsted acids or hydrogen peroxide [45]. What forms of protoferredoxins could have been synthesized under prebiotic conditions? In the absence of proteins, could naturally occurring hydrothermal redox couples have instead driven the reductive carboxylation steps of early metabolic networks?

Although hydrothermal redox couples such as mackinawite/greigite and pyrrhotite/pyrite can promote carbon-fixation reactions similarly to those in acetogens [8] and methanogens [46], whether they could have played a significant role at early evolutionary stages of metabolism is in question. Pyrite is a highly stable form of iron sulfides with a complex crystal structure, making it hard to reduce. Once it is produced, it can no longer participate in abiotic metabolic reactions, nor can it be regenerated by reducing agents, such as H_2 or H_2S [19]. Regenerability is an important feature of extant redox couples (*e.g.*, NAD/NADH and FAD/FADH), underlying the autonomy and robustness of modern organisms. In contrast, mackinawite and greigite are electroconductive with readily switchable iron valence; thus, in principle, mackinawite can be regenerated by reducing greigite [19]. Interestingly, the crystal structure of greigite resembles that of iron-sulfur centers in modern ferredoxins, so it is expected to exhibit similar catalytic characteristics [19]. However, both amorphous FeS and greigite are prone to oxidative degradation. That is, they are both oxidized to pyrite if exposed to H_2S and H^+ [47]. This last point is particularly problematic in the context of autotrophic origins of metabolism: On the one hand, acidic environments thermodynamically favor reductive carboxylation. On the other hand, they degrade the crystal structure of mackinawite and greigite. Therefore, preserving the structural integrity and catalytic capacity of hydrothermal minerals through oxidation and reduction cycles under conditions that allowed abiotic metabolism to spontaneously emerge would have been challenging on the early Earth, necessitating the formation of stabilizing metal-ligand bonds.

Comparative studies of amino-acid sequences of ferredoxins across diverse organismal groups suggest that these archaic proteins have evolved through a sequence-doubling process, whereby their stability and catalytic activity were improved in a stepwise manner

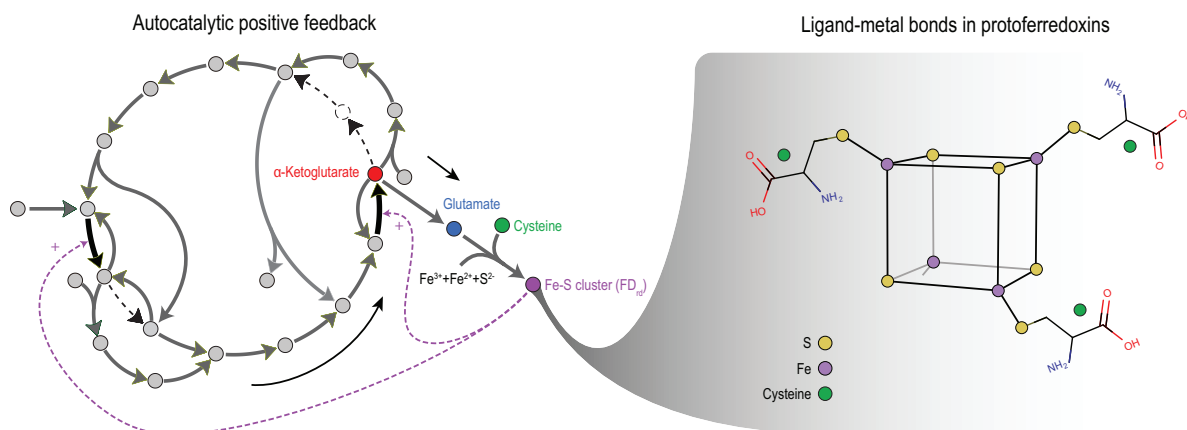


Figure S5: Autocatalytic positive feedback underlying the evolutionary selection of protoferredoxins and amino-acid synthesis pathways in first metabolic cycles. Once their synthesis pathways had been activated, ancient amino acids, such as cysteine, could interact with the products of reductive carboxylation reactions, to chelate Fe-S clusters in the presence of iron and sulfur ions, synthesizing protoferredoxins. The crystal structure of these protoferredoxins is more stable and their catalytic efficiency is higher than those of naturally occurring hydrothermal minerals, such as mackinawite and greigite. Protoferredoxins in turn can accelerate reductive carboxylation reactions more efficiently than alternative mineral catalysts. Thus, the formation of protoferredoxins could have generated positive feedbacks by amplifying the reductive carboxylation steps of first metabolic cycles—a nongenetic mechanism underlying their evolutionary selection.

[48]. This process started from a four amino-acid peptide and continued by repeating longer and longer sequences of amino acids. Extrapolating this process to the most ancient ferredoxins, it follows that the earliest redox-active iron-sulfur assemblages at the origin of metabolism, which we refer to as protoferredoxins, were stabilized by primitive, short ligands that would have been present on the early Earth. The organic intermediates of the rTCA cycle (*e.g.*, pyruvate) [47] and prebiotic thiolates [19] are examples of such ligands that could have stabilized primitive Fe-S clusters. Ancient peptides could also have played the same role. As has been experimentally demonstrated, in the presence of iron and sulfur ions, micromolar concentrations of ancient peptides (*e.g.*, cysteine and glutathione) can chelate redox-active Fe-S clusters [45, 49]. Their constituent amino acids in turn can be synthesized nonenzymatically through metal-catalyzed reductive amination or transamination of α -keto acids of the rTCA and glyoxylate cycle under prebiotic conditions [11, 16, 50, 51]. The formation of these amino acids hinge on abiotic carbon-fixing cycles that can produce their precursors in sufficiently large concentrations. To avoid the circular interdependence between amino-acid synthesis and carbon fixation, reductive amination and transamination reactions are assumed not to be activated in the earliest carbon-fixing cycles represented by our protometabolic network, although they could have become activated at later stages. Accordingly, protoferredoxins in our model are synthesized from either the intermediates of the rTCA cycle or primitive thiolates. These protoferredoxins are mixed with other hydrothermal redox couples (*e.g.*, mackinawite/greigite), and they constitute a fraction of the iron-sulfide membrane of the protocell (see Fig. 1B). As previously stated, these hydrothermal minerals could have promoted the reductive carboxylation steps of early metabolic networks. However, in the

protometabolic network, these steps are assumed to be driven only by protoferredoxins for simplicity.

Autocatalysis could have played an important role in the evolution of first metabolic cycles, such as our protometabolic network, before the emergence of genetic codes [52]. Once these cycles had stabilized, they could produce their intermediates like α -ketoacids in large concentrations due to their self-amplifying characteristics. Sufficiently large concentrations of α -ketoacids then could have activated amino-acid synthesis pathways in the presence of nitrogen sources. As stated above, ancient amino acids, such as cysteine and glutamate, could have facilitated the formation of more stable and efficient protoferredoxins than those chelated by the intermediates of the rTCA cycle or primitive thiolates. These protoferredoxins in turn could have generated positive feedback, amplifying reductive carboxylation reactions (Fig. S5). Such positive feedback likely underlay the evolutionary selection of amino-acid synthesis pathways and iron-sulfide crystal structures that were optimal for the operation of first carbon-fixing cycles before genes.

Next, consider the neutral carboxylation steps. Without neutral carboxylation, some intermediates of the enzymatic rTCA cycle cannot be accessed. In the absence of energetic cofactors of extant biology, abiotic carboxylation in primitive pathways would have been highly endergonic and challenging to run nonenzymatically. Thus, a natural question is whether nonenzymatic carbon-fixing pathways similar to the rTCA cycle could have operated prebiotically using inorganic energy sources and reducing agents—pathways that were fully cyclic and contained most or all the intermediates of the enzymatic rTCA cycle. Examples of such pathways were examined in the glyoxylate scenario [53]. The pathways considered in this scenario rely on the chemistry of HCN. Glyoxylate is the central intermediate of these pathways, realizing aldol condensations that connect simple C1 carbon sources to more complex intermediates of the rTCA cycle. A recent theoretical analysis showed that a phosphate-free core metabolic network enriched with major carbon-fixing cycles can emerge from simple carbon sources using network expansion algorithms [5]. Interestingly, glyoxylate and pyruvate were among the largest branching points of this core network, suggesting their prominent role in the earliest carbon-fixing pathways. Subsequent experiments confirmed that an abiotic reaction network, comprising a large portion of the TCA and glyoxylate cycle, can be generated using glyoxylate and pyruvate [11]. This network was promoted by Fe^{2+} and produced nine of the thirteen intermediates of the rTCA cycle.

The foregoing studies highlight the importance of glyoxylate for prebiotic production of the intermediates of the rTCA cycle—primordial hubs from which the earliest building blocks of life were likely synthesized. Thus, we searched for nonenzymatic and fully cyclic carbon-fixing pathways among the reactions of the core metabolic network that have been previously demonstrated to proceed spontaneously only using transition metals as catalysts [10, 11]. We considered pathways that could bypass the neutral carboxylation steps to connect simple carbon sources, such as CO_2 or acetate, to the intermediates of the rTCA cycle if glyoxylate was included in the food set. We found that both neutral carboxylation steps of the enzymatic rTCA cycle can be bypassed through aldol reactions involving glyoxylate followed by oxidative decarboxylation and thioester hydrolysis steps. Specifically, the carboxylation of pyruvate to oxaloacetate (Fig. S3, step B) and of α -ketoglutarate to oxalosuccinate (Fig. S3, step H) can be bypassed through aldol reaction of glyoxylate with pyruvate (Fig. 1C, step O) and α -ketoglutarate (Fig. 1C, step R), respectively. These off-cycle reactions yield two intermediates that are not produced by the enzymatic rTCA cycle, namely hydroxyketoglutarate and oxalohydrox-

γ-glutarate [11]. These intermediates in turn can undergo an oxidative decarboxylation and a thioester hydrolysis to produce malate and isocitrate, respectively. Note that Fe^{2+} can also promote the direct synthesis of malate and isocitrate from hydroxyketoglutarate and oxalohydroxyglutarate without thioester hydrolysis steps if thiols are not present in the reaction environment [11]. However, it is likely that the foregoing bypass pathways occurred predominantly through thiol-mediated oxidative decarboxylation because the alkaline hydrothermal environments of the primitive ocean would have been rich in sulfide and thiols. Accordingly, we neglect the direct decarboxylation reactions in our protometabolic network for simplicity (see Fig. 1C).

Role of Phosphate at the Origin of Metabolism

Phosphate is a central component of biochemical transformations that are essential to all living systems, such as energy transduction and information storage. Phosphate and thioester chemistries are intimately connected in biology, and they are both involved in many bioenergetic processes. However, it is unclear when and how these processes were incorporated into protometabolic networks during their early evolutionary stages [34]. Phosphate, in particular, has long been considered as limiting compound in prebiotic systems due to its geological scarcity [54] and thus believed not to have been essential for the emergence of first metabolic network [5].

In our model, we adopted the hypothesis that the first carbon-fixing cycles at the origin of metabolism did not require phosphate-dependent energy-coupling mechanisms. As discussed in the main text, we reconstructed a phosphate-free protometabolic network, demonstrating its feasibility under AHV conditions of the early Earth. However, recent theoretical and experimental studies question the poor geochemical accessibility of phosphate, showing that ferrous iron present in the Hadean and Archean oceans could have significantly increase the solubility of phosphate minerals [55]. High concentrations of phosphate in the early ocean would have facilitated the formation primordial energy currencies, such as acetyl phosphate, connecting phosphate and thioester chemistries during the early evolution of metabolic pathways [56]. To what extent the presence of phosphate-containing cofactors could have affected the feasibility of first-carbon-fixing cycles in AHVs is an important question, which warrants further research in the future.

Abiotic Thioesterification

Thioesters play a prominent role in biochemistry, especially in the bioenergetics of metabolic networks [25]. Being coupled to ATP synthesis and energy-demanding condensation reactions [34], they are involved in both catabolic and anabolic processes. In the context of carbon assimilation, thioesters are high-energy and activated forms of carboxylic acids. Specifically, in the rTCA cycle, their synthesis precedes the carboxylation steps to facilitate energy-demanding C-C bond formations. In biological systems, thioesterification is coupled to ATP hydrolysis because it is an endergonic process. However, thioesterification can also be accomplished inorganically using Brønsted acids [40] or metal ions [41] as catalysts. As previously stated, thioesterification steps are driven by a hypothetical thiol HS-R in our protocell model. All the thioesterification steps of our protometabolic network can be represented by the following general reaction

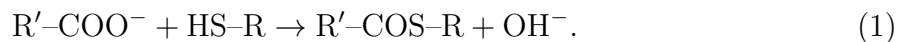


Table S1: Parameters used to estimate the Gibbs energy of formation of hypothetical thioesters in the protometabolic network.

Carboxylic acid	Identifier	R'	$\Delta\Delta_f G_{R'}^{\circ\circ}$ (kJ/mol)
Acetic acid	ac	CH3	-161.5
Succinic acid	suc	COOCH2CH2	-152.0
Citric acid	cit	COOCH2COHCOOCH2	-161.5
Isocitric acid	icit	COOCH2CHCOOCHOH	-161.5
Malic acid	mal	COOCHOHCH2	-142.5

Because the molecular structure of the hypothetical substituent R in this equation is unspecified, the Gibbs energy of formation of HS-R and R'-COS-R are unknown. To estimate the Gibbs energy of thioesterification reactions in our model, the difference between the reference Gibbs energy of formation of the minimum charge state of each carboxylic acid and its respective thioester, defined by

$$\Delta\Delta_f G_{R'}^{\circ\circ} := \Delta_f G_{R'-COO^-}^{\circ\circ} + \Delta_f G_{HS-R}^{\circ\circ} - \Delta_f G_{R'-COS-R}^{\circ\circ}, \quad (2)$$

are given as model parameters (Table S1). Here, the superscript $\circ\circ$ denotes the reference Gibbs energy of formation of the minimum charge state of the species indicated by the subscript. Since these parameters only provide the Gibbs energy of formation of the unspecified thioesters R'-COS-R relative to that of HS-R, we use the reference Gibbs energy of formation of the minimum charge state of ethanethiol as a basis for the Gibbs energy of formation of HS-R.

