Catalysts, autocatalysis and the origin of metabolism

Martina Preiner1, Joana C. Xavier1, Andrey do Nascimento Vieira1, Karl Kleinermanns2, John F. Allen3 and William F. Martin1

1Institute for Molecular Evolution, and 2Institute for Physical Chemistry, Heinrich-Heine-University, 40225 Düsseldorf, Germany
3Research Department of Genetics, Evolution and Environment, University College London, Darwin Building, Gower Street, London WC1E 6BT, UK

If life on Earth started out in geochemical environments like hydrothermal vents, then it started out from gasses like CO2, N2 and H2. Anaerobic autotrophs still live from these gasses today, and they still inhabit the Earth’s crust. In the search for connections between abiotic processes in ancient geological systems and biotic processes in biological systems, it becomes evident that chemical activation (catalysis) of these gasses and a constant source of energy are key. The H2–CO2 redox reaction provides a constant source of energy and anabolic inputs, because the equilibrium lies on the side of reduced carbon compounds. Identifying geochemical catalysts that activate these gasses en route to nitrogenous organic compounds and small autocatalytic networks will be an important step towards understanding prebiotic chemistry that operates only on the basis of chemical energy, without input from solar radiation. So, if life arose in the dark depths of hydrothermal vents, then understanding reactions and catalysts that operate under such conditions is crucial for understanding origins.

1. Introduction

When the Earth was formed 4.5 billion years ago, it was formed without life, we can safely presume. If there was any life on the freshly accreted Earth, it was destroyed at the moon forming impact, which converted the Earth into a ball of boiling magma [1]. By about 3.95 billion years ago, there was life on Earth [2]. The question of how it arose is of substantial interest. Hydrothermal vents play an important role in the question of life’s origin, because they were present on the early Earth [3–7] and because they harbour continuously far-from-equilibrium conditions in an environment where H2 and CO2 interact in such a way as to generate reduced carbon compounds [8–15]. In the discussion of possible sites for life’s origin, hydrothermal vents are unique by that criterion: hydrothermal vents harbour far from equilibrium conditions over geological timescales, and the approach towards equilibrium releases energy in the synthesis of reduced carbon compounds. This sets hydrothermal vents apart from all other physicochemical settings [16]. Moreover, the release of free energy and the synthesis of reduced carbon compounds at vents are united in a common reaction sequence that operates in the laboratory without enzymes [15] and that is simultaneously the core of carbon and energy metabolism in real bacteria and archaea—acetogens and methanogens. Vents are unique among settings for the origin of metabolism (as opposed to the origin of life), because no other site for life’s origin harbours chemical reactions that resemble real microbial carbon and energy metabolism.

The far-from-equilibrium conditions at alkaline hydrothermal vents entail steep redox gradients owing to a constant flux of H2-rich effluent over geological timescales [17]. The main redox reaction they harbour is the H2–CO2 system,
in which the equilibrium lies far on the side of organic compounds [18], such that the reaction can proceed spontaneously as long as suitable catalysts are available and strictly reducing conditions are maintained [10,15,19,20]. In the presence of activated nitrogen species, hydrothermal vents can synthesize the building blocks of life [12,13]. Because of their abundance of chemical energy, and despite the absence of light, modern alkaline hydrothermal vents are teeming with microbial life [21,22], life that is ultimately fuelled by the reaction of \( \text{H}_2 \) with \( \text{CO}_2 \).

The \( \text{H}_2–\text{CO}_2 \) redox reaction is an attractive source of energy for the first chemical reactions en route to life, because it provides direct links between a known geochemical process (serpentinization) and known biochemical processes. These are most notably the reactions of core carbon and energy metabolism in acetogens and methanogens, anaerobic autotrophs that live from the reduction of \( \text{CO}_2 \) with \( \text{H}_2 \). Acetogenesis and methanogenesis represent the most primordial forms of metabolism in bacteria and archaea [23,24], rooting life’s chemistry to reactions of gasses, rocks and water.

The continuity between exergonic geochemical and biochemical reactions can be seen as a virtue of hydrothermal origin theories, because it generates concrete mechanistic links between processes catalysed by minerals in the Earth’s crust (exergonic \( \text{CO}_2 \) reduction) [25] and processes catalysed by enzymes in the metabolism of prokaryotic lineages [26]. At hydrothermal vents, life as we know it connects to geochemistry as we know it.