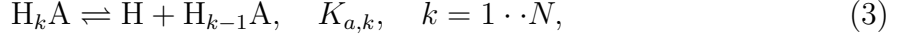
In our model, thioesterification reactions are catalyzed by both acids and Fe^{2+} . Similarly to biochemical reactions, we assume that the nonenzymatic reactions in our protometabolic network occur at a fixed pH. When pH is a control parameter besides temperature and pressure, the spontaneity of chemical reactions is determined by the transformed Gibbs energy of reaction [57, 58]. This Gibbs energy accounts for acid dissociation constants and depends on the concentration of H^+ (see ‘‘Thermodynamic Constraints’’). In this formulation, the catalytic effect of Brønsted acids is captured by the change in the transformed Gibbs energy of reaction because of variations in the acid concentration. According to this formulation, thioesterification reactions have a larger negative transformed Gibbs energy of reaction and are more thermodynamically favorable in more acidic environments [58] because they consume H^+ . As stated before, in our models, thioesterification reactions occur inside the protocell in an acidic environment, where they are both thermodynamically and kinetically favored. Accordingly, protocells with an acidic interior would have provided a desirable environment for first carbon-fixing cycles to emerge by alleviating their thermodynamic bottlenecks.

Once the transformed reference Gibbs energy of formation of all the species in Eq. (1) and their acid dissociation constants have been determined, the transformed Gibbs energy of the thioesterification steps at a given pH can be ascertained. Note that, besides the reference Gibbs energy of formation of the minimum charge state, we used the acid dissociation constants of ethanethiol to estimate those of the hypothetical thiol in our model.

Mass Balance Constraints and Maxwell’s first law

Conservation of mass is a fundamental constraint that we impose in our protocell model. We solve the mass-balance equations to determine how the concentrations of

different compounds in the protometabolic network vary with time. Before proving the detailed formulation of these equations, we introduce a few concepts that are important in the forthcoming discussion. One is the idea of combining the protonation states of a molecule into an aggregate compound, as defined by Alberty [57]. Most molecules that participate in protometabolism behave like weak acids in water, releasing hydrogen ions in multiple steps



where N is the number of hydrogen-dissociation steps, and A represents the minimum charge state of a molecule in the system with $K_{a,k}$ the respective hydrogen-dissociation constants. These protonation or charge states are referred to as species [57, 58]. Their distribution can vary depending on the pH and ionic strength of the solution. In general, all these charge states are utilized by metabolic networks, so hydrogen-dissociation reactions must be incorporated into metabolic reactions to account for the effect of pH.

Hydrogen-dissociation reactions tend to equilibrate faster than metabolic reactions [57]. As a result, the distribution of species vary independently from the extension of metabolic reactions. Therefore, we simplify our analysis by lumping all the species together, representing them as a single reactant with respect to which mass-balance equations are expressed

$$\mathcal{A} := \{\text{A}, \dots, \text{H}_N\text{A}\}. \quad (4)$$

Here, \mathcal{A} represent a reactant in the system the concentration of which is given by

$$[[\mathcal{A}]] := [\text{A}] + \dots + [\text{H}_N\text{A}]. \quad (5)$$

In subsequent sections, we sometimes denote the concentration of \mathcal{A} by $C_{\mathcal{A}}$ to simplify the notation. For reactant \mathcal{A} with only one species A , $[[\mathcal{A}]]$ and $[\text{A}]$ are indistinguishable. For each species $\text{A}' \in \mathcal{A}$, the corresponding model fraction $\rho_{\text{A}'} := [\text{A}']/[[\mathcal{A}]]$ is calculated from hydrogen-dissociation constants irrespective of the reactions A' participates in [58].

As previously stated, the protocell membrane in our model is made of semi-permeable minerals to allow the exchange of materials between the protocell and its environment. The positive membrane potential developed across the membrane due to a pH gradient affects the transport rate of all charged species. The transport rate of charged or uncharged species through such membranes can be determined by solving the species mass-balance equation and Maxwell's first law [1, 59]. We approximate the diffusive transport rates by [59]

$$J_i = \frac{\vartheta D_i}{d} z_i \Delta \hat{\psi} \frac{C_i^{\text{out}} - C_i \exp(z_i \Delta \hat{\psi})}{\exp(z_i \Delta \hat{\psi}) - 1}, \quad (6)$$

where $\Delta \hat{\psi} := F \Delta \psi / (RT)$ with $\Delta \psi$ the membrane potential, ϑ tortuosity coefficient, F Faraday constant, T temperature, and R the universal gas constant. Moreover, J_i denotes the diffusive flux, D_i the diffusivity coefficient, z_i the valence, C_i intracellular concentration, and C_i^{out} extracellular concentration of species i . Note that Eq. (6) is an approximate expression derived for a flat membrane, where the electric potential field ψ varies linearly along its thickness [59]. We may derive a more general expression for the membrane transport rate of reactant \mathcal{A} from Eq. (6) that accounts for other factors that contribute to its import flux

$$J_{\mathcal{A}} = M_{\mathcal{A}}^{\text{out}} C_{\mathcal{A}}^{\text{out}} - M_{\mathcal{A}} C_{\mathcal{A}} + J_{\mathcal{A}}^0, \quad (7)$$

where

$$J_{\mathcal{A}} := \sum_{A' \in \mathcal{A}} J_{A'}, \quad (8a)$$

$$M_{\mathcal{A}}^{\text{out}} := \frac{\vartheta}{d} \sum_{A' \in \mathcal{A}} D_{A'}^{\text{out}} \Psi_{A'} \rho_{A'}^{\text{out}}, \quad (8b)$$

$$M_{\mathcal{A}} := \frac{\vartheta}{d} \sum_{A' \in \mathcal{A}} D_{A'} \Psi_{A'} \exp\left(z_{A'} \Delta \hat{\psi}\right) \rho_{A'}, \quad (8c)$$

with $\rho_{A'}^{\text{out}}$ the mole fraction of A' in the extracellular environment, $D_{A'}^{\text{out}}$ diffusivity coefficient of A' in the extracellular environment, $\rho_{A'}$ the mole fraction of A' in the intracellular environment, $D_{A'}$ diffusivity coefficient of A' in the intracellular environment, and

$$\Psi_{A'} := \begin{cases} \frac{z_{A'} \Delta \hat{\psi}}{\exp(z_{A'} \Delta \hat{\psi}) - 1}, & z_{A'} \neq 0, \\ 1, & z_{A'} = 0. \end{cases} \quad (9)$$

The constant term $J_{\mathcal{A}}^0$ in Eq. (7) is a given parameter of the model. For all reactants except Fe^{2+} , $J_{\mathcal{A}}^0 = 0$. However, for Fe^{2+} , $J_{\mathcal{A}}^0$ is the rate at which the iron-sulfide membrane dissolves into the intracellular environment (see Fig. S1C).

Assuming that all the species occupy a spherical region as they move, we approximate their diffusivity coefficient by Stokes–Einstein equation

$$D_{A'} = \frac{k_{\text{B}} T}{6\pi\mu a_{A'}}. \quad (10)$$

Here, k_{B} is the Boltzmann constant, μ the viscosity of water in the membrane, and $a_{A'}$ the radius of the sphere that characterize A' . We used the Van der Waals radius of A' to estimate $a_{A'}$. Van der Waals radii were computed with Calculator Plugins, Marvin 20.21.0, 2020, ChemAxon (<http://www.chemaxon.com>). The Van der Waals radii of the unspecified thiol HS–R and its thioesters derivatives R'–COS–R were approximated assuming that the hypothetical substituent R is the alkyl group of ethanethiol.

Having introduced the definition of diffusive flux and concentration for species and reactants, we now can proceed to the formulation of mass-balance equations. Here, the idea is to express these equations with respect to reactants instead of species, solving them for reactant concentrations. A balance between reaction and membrane transport rates determines these concentrations. Reactions in our protometabolic network are divided into two groups: (i) Membrane reactions. The reductive carboxylation steps and the reaction for the regeneration of FD_{rd} are in this group (see “Model Description”). These reactions occur inside the protocell membrane. (ii) Bulk reactions. All reactions of the protometabolic network besides the membrane reactions belong to this group. These reactions occur inside the protocell. Accounting for membrane transport and both groups of reactions, the mass-balance equations for all the reactants involved in the protometabolic network are written

$$\frac{dC_{\mathcal{A}}}{dt} = \frac{A_c}{V_c} J_{\mathcal{A}} + \eta r_{a,\mathcal{A}} + r_{v,\mathcal{A}}, \quad \mathcal{A} \in \text{MET}, \quad (11)$$

where $\eta := \mathcal{V}/V_c$ with A_c the inner surface area of the membrane, V_c volume of the protocell, \mathcal{V} volume of the membrane, $r_{a,\mathcal{A}}$ production rate of \mathcal{A} by the membrane reactions,

$r_{v,\mathcal{A}}$ production rate of \mathcal{A} by bulk reactions, t time, and MET the index set of reactants that participate in metabolic reactions. Using the diffusive transport rates given by Eq. (7), Maxwell’s first law is implicitly applied in Eq. (11). Accordingly, the solutions of Eq. (11) simultaneously satisfy mass-balance constraints and Maxwell’s first law.

Thermodynamic Constraints

Thermodynamic laws play a fundamental role in constraint-based models of metabolism [58, 60]. They dictate the limiting behavior of chemical reactions as they approach their equilibrium. The rate of reactions that are not in equilibrium are affected by their thermodynamic driving force [61]. As will be discuss in subsequent sections (see “Kinetic Constraints”), thermodynamic driving forces are determined by the equilibrium constants of chemical reactions.

Nonenzymatic reactions in the protometabolic network occur on catalytic surfaces similarly to enzymatic reactions of biochemistry. Mineral surfaces that catalyze these reactions in our model are made of transitions metals, which are positively charged in most cases [2]. The energetics of these reactions are affected by pH due to the interactions of hydrogen ions, protonation states of reactants, and charged surfaces. To determine the spontaneity of these reaction types, their Gibbs energy must be evaluated at a fixed pH [57]. This constraint can be incorporated into the Gibbs energy of reaction using the Legendre transformation. The quantity that results from the Legendre transformation of the Gibbs energy of reaction is referred to as the transformed Gibbs energy of reaction [57]. Its sign determines the spontaneity reactions for systems, where the temperature, pressure, and pH are the control parameters. As with the Gibbs energy of reaction, the transformed Gibbs energy of reaction at a given condition is evaluated with respect to a standard condition according to

$$\frac{\Delta_r G'_j}{RT} = \frac{\Delta_r G_j^{\prime\circ}}{RT} + \ln \Gamma_j, \quad \Gamma_j := \prod_{i \in \text{MET}} \left(\frac{C_i}{C^\circ} \right)^{S_{ij}}, \quad (12)$$

$$\frac{\Delta_r G'_j}{RT} = \sum_{i \in \text{MET}} S_{ij} \frac{\Delta_f G_i^{\prime\circ}}{RT} + S_{w,j} \frac{\Delta_f G_w^{\prime\circ}}{RT}, \quad (13)$$

where C° is a reference concentration, S_{ij} are the components of the stoichiometry matrix of the protometabolic network, $S_{w,j}$ stoichiometric coefficient of water in reaction j , $\Delta_f G_w^{\prime\circ}$ standard transformed Gibbs energy of formation of water, and $\Delta_f G_i^{\prime\circ}$ standard transformed Gibbs energy of formation of reactant i with $\Delta_r G'_j$ the transformed Gibbs energy, $\Delta_r G_j^{\prime\circ}$ standard transformed Gibbs energy, and Γ_j quotient of reaction j . Note that, in this formulation, the concentrations of hydrogen ions and water do not affect the reaction quotients. Their stoichiometric coefficients do not contribute to S either. The standard transformed Gibbs energy of formation of a reactant \mathcal{A} is given by [58]

$$\begin{aligned} \frac{\Delta_f G_{\mathcal{A}}^{\prime\circ}}{RT} &= \frac{\Delta_f G_{\mathcal{A}}^{\circ\circ}}{RT} + N_{\text{H},\mathcal{A}} [\ln(10)\text{pH} - \ln \gamma_{\text{H}}] \\ &\quad - \ln \left\{ \sum_{k=0}^N \frac{\gamma_{\text{H}}^k 10^{-k\text{pH}}}{\gamma_{\text{H}_k\mathcal{A}} \prod_{i=0}^k K_{a,i}^{\circ\circ}} \exp \left[-(z_{\mathcal{A}} + k) \frac{F\psi}{RT} \right] \right\}. \end{aligned} \quad (14)$$

Here, $N_{\text{H},\mathcal{A}}$ denotes the hydrogen content of the minimum charge state, $\Delta_f G_{\mathcal{A}}^{\circ\circ}$ reference Gibbs energy of formation of the minimum charge state, γ activity coefficient of the species

indicated by the subscript, $K_{a,i}^{\circ\circ}$ reference hydrogen-dissociation constant of species i , and ψ electric potential at which $\Delta_f G_A^{\prime\circ}$ is evaluated. Note that, for each species i , the activity coefficient is a function of the ionic strength I of the solution it is in and its valence z_i [58]. Accordingly, the apparent equilibrium constant of reaction j is defined as

$$K'_j := \exp\left(-\frac{\Delta_r G_j^{\prime\circ}}{RT}\right). \quad (15)$$

The activity coefficients in Eq. (13) depend on the ionic strength of the solution I and are approximated by the extended Debye–Hückel as described elsewhere [57, 58]. Finally, the thermodynamic constraints of the protocell model is expressed as

$$\Delta_r G'_j \leq 0, \quad j \in \text{RXN} \quad (16)$$

with RXN the index set of reactions in the protometabolic network.

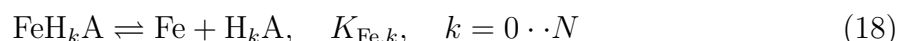
From Eqs. (13)–(15), it is evident that $\Delta_f G_A^{\circ\circ}$ and $K_{a,i}^{\circ\circ}$ are required for all reactants to ascertain K'_j . The value of these parameters for all the reactants of the protometabolic network is provided in Supplementary Data 1. We computed the reference Gibbs energy of formation of the minimum charge states using eQuilibrator (version 3.0) [62] and the reference hydrogen-dissociation constants using Calculator Plugins, Marvin (version 20.21.0, 2020), ChemAxon (<http://www.chemaxon.com>). In our protometabolic network, there are two nonstandard reactants for which $\Delta_f G_A^{\circ\circ}$ cannot be estimated from existing databases of tuned group-contribution parameters generated for standard metabolites of enzymatic metabolic pathways, namely hydroxyketoglutarate and oxalohydroxyglutarate. For these reactants, we estimated the reference Gibbs energy of formation of the minimum charge state using dGPredictor [63]. Both eQuilibrator and dGPredictor are web applications that estimate the standard transformed Gibbs energy of formation using group-contribution methods.

The energetics of the membrane reactions (*i.e.*, those involving protoferredoxins; see “Mass Balance Constraints and Maxwell’s first law”) depend on the standard transformed Gibbs energy of formation of FD_{rd} and FD_{ox} . In fact, the absolute values of these formation energies are not required. Providing the difference between these formation energies suffices to determine the standard transformed Gibbs energy of the membrane reactions. This difference may be expressed as

$$\Delta_f G_{\text{FD}_{\text{rd}}}^{\prime\circ} - \Delta_f G_{\text{FD}_{\text{ox}}}^{\prime\circ} = -F \Delta E_{\text{redox}}, \quad (17)$$

where ΔE_{redox} is the redox potential of protoferredoxins, which is a given parameter of the model. Once ΔE_{redox} has been provided, the standard transformed Gibbs energy of the membrane reactions can be specified.

Besides hydrogen ions, early metabolic reactions could also have been affected by metal-ion bindings. The early ocean is believed to have contained higher concentrations of several monovalent and polyvalent metal ions than modern seawater [64]. In particular, it would have contained appreciable amounts of ferrous iron in the absence of oxygen [19], which in turn could have interacted with other reactants that participated in early metabolic networks. These interactions would have altered the formation energy of reactants and the energetics of metabolic reactions. In this work, we only account for the binding of ferrous iron, as represented by



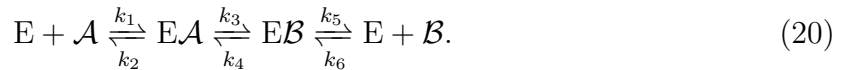
with $K_{\text{Fe},k}$ ferrous-iron binding constants. Note that, in this equation, the charge of compounds has not been shown for brevity. Accounting for the charge states arising from ferrous-iron binding reactions, we arrive at

$$\begin{aligned} \frac{\Delta_f G'_{\mathcal{A}}|_{\text{Fe}}}{RT} &= \frac{\Delta_f G_A^{\circ\circ}}{RT} + N_{\text{H},\text{A}}[\ln(10)\text{pH} - \ln \gamma_{\text{H}}] \\ &\quad - \ln \left\{ \sum_{k=0}^N \frac{\gamma_{\text{H}}^k 10^{-k\text{pH}}}{\gamma_{\text{H},k\text{A}} \prod_{i=0}^k K_{a,i}^{\circ\circ}} \exp \left[-(z_{\text{A}} + k) \frac{F\psi}{RT} \right] \right. \\ &\quad \left. + \sum_{k=0}^N \frac{\gamma_{\text{H}}^k 10^{-k\text{pH}} \gamma_{\text{Fe}} 10^{-\text{pFe}}}{\gamma_{\text{FeH}_k\text{A}} K_{\text{Fe},k}^{\circ\circ} \prod_{i=0}^k K_{a,i}^{\circ\circ}} \exp \left[-(z_{\text{A}} + k + 2) \frac{F\psi}{RT} \right] \right\}, \end{aligned} \quad (19)$$

where $\Delta_f G'_{\mathcal{A}}|_{\text{Fe}}$ denotes the standard transformed Gibbs energy of formation of \mathcal{A} when the binding reactions of Fe^{2+} are accounted for. We use the notations $\Delta_r G'|_{\text{Fe}}$ and $K'|_{\text{Fe}}$ in a similar manner. Note that, in the main article, we dropped the subscript Fe when referring to $\Delta_r G'|_{\text{Fe}}$. Accordingly, the transform Gibbs energy of reactions, which are shown in Figs. 2E and 3C, are based on the expression given by Eq. (19).

Kinetic Constraints

All the reactions of the protometabolic network are surface-catalyzed. Therefore, to impose kinetic constraints, we require rate laws that account for reactant-catalyst bindings and surface reaction rates. We adopt Michaelis–Menten type expressions to quantify the kinetics of surface metabolic reactions in our model. There are eight reaction types in the protometabolic network: (i) thioesterification, (ii) membrane reactions (reductive carboxylation and protoferredoxin regeneration), (iii) retro-aldol and aldol, (iv) oxidative decarboxylation, (v) hydration and dehydration, (vi) reduction, (vii) decarboxylation, and (viii) solid catalyst synthesis. In this section, we present the rate laws that we used in our model for each reaction type to solve mass-balance equations Eq. (11). However, we first consider the Michaelis–Menten reaction mechanism



We use the general rate law for this reaction system to derive the rate laws for each reaction types in our model. Here, \mathcal{A} , \mathcal{B} , and E represent the substrate, product, and catalyst of a given reaction in the network with k_i the rate constants of the individual steps. The overall rate law for this reaction system reads [61]

$$v = \frac{k_{\text{cat}}}{K_{\text{m},\mathcal{A}}} [\text{E}] \left([\text{A}] - \frac{[\text{B}]}{K'|_{\text{Fe}}} \right), \quad (21\text{a})$$

$$k_{\text{cat}} := \frac{k_3 k_5}{k_3 + k_4 + k_5}, \quad (21\text{b})$$

$$K_{\text{m},\mathcal{A}} := \frac{k_2 k_4 + k_2 k_5 + k_3 k_5}{k_1 (k_3 + k_4 + k_5)}, \quad (21\text{c})$$

where $[\text{E}]$ is the concentration of unbound catalyst, k_{cat} the turnover number, and $K_{\text{m},\mathcal{A}}$ is the Michaelis constant of \mathcal{A} . As mentioned in the previous section, ferrous-iron binding constants are required to determine $K'|_{\text{Fe}}$. Since these binding constants are generally unknown for most reactants in our protometabolic network, we simplify Eq. (19) in a way

that allows us to parametrize them. To this end, we first write $\Delta_f G'_{\mathcal{A}}|_{\text{Fe}}$ and $\Delta_f G'_{\mathcal{B}}|_{\text{Fe}}$ in terms of $\Delta_f G'_{\mathcal{A}}$ and $\Delta_f G'_{\mathcal{B}}$, respectively. Let $T_{\text{H},\mathcal{A}}$ and $T_{\text{Fe},\mathcal{A}}$ stand for the first and second sum in the argument of the logarithm function in Eq. (19) for \mathcal{A} . Let also Let $T_{\text{H},\mathcal{B}}$ and $T_{\text{Fe},\mathcal{B}}$ the same sums for \mathcal{B} . Then, we obtain the following standard transformed Gibbs energy of formation

$$\begin{aligned}\frac{\Delta_f G'_{\mathcal{A}}|_{\text{Fe}}}{RT} &= \frac{\Delta_f G_{\mathcal{A}}^{\circ\circ}}{RT} + N_{\text{H},\mathcal{A}}[\ln(10)\text{pH} - \ln \gamma_{\text{H}}] - \ln(T_{\text{H},\mathcal{A}} + T_{\text{Fe},\mathcal{A}}), \\ \frac{\Delta_f G'_{\mathcal{B}}|_{\text{Fe}}}{RT} &= \frac{\Delta_f G_{\mathcal{B}}^{\circ\circ}}{RT} + N_{\text{H},\mathcal{B}}[\ln(10)\text{pH} - \ln \gamma_{\text{H}}] - \ln(T_{\text{H},\mathcal{B}} + T_{\text{Fe},\mathcal{B}}).\end{aligned}$$

Assuming

$$T_{\text{Fe},i} \ll T_{\text{H},i}, \quad i \in \{\mathcal{A}, \mathcal{B}\}, \quad (22)$$

we have

$$\ln(T_{\text{H},i} + T_{\text{Fe},i}) \approx \ln T_{\text{H},i} + \frac{T_{\text{Fe},i}}{T_{\text{H},i}}.$$

This assumption is justified by comparing the values of hydrogen-dissociation and metal-ion binding constants. Although we do not have all the binding constants for ferrous iron, we may estimate them from the binding constants of similar metal ions for which experimental data are available. The binding constants of similar divalent ions, such as Zn^{2+} and Mg^{2+} , to a wide range of metabolites participating in biochemical reactions have been measured, and they are generally several orders of magnitude larger than hydrogen-dissociation constants of the same metabolites [65]. Thus, we expect that the binding constants of ferrous iron exhibit the same behavior for the reactants of the protometabolic network. Using the foregoing assumption, we obtain

$$K'|_{\text{Fe}} = wK', \quad (23a)$$

$$w := \exp[\Delta_r(T_{\text{Fe}}/T_{\text{H}})], \quad (23b)$$

$$\Delta_r(T_{\text{Fe}}/T_{\text{H}}) := \frac{T_{\text{Fe},\mathcal{B}}}{T_{\text{H},\mathcal{B}}} - \frac{T_{\text{Fe},\mathcal{A}}}{T_{\text{H},\mathcal{A}}}. \quad (23c)$$

Substituting Eq. (23a) in Eq. (21a) yields

$$v = k_{\text{cat}}[\text{E}] \left(\frac{[\text{A}]}{K_{\text{m},\mathcal{A}}} - \frac{[\text{B}]}{K' \tilde{K}_{\text{m},\mathcal{B}}} \right), \quad (24a)$$

$$\tilde{K}_{\text{m},\mathcal{B}} := wK_{\text{m},\mathcal{A}}. \quad (24b)$$

Here, $w \sim \mathcal{O}(1)$ because $|\Delta_r(T_{\text{Fe}}/T_{\text{H}})| \ll 1$ according to the assumption stated in Eq. (22). Therefore, $\tilde{K}_{\text{m},\mathcal{B}}$ is of the same order of magnitude as $K_{\text{m},\mathcal{A}}$. Note that $\tilde{K}_{\text{m},\mathcal{B}}$ is not the same as the Michaelis constant $K_{\text{m},\mathcal{B}}$ that results from the reaction system Eq. (20). Thus, to avoid confusion, we refer to $\tilde{K}_{\text{m},\mathcal{B}}$ as the apparent Michaelis constant of \mathcal{B} at this point. However, as will be explained shortly, distinguishing between Michaelis and apparent Michaelis constants will not be necessary in the remainder of this document.

For a given reactant \mathcal{X} , $K_{\text{m},\mathcal{X}}$ and $\tilde{K}_{\text{m},\mathcal{X}}$ generally depend on the kinetic and Fe^{2+} -binding constants of the reaction that \mathcal{X} participates in. However, to minimize the number of model parameters, we assume that the Michaelis and apparent Michaelis constants of a given reactant in the protometabolic network is the same for all reactions.

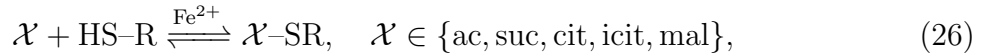
Because of this assumption, we no longer need to distinguish between Michaelis and apparent Michaelis constants in the rate laws that will be discussed later in this section. Accordingly, we rewrite Eq. (24a) as

$$v = k_{\text{cat}}[\text{E}] \left(\frac{[\text{A}]}{K_{\text{m},\text{A}}} - \frac{[\text{B}]}{K'K_{\text{m},\text{B}}} \right) \quad (25)$$

with $K_{\text{m},\text{A}}$ and $K_{\text{m},\text{B}}$ the Michaelis constant of the substrate and product, which are regarded as model parameters. In the following, we derive rate laws for each reaction type using Eq. (25). Note that reaction rates that are ascertained from $[\text{A}]$ and $[\text{B}]$ according to this rate law automatically satisfy the thermodynamic constraints, irrespective of the flux direction.

1. Thioesterification:

As previously stated, thioesterification reactions are catalyzed by H^+ and Fe^{2+} in our model. Accordingly, the general reaction equation for thioesterification steps can be written



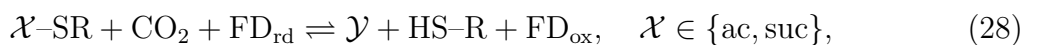
where \mathcal{X} represent a carboxylic acids and $\mathcal{X}\text{-SR}$ the respective thioester. Reactant identifiers are given in Supplementary Data 1. Following the development of Alberty [57], we do not explicitly state the stoichiometric coefficients of H^+ and H_2O that are consumed or produced in reaction equations—a convention we adhere to in the remainder this document. Given that all reactions occur in a dilute aqueous environment, the concentration of water can be regarded as a constant. Therefore, it only affects the reaction rates indirectly through their thermodynamic driving forces, or apparent equilibrium constants to be more specific (see Eq. (13)). Similarly, we assume that H^+ only affects the reaction rates by altering the thermodynamic driving forces. The contribution of H^+ is captured in Eq. (25) by using the apparent equilibrium constant to evaluate the thermodynamic driving force. The apparent equilibrium constant in turn accounts for the energetics of H^+ being exchanged between reactants and the aqueous environment. Taking all these considerations into account, the rate of thioesterification reactions can be expressed as

$$v = k_{\text{cat}}[\text{Fe}^{2+}] \left(\frac{[\text{X}][\text{SR}]}{K_{\text{m},\mathcal{X}}K_{\text{m},\text{HS-R}}} - \frac{[\text{X-SR}]}{K'K_{\text{m},\mathcal{X}\text{-SR}}} \right) \quad (27)$$

with X, SR, and X-SR the minimum charge states of \mathcal{X} , HS-R, and $\mathcal{X}\text{-SR}$, respectively. Recall, we adopted the notation $[\text{minimum charge state}]$ to represent the concentration of a reactant (see “Mass Balance Constraints and Maxwell’s first law”).

2. Membrane reactions:

The reductive carboxylation steps of the protometabolic network occur inside the membrane on its acidic side. They are driven by protoferredoxin redox couples according to



where $\mathcal{X}\text{-SR}$ is the thioester derived from the carboxylic acid \mathcal{X} and \mathcal{Y} the carboxylic acid resulting from the carboxylation of \mathcal{X} . These reactions consume H^+ , so they are thermodynamically favorable in acidic solutions. Decreasing the intracellular pH increases their apparent equilibrium constants and their rates as a result. Moreover, protoferredoxins

serve as both catalyst and reducing/oxidizing agents for these reactions. We use the rate law

$$v = k_{\text{cat}} \left(\frac{[\text{FD}_{\text{rd}}][[\text{X-SR}][\text{CO}_2]}{K_{\text{m},\mathcal{X}\text{-SR}}K_{\text{m},\text{CO}_2}} - \frac{[\text{FD}_{\text{ox}}][[\text{Y}][[\text{SR}]]]}{K'K_{\text{m},\mathcal{Y}}K_{\text{m},\text{HS-R}}} \right) \quad (29)$$

to describe the kinetics of reductive carboxylation reactions. Here, Y is the minimum charge state of \mathcal{Y} with $[\text{FD}_{\text{rd}}]$ and $[\text{FD}_{\text{ox}}]$ the concentrations of FD_{rd} and FD_{ox} , respectively. These concentrations are defined as the number of moles of the respective protoferredoxin unit cell per unit volume of the membrane. In this equation, we assume that the rate of forward and reverse reactions are proportional to the amount of FD_{rd} and FD_{ox} that are present in the membrane.

The reduced protoferredoxin FD_{rd} , which is consumed by the reductive carboxylation reactions, is regenerated inside the membrane on its alkaline side

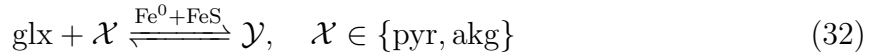


using H_2 as the reducing agent. This reaction produces H^+ , and it is thermodynamically favorable in alkaline environments. The rate law for this reaction reads

$$v = k_{\text{cat}} \left(\frac{[\text{FD}_{\text{ox}}][\text{H}_2]}{K_{\text{m},\text{H}_2}} - \frac{[\text{FD}_{\text{rd}}]}{K'} \right). \quad (31)$$

3. Retro-aldol and aldol:

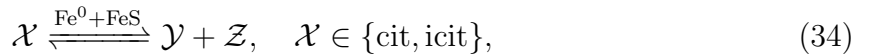
Aldol and retro-aldol reactions play a significant role in biochemistry. They can be catalyzed by a wide range of organic and inorganic catalysts [66]. They can also be promoted by acids or bases depending on the tendency of the participating compounds to accept or donate H^+ [67]. Among inorganic catalysts, transition metals [68] and chiral metal complexes [67] are of particular interest in the context of life's metabolic origins. With regards to the rTCA cycle, FeS was shown to promote the retro-aldol cleavage of citrate into acetate and oxaloacetate [39]. Similarly, the $\text{Fe}^{2+}/\text{Fe}^0$ system was observed to catalyze several aldol reactions involving the intermediates of the rTCA cycle [11]. In our model, aldol and retro-aldol reactions are catalyzed by both FeS and Fe^0 , which are synthesized through the interaction of Fe^{2+} and HS^- in the protocell. Aldol reactions in the protometabolic network can be described by the general equation



with \mathcal{X} a carboxylic acid and \mathcal{Y} the carboxylic acid that results from the aldol condensation of glyoxylate and \mathcal{X} . The rate law for these reactions is

$$v = k_{\text{cat}}\beta_{\text{aldol}} ([\text{Fe}^0] + [\text{FeS}]) \left(\frac{[\text{GLX}][[\text{X}]]}{K_{\text{m},\text{glx}}K_{\text{m},\mathcal{X}}} - \frac{[\text{Y}]}{K'K_{\text{m},\mathcal{Y}}} \right), \quad (33)$$

where GLX is the minimum charge state of glyoxylate. Retro-aldol reactions are described by



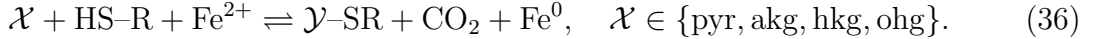
where \mathcal{X} is a carboxylic acid with \mathcal{Y} and \mathcal{Z} the carboxylic acids that results from the retro-aldol cleavage of \mathcal{X} . The rate law for these reactions read

$$v = k_{\text{cat}}\beta_{\text{retro-aldol}} ([\text{Fe}^0] + [\text{FeS}]) \left(\frac{[\text{X}]}{K_{\text{m},\mathcal{X}}} - \frac{[\text{Y}][[\text{Z}]]}{K'K_{\text{m},\mathcal{Y}}K_{\text{m},\mathcal{Z}}} \right) \quad (35)$$

with Z the minimum charge state of \mathcal{Z} . In these rate laws, β_{aldol} and $\beta_{\text{retro-aldol}}$ are mole-fractions of FeS-Fe^0 assemblages that catalyze aldol and retro-aldol reactions, respectively.

4. Oxidative decarboxylation:

Oxidative decarboxylation steps are one group of off-cycle reactions that we accounted for in the protometabolic network. They are parasitic reactions since they counteract the reductive carboxylation steps, diminishing the efficiency of the carbon fixation. The competition between oxidative decarboxylation and reductive carboxylation reactions is an important factor, determining whether a complete-cycle flux can be achieved. Previous studies on nonenzymatic analogues the TCA cycle detected these reactions in the presence of Fe^{2+} [11]. In the protocell model, these reactions are catalyzed by Fe^{2+} as indicated by

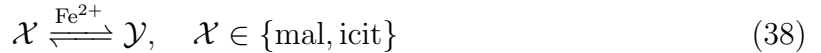


Here, \mathcal{X} represent a carboxylic acid and \mathcal{Y} the carboxylic acid that results from the decarboxylation of \mathcal{X} . Note that Fe^{2+} and Fe^0 simultaneously serve as reducing/oxidizing agents and catalysts in these equations. The following rate law describes the kinetics of these reactions

$$v = k_{\text{cat}} \left(\frac{[\text{Fe}^{2+}][\mathcal{X}][\text{SR}]}{K_{\text{m},\mathcal{X}}K_{\text{m},\text{HS-R}}} - \frac{[\text{Fe}^0][\mathcal{Y}\text{-SR}][\text{CO}_2]}{K'K_{\text{m},\mathcal{Y}\text{-SR}}K_{\text{m},\text{CO}_2}} \right). \quad (37)$$

5. Hydration and dehydration:

Previous experimental studies showed that the hydration and dehydration steps of the rTCA and TCA cycles can be promoted by Fe^{2+} , Zn^{2+} , and Cr^{3+} without enzymes [10, 11]. Specifically, both Fe^{2+} and Zn^{2+} were observed to promote the dehydration of malate and isocitrate to fumarate and aconitate, respectively [10, 11]. However, the hydration of aconitate to citrate was observed only in the presence of Cr^{3+} [10]. In the protocell model, we consider the dehydration reactions



with the rate law

$$v = k_{\text{cat}}[\text{Fe}^{2+}] \left(\frac{[\mathcal{X}]}{K_{\text{m},\mathcal{X}}} - \frac{[\mathcal{Y}]}{K'K_{\text{m},\mathcal{Y}}} \right), \quad (39)$$

where \mathcal{X} is a carboxylic acid and \mathcal{Y} the product of its dehydration. The only hydration reaction in the protometabolic network is



the rate law of which is given by

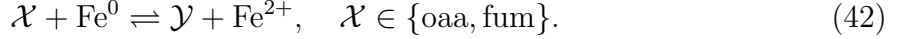
$$v = k_{\text{cat}}[\text{Cr}^{3+}] \left(\frac{[\text{ACON}]}{K_{\text{m},\text{acon}}} - \frac{[\text{CIT}]}{K'K_{\text{m},\text{cit}}} \right) \quad (41)$$

with ACON and CIT the minimum charge states of acon and cit, respectively.

6. Reduction:

In the absence of oxygen, the early Earth's crust is believed to have been made of reducing mineral assemblages, mostly containing silicate, iron, nickel, and magnesium

[19, 69]. Among these minerals, zero valent iron Fe^0 and iron monosulfides FeS have been extensively studied in the origins-of-life literature as the main reducing agent [19, 70]. Specifically, Fe^0 were shown to promote the reduction steps of the rTCA cycle [10]. In the protometabolic network, the reduction steps are also driven by Fe^0 according as described by the following equation



Here, \mathcal{X} is a carboxylic acid and \mathcal{Y} the product of its reduction. The rate law for these reactions is

$$v = k_{\text{cat}} \left(\frac{[\text{Fe}^0][\mathcal{X}]}{K_{\text{m},\mathcal{X}}} - \frac{[\text{Fe}^{2+}][\mathcal{Y}]}{K'K_{\text{m},\mathcal{Y}}} \right). \quad (43)$$

In these reactions, Fe^0 and Fe^{2+} are both catalysts and reducing/oxidizing agents.

7. Decarboxylation:

Neutral decarboxylation steps are another group of off-cycle reactions that we considered in the protometabolic network. Nonenzymatic decarboxylation reactions in the presence of Fe^{2+} were observed previously in the prebiotic chemistry literature [11]. According, the only neutral decarboxylation step of the protometabolic network is catalyzed by Fe^{2+} as



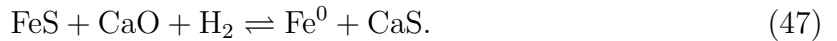
the rate law of which is

$$v = k_{\text{cat}}[\text{Fe}^{2+}] \left(\frac{[\text{OAA}]}{K_{\text{m},\text{oaa}}} - \frac{[\text{PYR}][\text{CO}_2]}{K'K_{\text{m},\text{pyr}}K_{\text{m},\text{CO}_2}} \right), \quad (45)$$

where OAA and PYR are the minimum charge states of oaa and pyr, respectively.

8. Solid catalyst synthesis:

In the discussion of the rate laws so far, we noted that Fe^0 and FeS play the role of catalysts and reducing agents in the nonenzymatic reactions of the protometabolic network. As discussed in previous sections (see ‘‘Model Description’’), FeS-Fe^0 assemblages that precipitate on lime minerals inside the protocell as a result of the interaction of Fe^{2+} and HS^- are the sources of Fe^0 and FeS that promote the foregoing protometabolic reactions. In the protocell model, assemblages of FeS-Fe^0 in turn are synthesized from Fe^{2+} through



Reactions of the type in Eq. (47) were previously examined in the context of extractive metallurgy [71]. The rate laws for these reactions are given by

$$v = k_{\text{cat}} \left(\frac{[\text{Fe}^{2+}][\text{S}]}{K_{\text{m},\text{H}_2\text{S}}} - \frac{[\text{FeS}]}{K'} \right), \quad (48)$$

$$v = k_{\text{cat}} \left(\frac{[\text{FeS}][\text{H}_2]}{K_{\text{m},\text{H}_2}} - \frac{[\text{Fe}^0]}{K'} \right), \quad (49)$$

where S in Eq. (48) is the minimum charge state of H_2S . Note that, CaO and CaS are produced and consumed only in Eq. (47), so $[\text{CaO}] + [\text{CaS}]$ remains constant at all times.

Table S2: Inorganic reagents, promoting the nonenzymatic reactions of the protometabolic network. The concentration of these reagents are either the variables of the model, which are determined from the mass-balance equations, or fixed parameters of the model.

Reagent	Reducing	Oxidizing	Energy source	Catalyst	Type
ΔpH	✗	✗	✓	✗	Parameter
Cr^{3+}	✗	✗	✗	✓	Parameter
H^+	✗	✓	✗	✓	Parameter
H_2	✓	✗	✓	✗	Variable
H_2S	✓	✗	✓	✗	Variable
FeS	✓	✗	✓	✓	Variable
Fe^0	✓	✗	✓	✓	Variable
Fe^{2+}	✗	✓	✗	✓	Variable
FD_{rd}	✓	✗	✓	✓	Variable
FD_{ox}	✗	✓	✗	✓	Variable

Therefore, to simplify our analysis, we do not solve their mass-balance equations to determine their concentrations. Since we are only interested in the solution of mass-balance equations near steady states in this work, we assume that CaO and CaS are abundantly available in the protocell, so their concentrations do not deviate from steady-state values significantly, and they approximately remain constant in time. Accordingly, $[\text{CaO}]$ and $[\text{CaS}]$ can be factored out and implicitly accounted for in k_{cat} and K_{m,H_2} in Eq. (49).

Having detailed the kinetic constraints of the protocell model, we conclude this section by emphasizing two points, namely (i) inorganic reagents driving the protometabolic network and (ii) parametrization of the rate laws. With regards to the first point, all the steps of the protometabolic network are promoted by purely inorganic reagents. Catalysts, reducing/oxidizing agents, and energy sources are all either solid transition metals or inorganic ions in aqueous phase. Some of the reagents serve multiple purpose. For example, Fe^0 is a catalysts in some reactions and a reducing agents in others. Table S2 summarizes the characteristics of the inorganic reagents in the protocell model. Regarding the second point, in our formulation, the turnover number k_{cat} and Michaelis constant K_{m} are the key parameters that characterize the kinetics of nonenzymatic reactions. As will be explained in the next section, we determine the Michaelis constants by solving an optimization problem, which ensures that the mass-balance equations have a steady-state solution. However, we treat the turnover numbers as free parameters with respect to which the feasibility of first carbon-fixing cycles is to be investigated. Turnover numbers in turn measure the efficiency of catalysts that promote metabolic reactions. Therefore, we adopt this formulation to better understand the role of turnover numbers in the evolution of enzymes and genetic codes. The idea is to identify ranges turnover numbers that furnish stable, self-sustaining carbon-fixing cycles using parameter-sweep computations.

Although we used the same symbol to denote the turnover number and Michaelis constants of every reaction, the values of these parameters are generally not the same for all reactions. As previously stated, for a given reactant \mathcal{A} in the protometabolic network, we assume that $K_{\text{m},\mathcal{A}}$ is the same for all the reactions that \mathcal{A} participates in. However, for turnover numbers, our goal is to examine the feasibility of the protometabolic network with respect to a single parameter k_{cat} —the characteristic turnover number—

that captures the order of magnitude of the turnover numbers for all the reactions in the network. The turnover number $k_{\text{cat},j}$ of a given reaction j in the network can deviate from the characteristic turnover number. However, the deviations

$$\delta k_{\text{cat},j} := k_{\text{cat},j} - k_{\text{cat}}$$

are assumed not to change the order of magnitude of $k_{\text{cat},j}$. That is, we assume that

$$\frac{\delta k_{\text{cat},j}}{k_{\text{cat}}} \ll 1, \quad j \in \text{RXN}.$$

The underlying assumption of this approach is that the turnover numbers of the nonenzymatic reactions in the protometabolic network, which are all catalyzed by transition metals, are of the same order of magnitude. Although experimental data on the kinetic parameters of abiotic reactions in the context of life’s metabolic origins are scarce, this assumption is partly supported by a few kinetic studies on nonenzymatic analogues of core metabolic reactions. For example, Keller et al. [72] reproduced several steps of the TCA cycle, glyoxylate shunt, and succinic-semialdehyde pathway using a combination of iron sulfides and sulfate radicals, showing that the rate constants of the majority of the TCA reactions were in the same order of magnitude. In another work, Mayer et al. [16] studied the transamination reactions of glyoxylate with several α -keto acids using transition metals as catalysts and found turnover numbers of the order $k_{\text{cat}} \approx 10^{-5}$ 1/s for most cases. Moreover, the transition metals examined in this study enhanced the turnover number of glutamate transamination by a factor $\sim 10^2$ compared to the case where no catalyst was used.

Drawing on the results from previous kinetic studies, we simplify our analysis by assuming that the turnover number of every reaction in the protometabolic network is the same as the characteristic turnover number k_{cat} , neglecting all the turnover-number deviations ($\delta k_{\text{cat},j} \approx 0$, $j \in \text{RXN}$). Kinetic studies of transition-metal catalyzed counterparts of core metabolic reactions could provide more accurate estimates for $\delta k_{\text{cat},j}$ and benefit future systems-level kinetic models, such as that presented in this work. Such experimental efforts could complement theoretical approaches and help elucidate the role of catalytic efficiency, as measured by turnover numbers, in the emergence and evolution of first metabolic cycles.

To investigate the role of turnover numbers, we also require a reasonable range for the turnover numbers of abiotic metabolic reactions catalyzed by transition metals. Reasonable estimates for these turnover numbers can be obtained by comparing the efficiency of transition-metal catalyzed and enzyme-catalyzed reactions. Enzymes are highly optimized catalysts. On average, they can enhance the rate of uncatalyzed chemical transformations approximately by 10^{12} – 10^{16} fold and 10^7 – 10^8 fold at 25°C and 100°C , respectively [73]. In contrast, inorganic catalysts have comparatively poor efficiencies, and generally they are not expected to increase the turnover numbers of uncatalyzed reactions by more than 10^2 fold [4]. Given the temperature of early AHVs (see Fig. 1A) and using the median turnover number $k_{\text{cat}} \approx 10$ 1/s of enzymatic reactions [74] as a basis for comparing inorganic and organic catalysts, we find the range $k_{\text{cat}} \approx 10^{-5}$ – 10^{-4} 1/s for the turnover number of nonenzymatic reactions in our model, which is the range in which we examined the feasibility of the protometabolic network (see Fig. 2A).

Acetate Concentration in Early Alkaline Hydrothermal Vents

In this section, we emphasize the importance of acetate in the amplification of carbon-fixing cycles in our model. As discussed in the main text, high concentrations of acetate ensure that fully-cyclic carbon-fixing pathways can be sustained in our protocell model. For example, the results in Fig. 2 were obtained for $C_{\text{ac}}^{\text{out}} = 200$ mM (see Supplementary Data 1) with $C_{\text{ac}}^{\text{out}}$ the extracellular concentration of acetate. However, previous studies suggested that acetate could have been synthesized in millimolar concentrations ($C_{\text{ac}} \approx 5$ mM) under AHV conditions [29] and smaller concentrations were reported for modern hydrothermal systems [75]. Nevertheless, given the high solubility of acetate in water, millimolar to molar concentrations of acetate could have been achieved theoretically in AHVs of the early ocean. In particular, because of the large partial pressures of CO_2 in the Hadean ocean, production of acetate in large concentrations through linear carboxylation pathways, such as the Wood-Ljungdahl pathway, would have been thermodynamically favored.

Highly elevated pressures at deep sea AHVs (*e.g.*, 75–100 bar at Lost City [76]) could also have affected the turnover number of inorganic catalysts. In this work, our analysis was based on turnover number estimates for inorganic catalysts, which mostly relied on experimental data at atmospheric pressures (see “Kinetic Constraints”). However, elevated pressures can significantly increase collision frequency and efficiency of reactants, enhancing the rate constants of chemical reactions. For example, using the SS-QRRK model, up to ~ 2 fold increase in rate constants of unimolecular tautomerization reactions was obtained for a 2-fold increase in pressure [77].

To evaluate the effect of elevated pressures on turnover numbers and the feasibility of cyclic carbon fixation, we performed additional parameter estimation analyses with modified upper bounds on the extracellular concentrations of reactants (see “Steady States of Protocell”). We solved Eq. (54) by imposing an upper bound on the extracellular concentration of acetate to examine whether lower extracellular concentrations of acetate can also sustain a stable carbon-fixing cycle in our protocell model but for larger turnover numbers than those considered in the main text. In these parameter-estimation studies, the optimization solver found a stable steady-state solution when the upper bound on the extracellular concentration of acetate was set to $C_{\text{ac}}^{\text{out}} \leq 20$ mM within a one-hour runtime limit for $k_{\text{cat}} = 10^{-2} \text{ s}^{-1}$. Although we cannot rule out the possibility that stable steady-state solutions also exist when $C_{\text{ac}}^{\text{out}} \leq 20$ mM for the same turnover number ranges as those considered in the main text, these results suggest that stable carbon-fixing cycles can be sustained for millimolar concentrations of acetate in AHVs if we account for pressure-dependent rate constants.

Steady States of Protocell

Once the rates of reactions in the protometabolic network have been ascertained from the rate laws outlined in the previous section, the production rates of reactants in the network are determined from

$$r_{a,i} := \sum_{j \in \text{RXN}} S_{ij}^{(a)} v_j, \quad i \in \text{MET}, \quad (50a)$$

$$r_{v,i} := \sum_{j \in \text{RXN}} S_{ij}^{(v)} v_j, \quad i \in \text{MET}, \quad (50b)$$

where $S_{ij}^{(a)}$ and $S_{ij}^{(v)}$ are the components of the stoichiometry matrices associated with the membrane and bulk reactions, respectively. Substituting Eq. (50) in Eq. (11), we obtain a closed system of ordinary differential equations from which the time-dependent concentrations of all the reactants in the protometabolic network are determined. This system has 27 variable concentrations and 24 metabolic fluxes. However, the system is not linearly independent when the rate laws outlined in the previous section are used in the mass-balance equations. The linear dependency arises because of the formulation of protoferredoxins in our model. Since the same amount of FD_{rd} is always consumed as FD_{ox} is produced or vice versa, the mass-balance equations for FD_{rd} and FD_{ox} are linearly dependent. To resolve this problem, we replace the mass-balance equation for FD_{ox} with a constraint on the total amount of protoferredoxins that the membrane can contain

$$[\text{FD}_{\text{rd}}] + [\text{FD}_{\text{ox}}] = [\text{FD}], \quad (51)$$

where

$$[\text{FD}] := \frac{\alpha_{\text{FD}} \rho_{\text{FD}}}{\text{MW}_{\text{FD}}}. \quad (52)$$

Here, α_{FD} the mass fraction of protoferredoxins in the membrane, which is a model parameter, with ρ_{FD} and MW_{FD} the density and molecular weight of protoferredoxins. Since the detailed structure and composition of protoferredoxins are unspecified in our model, we use the density and molecular weight of greigite to approximate ρ_{FD} and MW_{FD} .

The main objective of this paper is to answer whether there are conditions under which the protometabolic network can reach steady states in the protocell outlined in Section “Model Description” (see also Fig. 1B), and if so, whether state states are stable. In particular, we seek ranges of physico-chemical parameters for which the mass-balance equations have steady-state solutions. The membrane potential $\Delta\psi$ and characteristic turnover number k_{cat} are the two parameters with respect to which we study the feasibility of the protometabolic network. Accordingly, our goal is to determine whether the system

$$\mathcal{F}(\mathbf{C}, \mathbf{C}^{\text{out}}, \mathbf{K}_m, \Delta\psi, k_{\text{cat}}) = \mathbf{0} \quad (53)$$

has a solution for any ranges of parameters in the $(\Delta\psi, k_{\text{cat}})$ parameter space. Here, \mathcal{F} is a vector, containing the right-hand-sides of mass-balance equations in Eq. (11), \mathbf{C} the vector of concentrations, \mathbf{C}^{out} the vector of extracellular concentrations, and \mathbf{K}_m the vector of Michaelis constants. Since Eq. (53) is a nonlinear system of equations, we formulate a phase-I optimization problem to determine if it has a solution for any values of parameters [78, 79]. Our preliminary numerical experiments indicated that, regarding the extracellular concentrations of reactants, membrane potential, turnover numbers, and Michaelis constants as fixed parameters, the system Eq. (53) is highly constrained for ranges of these parameters that are pertinent to AHV conditions on the early Earth, and in most cases, solving the phase-I problem for Eq. (53) proved futile. Therefore, we sought to find extracellular concentrations and Michaelis constants within reasonable ranges, such that the fundamental constraints of the protocell model are not violated. This alternative viewpoint is motivated by the fact that the exact values of the extracellular concentrations and Michaelis constants of reactants involved in early carbon-fixing networks under the AHV conditions of the early Earth are unknown. Accordingly,

we solve the optimization problem

$$\begin{cases} [\mathbf{C}; \mathbf{C}^{\text{out}}; \mathbf{K}_m](\boldsymbol{\theta}) := \arg \min_{\mathbf{C}, \mathbf{C}^{\text{out}}, \mathbf{K}_m} \|\mathcal{F}(\mathbf{C}, \mathbf{C}^{\text{out}}, \mathbf{K}_m, \boldsymbol{\theta})\|_2 + \omega \mathbf{u}^T \mathbf{C} \\ \text{s.t.} \\ 0 \leq C_i \leq K_i^{\text{sol}}, & i \in \text{MET}, \\ 0 \leq C_i^{\text{out}} \leq K_i^{\text{sol}}, & i \in \text{MET}, \\ K_m^{\text{lb}} \leq K_{m,i} \leq K_m^{\text{ub}}, & i \in \text{MET}, \\ v_j^{\text{lb}} \leq v_j \leq v_j^{\text{ub}}, & j \in \text{RXN}, \end{cases} \quad (54)$$

where $\boldsymbol{\theta} := [\Delta\psi; k_{\text{cat}}]$ with K_i^{sol} the solubility of reactant i in water, K_m^{lb} lower bound on Michaelis constants, K_m^{ub} upper bound on Michaelis constants, v_j^{lb} lower bound on the flux of reaction j , and v_j^{ub} upper bound on the flux of reaction j . We approximate the lower and upper bounds on the Michaelis constants of the protometabolic network based on enzymatic Michaelis constants. Under physiological conditions, enzymatic Michaelis constants vary over a wide range, spanning several orders of magnitude from $\sim 10^{-8}$ M to $\sim 10^{-1}$ M [80]. Using these values as guide, we found steady-state solutions for Eq. (11) within reasonable physical range by choosing $K_m^{\text{lb}} = 10^{-9}$ M and $K_m^{\text{ub}} = 10^{-3}$ M. We also impose upper and lower bounds on some metabolic fluxes to eliminate inconsistent solutions that are not ruled out by the physico-chemical constraints of the protocell model. The only examples are the oxidative decarboxylation of pyruvate to acetate and of α -ketoglutarate to succinate. These steps like others in the protometabolic network are treated as reversible reactions in the protocell model. Although, in principle, they can proceed in the reverse direction without violating the thermodynamic constraints, as were shown for the reduction of acetate to pyruvate [81], protoferredoxins could likely have driven the reverse reactions (*i.e.*, the reductive carboxylation steps) preferentially and more efficiently than the $\text{Fe}^0/\text{Fe}^{2+}$ redox couple due to their physico-chemical characteristics and crystal structures. Accordingly, for these reactions, we impose the lower bound $v_j^{\text{lb}} = 0$ M/s to ensure that the flux from acetate to pyruvate and from succinate to α -ketoglutarate only passes through the reductive carboxylation reactions (Fig. 1C, steps A and G), which are mediated by protoferredoxins, and not the reverse direction of the oxidative decarboxylation steps (Fig. 1C, steps V and X). We impose no upper or lower bound on the metabolic flux of other reactions (*i.e.*, $v_j^{\text{lb}} = -\infty$ and $v_j^{\text{ub}} = \infty$). For these reactions, because the concentrations are bounded, the rate laws automatically ensure that the metabolic fluxes are also bounded.

Although Eq. (53) is a balanced system of equations, it could have multiple solutions for a given set of parameters because it is a nonlinear system. Therefore, we incorporate a regularizer into the optimization problem Eq. (54) that allows us to choose a unique solution based on a given criterion. In the objective function of Eq. (54), the first term measures deviations from the steady-state mass-balance constraints Eq. (53) and the second term is a linear regularizer, which penalizes deviations from the criterion based on which a unique solution is to be determined. Here, ω is the weight of the regularizer and \mathbf{u} is a constant vector that specifies the criterion. For example, through numerical experiments, we found that choosing a regularizer that maximizes the concentration of α -ketoglutarate leads to steady-state solutions that correspond to a complete-cycle flux in most cases.

We determine the stability of steady-state solutions obtained from Eq. (54) by com-

Table S3: Model parameters.

Parameter	Value	Parameter	Value
T_{out}	150° C	R_c	10^{-7} m
T_{in}	90° C	d	2×10^{-8} m
pH_{out}	9	$[\text{Cr}^{3+}]$	10^{-3} M
pH_{in}	5	α_{FD}	0.8
I_{out}	0.3 M	β_{aldol}	0.8
I_{in}	0.3 M	$\beta_{\text{retro-aldol}}$	0.1
$J_{\text{Fe}^{2+}}^0$	10^{-2} mol/(m ² s)	ϑ	10^{-5}
ΔE_{redox}	1 V	C_t^{out}	0.4 M

puting the eigenvalues of the Jacobian matrix \mathcal{J} defined as

$$\mathcal{J}_{ij} := \frac{\partial \mathcal{F}_i(\mathbf{C}, \mathbf{C}^{\text{out}}, \mathbf{K}_m, \Delta\psi, k_{\text{cat}})}{\partial C_j}, \quad i, j \in \{\text{MET}\}. \quad (55)$$

Note that, once we have found \mathbf{C}^{out} and \mathbf{K}_m that allow the protometabolic network to attain a steady state, we treat them as fixed parameters in the stability analysis. According to linear stability theory [82], a steady-state solution of Eq. (11) is linearly stable if the eigenvalues of the Jacobian matrix all have negative real parts. Otherwise, the steady-state solution is unstable.

A stable steady-state solution of Eq. (11) is an interior point of a stability region in the parameter space. To determine parts of the $(\Delta\psi, k_{\text{cat}})$ parameter space in which steady-state solutions are stable and those in which steady-state solutions are unstable, we use branch-continuation methods [82]. Accordingly, we compute steady-state solutions of Eq. (11) over a nonuniform grid in the parameter space by performing branch continuation along $\Delta\psi$ - and k_{cat} -directions from a starting interior point provided by the optimization problem Eq. (54). Steady-state solution branches along $\Delta\psi$ - and k_{cat} -directions are obtained by solving

$$\mathcal{F}(\mathbf{C}(s), \mathbf{C}^{\text{out}}, \mathbf{K}_m, \Delta\psi(s), k_{\text{cat}}) = \mathbf{0} \quad (56)$$

and

$$\mathcal{F}(\mathbf{C}(s), \mathbf{C}^{\text{out}}, \mathbf{K}_m, \Delta\psi, k_{\text{cat}}(s)) = \mathbf{0} \quad (57)$$

using arclength continuation methods, respectively. Here, s denotes the arclength along solution the respective branch. We refer the reader to the work of Seydel [82] for a comprehensive overview of arclength-continuation methods, and how solution branches are constructed from Eqs. (56) and (57). Using these techniques, we constructed the stability region discussed in the main article (see Fig. 2A) and characterized the bifurcation of steady-state solution branches along its boundaries. We found that the stability region is bounded in the membrane potential and unbounded in the turnover number in the ranges of parameters that we studied. Specifically, solution branches along $\Delta\psi$ -directions bifurcate and lose stability at both small and large membrane potentials but not along k_{cat} -directions. The stability-region boundary at small membrane potentials correspond to turning points and at large membrane potentials to simple branch points. The Jacobian matrix is singular at both boundaries, that is $\text{rank}(\mathcal{F}_{\mathbf{C}}) = n - 1$ with $\mathcal{J} := \mathcal{F}_{\mathbf{C}}$ (subscript indicates the variables with respect to which derivative is taken; see Eq. (55)) and $n := |\text{MET}|$ the number of variable concentrations. We can characterize the type

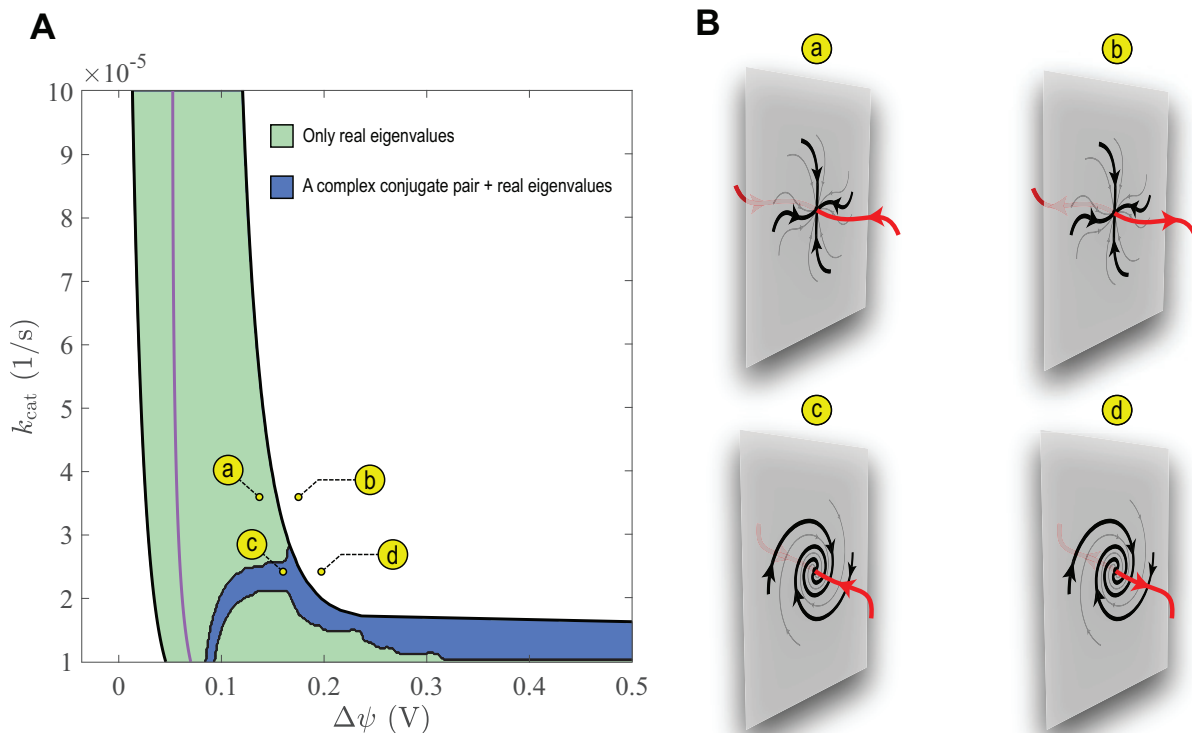


Figure S6: **Near-steady-state dynamics of the protocell model.** (A) Characteristics of the eigenvalues of the Jacobian matrix in Eq. (55) in the stability region of the protocell discussed in the main article (see Fig. 2A). (B) Schematic representation of dynamic trajectories near steady-state solutions in a three-dimensional projection of the phase space at selected points of the parameter space.

of bifurcation along the stability-region boundaries based on the rank of the following concatenated matrix for solution branches that are constructed from Eq. (56)

$$\Xi := [\mathcal{F}_C | \mathcal{F}_{\Delta\psi}]. \quad (58)$$

Accordingly, $\text{rank}(\Xi) = n$ at turning points and $\text{rank}(\Xi) = n - 1$ at simple branch points (see Definitions 2.7 and 2.8 of Seydel [82]). We can follow a similar approach for solution branches that are constructed from Eq. (57). We also characterized the eigenvalues of the Jacobian matrix throughout the stability region and found that dynamic trajectories around steady-state solutions can exhibit oscillatory or nonoscillatory behaviors at different parts of the stability region (Fig. S6). All the model parameters we used for computations performed in this paper are summarized in Table S3.

The solubility of reactants in Eq. (54) were computed with Calculator Plugins, Marvin (version 20.21.0, 2020), ChemAxon (<http://www.chemaxon.com>). We formulated the global optimization problem Eq. (54) in GAMS (version 28.2.0, 2019) and solved it using the global solver BARON (version 19.7.13, 2019) equipped with the local nonlinear programming solver CONOPT (version 3.17K, 2019) and the linear programming solver CPLEX (version 12.9.0, 2019). Branch-continuation problems Eqs. (56) and (57) were solved using custom codes in MATLAB (version R2019b). All the custom codes written in GAMS and MATLAB to generate the results presented in this paper are provided in Supplementary Code 1.

Osmotic Pressure Drives Protocell Growth

As we stated in the main article, the membrane potential is the main driving force of growth in our model. The membrane potential is positive due to the direction of the pH gradient across the protocell membrane. Because the positive membrane potential nonuniformly distributes ions across the membrane, an osmotic pressure differential is developed in the protocell. Note that the osmotic pressure differential is generated because of the uneven movements of ions to one side of the membrane whether or not metabolic reactions occur inside the protocell. Accordingly, our goal in this section is to estimate the change in the osmotic pressure differential as a result of metabolic reactions.

We approximate the osmotic pressure of an aqueous mixture of solutes in water using a truncated Pitzer model [83]

$$\Pi = RTC_t + \mathcal{O}(C_i), \quad (59)$$

where Π is the osmotic pressure of the aqueous solution of interest, C_t total solute concentrations, and C_i passed to the order-of-magnitude operator \mathcal{O} is a representative solute concentration. Following the approach we introduced in our earlier work [1], we partition all solutes that contribute to the osmotic pressure into reactive and nonreactive solutes. Since Eq. (59) is linear in solute concentrations, we can decompose the total osmotic pressure into a component induced by reactive and another induced by nonreactive solutes and compute each component separately. Let C_t^\dagger and C_t^* denote the total concentration of reactive and nonreactive solutes in the intracellular environment. Let also $C_t^{\text{out},\dagger}$ and $C_t^{\text{out},*}$ denote the same quantities in the extracellular environment. Accordingly, we can consider the following decompositions

$$C_t = C_t^\dagger + C_t^*, \quad (60a)$$

$$C_t^{\text{out}} = C_t^{\text{out},\dagger} + C_t^{\text{out},*}. \quad (60b)$$

First, we estimate the osmotic pressure differential due to nonreactive species. To simplify the analysis, we assume that the contribution of reactive solutes to the osmotic pressure of the extracellular environment is negligible compared to that of nonreactive species, so that $C_t^{\text{out},\dagger} \ll C_t^{\text{out},*}$ and $C_t^{\text{out}} \approx C_t^{\text{out},*}$. We also assume that the extracellular environment is electroneutral and contains equal amounts of negatively and positively charged solutes, which are all monovalent. Thus,

$$C_{t,+}^{\text{out},*} = C_{t,-}^{\text{out},*} = \frac{1}{2}C_t^{\text{out},*} = \frac{1}{2}C_t^{\text{out}}$$

with $C_{t,+}^{\text{out},*}$ and $C_{t,-}^{\text{out},*}$ the total concentration of positively and negatively charged solutes in the extracellular environment. At steady state, the membrane transport rate of nonreactive species is zero [1]. Using this property and the expression for diffusive membrane transport rate given by Eq. (6), we can ascertain the total concentration of positively and negatively charged solutes in the intracellular environments

$$C_{t,+}^* = \exp(\Delta\hat{\psi})^{-1}C_{t,+}^{\text{out},*} \quad (61a)$$

$$C_{t,-}^* = \exp(\Delta\hat{\psi})C_{t,-}^{\text{out},*} \quad (61b)$$

from which it follows that

$$C_t^* = C_{t,+}^* + C_{t,-}^* = \Phi C_t^{\text{out}}, \quad (62)$$

where

$$\Phi := \frac{\exp(\Delta\hat{\psi})^2 + 1}{2\exp(\Delta\hat{\psi})} \quad (63)$$

with $C_{t,+}^*$ and $C_{t,-}^*$ the total concentration of positively and negatively charged solutes in the intracellular environment. Therefore, the contribution of nonreactive solutes to the osmotic pressure differential can be obtained using Eq. (59)

$$\Delta\Pi^0 = RT(C_t^* - C_t^{\text{out},*}) = RT(\Phi - 1)C_t^{\text{out}}. \quad (64)$$

Next, we obtain the following expression for the osmotic pressure differential due to both reactive and nonreactive solutes using Eq. (59)

$$\Delta\Pi = RT(C_t - C_t^{\text{out}}) = RT[C_t^\dagger + (\Phi - 1)C_t^{\text{out}}]. \quad (65)$$

The iron-sulfide membrane in our protocell model is rigid. Therefore, an increase in the osmotic pressure differential due to the accumulation of protometabolic network products, as measured by $\Delta\Pi/\Delta\Pi^0$, results in an excess in-plan stress throughout the membrane without changing the volume of the protocell. However, as we previously discussed (see ‘‘Model Description’’ and Fig. S2), at early stages of evolution, lipid membranes could have replaced iron-sulfide membranes once lipid-synthesis pathways had been incorporated into first carbon-fixing cycles. In these early cell-like structures with stretchable membranes, an increase in the osmotic pressure differential could have resulted in a volume expansion and been a driving force of growth. Therefore, it is useful to measure the ratio $\Delta\Pi/\Delta\Pi^0$ in terms of the volume expansion that it would cause in a deformable cell as it is more relevant to selective processes underlying transitions from iron-sulfide to lipid membranes. To this end, we compare the protocell model, which has a rigid membrane, to a hypothetical protocell with a deformable membrane. Let R_c^0 denote the radius of the protocell model with \bar{C}_t , \bar{C}_t^\dagger , and \bar{C}_t^* the total concentration of all solutes, reactive solutes, and nonreactive solutes in its intracellular environment. We then consider a hypothetical protocell with the same radius R_c^0 as the protocell model, only containing the nonreactive solutes of the protocell model. We assume that the membrane of the hypothetical protocell is elastic with a stretching rigidity γ . Regarding the total concentration of nonreactive solutes in the extracellular environment C_t^{out} as a fixed parameter (recall, we neglected $C_t^{\text{out},\dagger}$ compared to $C_t^{\text{out},*}$, so the total concentration of nonreactive solutes and the total concentration of all solutes in the extracellular environment are the same, that is $C_t^{\text{out},*} \approx C_t^{\text{out}}$), the total concentration of nonreactive solutes in the intracellular environment of the hypothetical protocell is the same as that of the protocell model, which is $\bar{C}_t^* = \Phi C_t^{\text{out}}$ (see Eq. (62)). The goal here is to determine how much this hypothetical protocell would expand, if it contained the total amount of the reactive solutes of the protocell model \bar{C}_t^\dagger .

To determine how much the volume of the hypothetical protocell expands as a result of the accumulation of the reactive solutes, we introduce two states: (i) The base state, which is the hypothetical protocell with an elastic membrane of stretching rigidity γ and radius R_c^0 . In this state, the total amount of solutes in the intracellular environment is $V_0\bar{C}_t^*$, the osmotic pressure differential is denoted $\Delta\Pi^0$, and $V_0 := 4\pi(R_c^0)^3/3$. (ii) The expanded state, which is the hypothetical protocell with an elastic membrane of the same stretching rigidity as in the base state and radius $R_c > R_c^0$, containing both the reactive and nonreactive solutes of the protocell model. In this case, the total amount of solutes in the intracellular environment is $V_0(\bar{C}_t^* + \bar{C}_t^\dagger)$ and the osmotic pressure differential is

denoted $\Delta\Pi$. Therefore, the total concentration of solutes in the intracellular environment is $C_t = V_0(\bar{C}_t^* + \bar{C}_t^\dagger)/V$ with $V := 4\pi R_c^3/3$, so we have

$$C_t = (R_c^0/R_c)^3(\bar{C}_t^* + \bar{C}_t^\dagger), \quad (66)$$

where C_t is the total concentration of solutes in the intracellular environment for the expanded state.

Assuming that the membrane in the hypothetical protocell is thin and elastic, the osmotic pressure differential is inversely proportional to the protocell radius [84]. Implementing this proportionality in Eqs. (64) and (65) yields

$$\frac{v\gamma}{R_c^0} = \Delta\Pi^0 = RT(\Phi - 1)C_t^{\text{out}}, \quad (67)$$

$$\frac{v\gamma}{R_c} = \Delta\Pi = RT(C_t - C_t^{\text{out}}) \quad (68)$$

with v the proportionality constant. Solving Eq. (67) for γ and substituting it in Eq. (68) along with C_t from Eq. (66) results in the cubic equation

$$(\bar{C}_t^\dagger + \Phi C_t^{\text{out}})X^3 - (\Phi - 1)C_t^{\text{out}}X - C_t^{\text{out}} = 0. \quad (69)$$

Here, $X := R_c^0/R_c$ is the ratio of the radii of the base and expanded states, serving as a measure for volume expansion due to the accumulation of reactive solutes in the protocell. Note that C_t^{out} in this equation is a fixed parameter, Φ is a function of $\Delta\psi$, and \bar{C}_t^\dagger is determined from the steady-state solutions of the mass-balance equations Eq. (11). Therefore, for any point in the $(\Delta\psi, k_{\text{cat}})$ parameter space, there is a corresponding \bar{C}_t^\dagger and X . Accordingly, as a specific example, we computed X throughout the stability region of the protocell model in the parameter space to gain general insights into the driving forces of growth for compartmentalized and self-sustaining carbon-fixing cycles at the origin of metabolism. The results are shown in Fig. 3A of the main article for $C_t^{\text{out}} = 0.4$ M.

References

1. Akbari, A., Palsson, B.O.. Positively charged mineral surfaces promoted the accumulation of organic intermediates at the origin of metabolism. *PLoS Comput Biol* 2022;18(8):e1010377. doi:10.1371/journal.pcbi.1010377.
2. Wächtershäuser, G.. Before enzymes and templates: Theory of surface metabolism. *Microbiol Rev* 1988;52(4):452–484. doi:10.1128/mmbr.52.4.452-484.1988.
3. Morowitz, H.J.. A theory of biochemical organization, metabolic pathways, and evolution. *Complexity* 1999;4(6):39–53. doi:10.1002/(SICI)1099-0526(199907/08)4:6;39::AID-CPLX8;3.0.CO;2-2.
4. Dyson, F.. Origins of life. Cambridge University Press; 1999.
5. Goldford, J.E., Hartman, H., Smith, T.F., Segrè, D.. Remnants of an ancient metabolism without phosphate. *Cell* 2017;168(6):1126–1134. doi:10.1016/j.cell.2017.02.001.

6. Xavier, J.C., Hordijk, W., Kauffman, S., Steel, M., Martin, W.F.. Autocatalytic chemical networks at the origin of metabolism. *Proc Royal Soc B* 2020;287(1922):20192377. doi:10.1098/rspb.2019.2377.
7. Huber, C., Wächtershäuser, G.. Activated acetic acid by carbon fixation on (Fe,Ni)S under primordial conditions. *Science* 1997;276(5310):245–247. doi:10.1126/science.276.5310.245.
8. Preiner, M., Igarashi, K., Muchowska, K.B., Yu, M., Varma, S.J., Kleinermanns, K., Nobu, M.K., Kamagata, Y., Tüysüz, H., Moran, J., et al. A hydrogen-dependent geochemical analogue of primordial carbon and energy metabolism. *Nat Ecol Evol* 2020;4(4):534–542. doi:10.1038/s41559-020-1125-6.
9. Beyazay, T., Belthle, K.S., Farès, C., Preiner, M., Moran, J., Martin, W.F., Tüysüz, H.. Ambient temperature co2 fixation to pyruvate and subsequently to citramalate over iron and nickel nanoparticles. *Nat Commun* 2023;14(1):570. doi:10.1038/s41467-023-36088-w.
10. Muchowska, K.B., Varma, S.J., Chevallot-Beroux, E., Lethuillier-Karl, L., Li, G., Moran, J.. Metals promote sequences of the reverse Krebs cycle. *Nat Ecol Evol* 2017;1(11):1716–1721. doi:10.1038/s41559-017-0311-7.
11. Muchowska, K.B., Varma, S.J., Moran, J.. Synthesis and breakdown of universal metabolic precursors promoted by iron. *Nature* 2019;569(7754):104–107. doi:10.1038/s41586-019-1151-1.
12. Stubbs, R.T., 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(11):1016–1022. doi:10.1038/s41557-020-00560-7.
13. Rauscher, S.A., Moran, J.. Hydrogen drives part of the reverse krebs cycle under metal or meteorite catalysis. *Angew Chem Int Ed* 2022;;e202212932doi:10.1002/anie.202212932.
14. Keller, M.A., Turchyn, A.V., Ralser, M.. Non-enzymatic glycolysis and pentose phosphate pathway-like reactions in a plausible archaean ocean. *Mol Syst Biol* 2014;10(4):725. doi:10.1002/msb.20145228.
15. Piedrafita, G., Varma, S.J., Castro, C., Messner, C.B., Szyrwił, L., Griffin, J.L., Ralser, M.. Cysteine and iron accelerate the formation of ribose-5-phosphate, providing insights into the evolutionary origins of the metabolic network structure. *PLoS Biol* 2021;19(12):e3001468. doi:10.1371/journal.pbio.3001468.
16. Mayer, R.J., Kaur, H., Rauscher, S.A., Moran, J.. Mechanistic insight into metal ion-catalyzed transamination. *J Am Chem Soc* 2021;143(45):19099–19111. doi:10.1021/jacs.1c08535.
17. Yi, J., Kaur, H., Kazöne, W., Rauscher, S.A., Gravillier, L.A., Muchowska, K.B., Moran, J.. A nonenzymatic analog of pyrimidine nucleobase biosynthesis. *Angew Chem Int Ed* 2022;;e202117211doi:10.1002/anie.202117211.

18. Messner, C.B., Driscoll, P.C., Piedrafita, G., De Volder, M.F., Ralser, M.. Nonenzymatic gluconeogenesis-like formation of fructose 1,6-bisphosphate in ice. *Proc Natl Acad Sci* 2017;114(28):7403–7407. doi:10.1073/pnas.1702274114.
19. Russell, M.J., Hall, A.J.. The emergence of life from iron monosulphide bubbles at a submarine hydrothermal redox and pH front. *J Geol Soc* 1997;154(3):377–402. doi:10.1144/gsjgs.154.3.0377.
20. Lane, N.. The vital question: Energy, evolution, and the origins of complex life. WW Norton & Company; 2015.
21. Cody, G.D., Boctor, N.Z., Filley, T.R., Hazen, R.M., Scott, J.H., Sharma, A., Yoder Jr, H.S.. Primordial carbonylated iron-sulfur compounds and the synthesis of pyruvate. *Science* 2000;289(5483):1337–1340. doi:10.1126/science.289.5483.1337.
22. Marín-Yaseli, M.R., González-Toril, E., Mompeán, C., Ruiz-Bermejo, M.. The role of aqueous aerosols in the “glyoxylate scenario”: An experimental approach. *Chem Eur J* 2016;22(36):12785–12799. doi:10.1002/chem.201602195.
23. Huber, C., Wächtershäuser, G.. α -hydroxy and α -amino acids under possible hadean, volcanic origin-of-life conditions. *Science* 2006;314(5799):630–632. doi:10.1126/science.1130895.
24. Swami, P., Yajurvedi, D., Mishra, P., Sharma, P.K.. Oxidation of some α -hydroxy acids by tetraethylammonium chlorochromate: A kinetic and mechanistic study. *Int J Chem Kinet* 2010;42(1):50–55. doi:10.1002/kin.20466.
25. de Duve, C., de Neufville, R.. Blueprint for a cell: the nature and origin of life. Carolina Biological Supply Company; 1991.
26. Baaske, P., Weinert, F.M., Duhr, S., Lemke, K.H., Russell, M.J., Braun, D.. Extreme accumulation of nucleotides in simulated hydrothermal pore systems. *Proc Natl Acad Sci* 2007;104(22):9346–9351. doi:10.1073/pnas.0609592104.
27. Camprubi, E., Jordan, S.F., Vasiliadou, R., Lane, N.. Iron catalysis at the origin of life. *IUBMB Life* 2017;69(6):373–381. doi:10.1002/iub.1632.
28. Kitadai, N., Nakamura, R., Yamamoto, M., Takai, K., Yoshida, N., Oono, Y.. Metals likely promoted protometabolism in early ocean alkaline hydrothermal systems. *Sci Adv* 2019;5(6):eaav7848. doi:10.1126/sciadv.aav7848.
29. Martin, W., Russell, M.J.. On the origin of biochemistry at an alkaline hydrothermal vent. *Philos Trans R Soc B* 2007;362(1486):1887–1926. doi:10.1098/rstb.2006.1881.
30. Russell, M.J., Hall, A.J., Turner, D.. In vitro growth of iron sulphide chimneys: possible culture chambers for origin-of-life experiments. *Terra Nova* 1989;1(3):238–241. doi:10.1111/j.1365-3121.1989.tb00364.x.
31. Sojo, V., Herschy, B., Whicher, A., Camprubi, E., Lane, N.. The origin of life in alkaline hydrothermal vents. *Astrobiology* 2016;16(2):181–197. doi:10.1089/ast.2015.1406.

32. Jordan, S.F., Ramm, H., Zheludev, I.N., Hartley, A.M., Maréchal, A., Lane, N.. Promotion of protocell self-assembly from mixed amphiphiles at the origin of life. *Nat Ecol Evol* 2019;3(12):1705–1714. doi:10.1038/s41559-019-1015-y.
33. Hanczyc, M.M., Fujikawa, S.M., Szostak, J.W.. Experimental models of primitive cellular compartments: Encapsulation, growth, and division. *Science* 2003;302(5645):618–622. doi:10.1126/science.1089904.
34. Muchowska, K.B., Varma, S.J., Moran, J.. Nonenzymatic metabolic reactions and life’s origins. *Chem Rev* 2020;120(15):7708–7744. doi:10.1021/acs.chemrev.0c00191.
35. Harrison, S.A., Lane, N.. Life as a guide to prebiotic nucleotide synthesis. *Nat Commun* 2018;9(1):1–4. doi:10.1038/s41467-018-07220-y.
36. Wächtershäuser, G.. Evolution of the first metabolic cycles. *Proc Natl Acad Sci* 1990;87(1):200–204. doi:10.1073/pnas.87.1.200.
37. Smith, E., Morowitz, H.J.. Universality in intermediary metabolism. *Proc Natl Acad Sci* 2004;101(36):13168–13173. doi:10.1073/pnas.0404922101.
38. Morowitz, H.J., Srinivasan, V., Smith, E.. Ligand field theory and the origin of life as an emergent feature of the periodic table of elements. *Biol* 2010;219(1):1–6. doi:10.1086/bblv219n1p1.
39. Cody, G.D., Boctor, N.Z., Hazen, R.M., Brandes, J.A., Morowitz, H.J., Yoder Jr, H.S.. Geochemical roots of autotrophic carbon fixation: Hydrothermal experiments in the system citric acid, H_2O -($\pm\text{Fe}$)-($\pm\text{Ni}$). *Geochim Cosmochim Acta* 2001;65(20):3557–3576. doi:10.1016/S0016-7037(01)00674-3.
40. Iimura, S., Manabe, K., Kobayashi, S.. Direct thioesterification from carboxylic acids and thiols catalyzed by a brønsted acid. *Chem Commun* 2002;(1):94–95. doi:10.1039/B109834A.
41. Sanden, S.A., Yi, R., Hara, M., McGlynn, S.E.. Simultaneous synthesis of thioesters and iron–sulfur clusters in water: Two universal components of energy metabolism. *Chem Commun* 2020;56(80):11989–11992. doi:10.1039/d0cc04078a.
42. Nunoura, T., Chikaraishi, Y., Izaki, R., Suwa, T., Sato, T., Harada, T., Mori, K., Kato, Y., Miyazaki, M., Shimamura, S., et al. A primordial and reversible TCA cycle in a facultatively chemolithoautotrophic thermophile. *Science* 2018;359(6375):559–563. doi:10.1126/science.aao3407.
43. Nanda, V., Senn, S., Pike, D.H., Rodriguez-Granillo, A., Hansen, W.A., Khare, S.D., Noy, D.. Structural principles for computational and de novo design of 4Fe–4S metalloproteins. *Biochim Biophys Acta* 2016;1857(5):531–538. doi:10.1016/j.bbabi.2015.10.001.
44. Swartz, P.D., Beck, B.W., Ichiye, T.. Structural origins of redox potentials in Fe-S proteins: electrostatic potentials of crystal structures. *Biophys J* 1996;71(6):2958–2969. doi:10.1016/s0006-3495(96)79533-4.

45. Scintilla, S., Bonfio, C., Belmonte, L., Forlin, M., Rossetto, D., Li, J., Cowan, J.A., Galliani, A., Arnesano, F., Assfalg, M., et al. Duplications of an iron-sulphur tripeptide leads to the formation of a protoferredoxin. *Chem Commun* 2016;52(92):13456–13459. doi:10.1039/c6cc07912a.
46. Thiel, J., Byrne, J.M., Kappler, A., Schink, B., Pester, M.. Pyrite formation from fes and h2s is mediated through microbial redox activity. *Proc Natl Acad Sci* 2019;116(14):6897–6902. doi:10.1073/pnas.1814412116.
47. Wang, W., Song, Y., Wang, X., Yang, Y., Liu, X.. Alpha-oxo acids assisted transformation of FeS to Fe3S4 at low temperature: implications for abiotic, biotic, and prebiotic mineralization. *Astrobiology* 2015;15(12):1043–1051. doi:10.1089/ast.2015.1373.
48. Eck, R.V., Dayhoff, M.O.. Evolution of the structure of ferredoxin based on living relics of primitive amino acid sequences. *Science* 1966;152(3720):363–366. doi:10.1126/science.152.3720.363.
49. Jordan, S.F., Ioannou, I., Ramm, H., Halpern, A., Bogart, L.K., Ahn, M., Vasiliadou, R., Christodoulou, J., Maréchal, A., Lane, N.. Spontaneous assembly of redox-active iron-sulfur clusters at low concentrations of cysteine. *Nat Commun* 2021;12(1):1–14. doi:10.1038/s41467-021-26158-2.
50. Huber, C., Wächtershäuser, G.. Primordial reductive amination revisited. *Tetrahedron Lett* 2003;44(8):1695–1697. doi:10.1016/s0040-4039(02)02863-0.
51. Barge, L.M., Flores, E., Baum, M.M., VanderVelde, D.G., Russell, M.J.. Redox and pH gradients drive amino acid synthesis in iron oxyhydroxide mineral systems. *Proc Natl Acad Sci* 2019;116(11):4828–4833. doi:10.1073/pnas.1812098116.
52. Kauffman, S.A.. The origins of order: Self-organization and selection in evolution. Oxford University Press, USA; 1993.
53. Eschenmoser, A.. On a hypothetical generational relationship between HCN and constituents of the reductive citric acid cycle. *Chem Biodivers* 2007;4(4):554–573. doi:10.1002/cbdv.200790050.
54. Schwartz, A.W.. Phosphorus in prebiotic chemistry. *Philos Trans R Soc Lond, B, Biol Sci* 2006;361(1474):1743–1749. doi:10.1098/rstb.2006.1901.
55. Brady, M.P., Tostevin, R., Tosca, N.J.. Marine phosphate availability and the chemical origins of life on earth. *Nat Commun* 2022;13(1):5162. doi:10.1038/s41467-022-32815-x.
56. Pinna, S., Kunz, C., Halpern, A., Harrison, S.A., Jordan, S.F., Ward, J., Werner, F., Lane, N.. A prebiotic basis for atp as the universal energy currency. *PLoS Biol* 2022;20(10):e3001437. doi:10.1371/journal.pbio.3001437.
57. Alberty, R.. Thermodynamics of biochemical reactions. John Wiley & Sons; 2005.
58. Akbari, A., Yurkovich, J.T., Zielinski, D.C., Palsson, B.O.. The quantitative metabolome is shaped by abiotic constraints. *Nat Commun* 2021;12(1):1–19. doi:10.1038/s41467-021-23214-9.

59. Stein, W.. Transport and diffusion across cell membranes. Elsevier; 2012.
60. Henry, C.S., Broadbelt, L.J., Hatzimanikatis, V.. Thermodynamics-based metabolic flux analysis. *Biophys J* 2007;92(5):1792–1805. doi:10.1529/biophysj.106.093138.
61. Noor, E., Flamholz, A., Liebermeister, W., Bar-Even, A., Milo, R.. A note on the kinetics of enzyme action: a decomposition that highlights thermodynamic effects. *FEBS Lett* 2013;587(17):2772–2777. doi:10.1016/j.febslet.2013.07.028.
62. Beber, M.E., Gollub, M.G., Mozaffari, D., Shebek, K.M., Flamholz, A.I., Milo, R., Noor, E.. equilibrator 3.0: a database solution for thermodynamic constant estimation. *Nucleic Acids Res* 2022;50(D1):D603–D609. doi:10.1093/nar/gkab1106.
63. Wang, L., Upadhyay, V., Maranas, C.D.. dgpredictor: Automated fragmentation method for metabolic reaction free energy prediction and de novo pathway design. *PLoS Comput Biol* 2021;17(9):e1009448. doi:10.1371/journal.pcbi.1009448.
64. Izawa, M.R.M., Nesbitt, H.W., MacRae, N.D., Hoffman, E.L.. Composition and evolution of the early oceans: evidence from the tagish lake meteorite. *Earth Planet Sci Lett* 2010;298(3-4):443–449. doi:10.1016/j.epsl.2010.08.026.
65. Du, B., Zhang, Z., Grubner, S., Yurkovich, J.T., Palsson, B.O., Zielinski, D.C.. Temperature-dependent estimation of gibbs energies using an updated group-contribution method. *Biophys J* 2018;114(11):2691–2702. doi:10.1016/j.bpj.2018.04.030.
66. Trost, B.M., Brindle, C.S.. The direct catalytic asymmetric aldol reaction. *Chem Soc Rev* 2010;39(5):1600–1632. doi:10.1039/b923537j.
67. Yamashita, Y., Yasukawa, T., Yoo, W.J., Kitanosono, T., Kobayashi, S.. Catalytic enantioselective aldol reactions. *Chem Soc Rev* 2018;47(12):4388–4480. doi:10.1039/c7cs00824d.
68. Murakami, K., Ohmiya, H., Yorimitsu, H., Oshima, K.. Generation of rhodium enolates via retro-aldol reaction and its application to regioselective aldol reaction. *Tetrahedron Lett* 2008;49(15):2388–2390. doi:10.1016/j.tetlet.2008.02.062.
69. Brack, A.. The molecular origins of life: Assembling pieces of the puzzle. Cambridge University Press; 1998.
70. Wächtershäuser, G.. Groundworks for an evolutionary biochemistry: The iron-sulphur world. *Prog Biophys Molec Biol* 1992;58(2):85–201. doi:10.1016/0079-6107(92)90022-X.
71. Udupa, A.R., Smith, K.A., Moore, J.J.. Lime-enhanced reduction of sulfide concentrates: a thermodynamic discussion. *Metall Trans B* 1986;17(1):185–196. doi:10.1007/bf02670832.
72. Keller, M.A., Kampjut, D., Harrison, S.A., Ralser, M.. Sulfate radicals enable a non-enzymatic krebs cycle precursor. *Nat Ecol Evol* 2017;1(4):1–9. doi:10.1038/s41559-017-0083.

73. Wolfenden, R.. Benchmark reaction rates, the stability of biological molecules in water, and the evolution of catalytic power in enzymes. *Annu Rev Biochem* 2011;80:645–667. doi:10.1146/annurev-biochem-060409-093051.
74. Bar-Even, A., Noor, E., Savir, Y., Liebermeister, W., Davidi, D., Tawfik, D.S., Milo, R.. The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry* 2011;50(21):4402–4410. doi:10.1021/bi2002289.
75. Chandru, K., Gilbert, A., Butch, C., Aono, M., Cleaves, H.J.. The abiotic chemistry of thiolated acetate derivatives and the origin of life. *Sci Rep* 2016;6(1):29883. doi:10.1038/srep29883.
76. Herschy, B.. Nature’s electrochemical flow reactors?: Alkaline hydrothermal vents and the origins of life. *The Biochemist* 2014;36(6):4–8. doi:10.1042/bio03606004.
77. Grajales-Gonzalez, E., Monge-Palacios, M., Sarathy, S.M.. Collision efficiency parameter influence on pressure-dependent rate constant calculations using the ss-qrrk theory. *J Phys Chem A* 2020;124(31):6277–6286. doi:10.1021/acs.jpca.0c02943.
78. Nocedal, J., Wright, S.. Numerical optimization. Springer Science & Business Media; 2006.
79. Akbari, A., Palsson, B.O.. Scalable computation of intracellular metabolite concentrations. *Comput Chem Eng* 2021;145:107164. doi:10.1016/j.compchemeng.2020.107164.
80. Bennett, B.D., Kimball, E.H., Gao, M., Osterhout, R., Van Dien, S.J., Rabinowitz, J.D.. Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat Chem Biol* 2009;5(8):593. doi:10.1038/nchembio.186.
81. Varma, S.J., Muchowska, K.B., Chatelain, P., Moran, J.. Native iron reduces CO₂ to intermediates and end-products of the acetyl-coa pathway. *Nat Ecol Evol* 2018;2(6):1019–1024. doi:10.1038/s41559-018-0542-2.
82. Seydel, R.. Practical bifurcation and stability analysis; vol. 5. Springer Science & Business Media; 2009.
83. Pitzer, K., Kim, J.. Thermodynamics of electrolytes. IV. Activity and osmotic coefficients for mixed electrolytes. *J Am Chem Soc* 1974;96(18). doi:10.1021/ja00825a004.
84. Timoshenko, S., Woinowsky-Krieger, S.. Theory of plates and shells; vol. 2. McGraw-hill New York; 1959.