2. Activation of \( \text{CO}_2 \) and \( \text{H}_2 \): the door to \( \text{CO}_2 \) fixation

In biology, acetogens and methanogens fix \( \text{CO}_2 \) via the \( \text{H}_2 \)-dependent reduction of \( \text{CO}_2 \) to a methyl group and \( \text{CO} \), followed by condensation of the methyl moiety and \( \text{CO} \) to a nickel bound acetyl group that is thiolytically cleaved from nickel to generate the thioester acetyl-CoA. The acetyl-CoA pathway is unique in microbial physiology, because it is carbon and energy metabolism in one. Carbon metabolism involves the \( \text{H}_2 \)-dependent reduction of \( \text{CO}_2 \) to acetyl-CoA. Under standard physiological conditions, the synthesis of the thioester is exergonic by about \(-59\,\text{kJ mol}^{-1}\) [27], while there is not enough energy to generate thiosteres and synthesise ATP via substrate level phosphorylation [28]. Thus, for energy metabolism, acetogens that lack cytochromes and quinones couple methyl synthesis to the generation of ion gradients via electron bifurcation and ferredoxin oxidation at the membrane-bound Rnf complex [29], while methanogens that lack cytochromes generate their ion gradient by coupling the transfer of the methyl group from a nitrogen atom in methyl-tetrahydromethanopterin to a sulfur atom in coenzyme M [30].

If the acetyl-CoA pathway is the most ancient carbon fixation pathway, and various lines of evidence indicate that to be the case [14,15,23,24,27,31], there are still some dots that need to be connected. For \( \text{H}_2 \) to have played a role in early chemical evolution, it required activation—it required catalysis. It is noteworthy that \( \text{H}_2 \) never interacts directly with any organic oxidant (substrate) in metabolism, it always releases electrons into metabolism via a catalyst: hydrogenase. There are only three classes of hydrogenases known. All three harbour Fe atoms at their active site [32,33], all three harbour carbon metal bonds at their active site [26]. The central enzyme of the acetyl-CoA pathway, the only exergonic \( \text{CO}_2 \) fixation pathway known [34,35], is bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase (CODH/ACS), which also harbours carbon metal bonds. These two activities, hydrogenase and CODH/ACS, trace to the last universal ancestor, LUCA [26]. Organisms that use the acetyl-CoA pathway employ flavin-based electron bifurcation to generate ferredoxins with a lower reducing potential than \( \text{H}_2 \) [36–38]. Flavin-based electron bifurcation thus accounts for the thermodynamics of \( \text{H}_2 \) oxidation, but what about the kinetics? In kinetically controlled reactions, catalysts can have an important influence on the nature of the products that accumulate—and the same is true for geochemical \( \text{CO}_2 \) fixation with \( \text{H}_2 \).

The \( \text{H}_2 \)-dependent reaction from the most oxidized form of carbon, \( \text{CO}_2 \), to its most reduced form, methane (\( \text{CH}_4 \)), is thermodynamically favourable under reducing conditions. However, in serpentinizing, alkaline hydrothermal systems [39] the direct transfer of electrons from \( \text{H}_2 \) to \( \text{CO}_2 \) has a large activation energy and requires either high temperatures and high pressures [40] or, at milder conditions, chemical activation and catalysis [41,42]. The requirement for catalysts stems from kinetic barriers in the sequence of reactions from \( \text{CO}_2 \) to \( \text{CH}_4 \). Catalysts decrease the activation energy and thus the kinetic barrier, allowing intermediate products such as formate, acetate, methanol and pyruvate to accumulate after a short time under mild conditions [15] rather than the thermodynamically favoured end product \( \text{CH}_4 \). While high temperatures, high pressures and long reaction times lead to the accumulation of \( \text{CH}_4 \), the most stable product [40,43], catalysts influence the product distribution in the short term. In biology, enzymes effect such shifts from thermodynamically controlled reactions to kinetically controlled reactions [44]. In purely geological settings, however, heterogeneous catalysis may have been the key for early processes of protometabolism. Dashed lines indicate physisorption, non-dashed lines indicate chemisorption on the surface.
the chemical activation that enzymes provide in modern organisms [46,47].

3. Adding nitrogen

In order to synthesize amino acids and nucleic acid bases, living cells have to incorporate dinitrogen (N2) into biosynthetic pathways. From a chemical point of view, N2 as a starting material is not the easiest choice in comparison to more oxidized or reduced nitrogen compounds [48]. Nevertheless, looking at early Earth’s conditions, an atmosphere filled with N2 would have led to an ocean with dissolved N2 and thus—via sequestration through the Earth’s crust—to a nitrogen source in serpentinizing systems [49,50]. Looking at biology, N2 fixation is considered ancient [50,51]. There is only one way for N2 to enter metabolism: via the nitrogenase complex. Nitrogenase consists of two proteins, dinitrogenase reductase, which contains an FeS-based active centre and the dinitrogenase protein, harbouring an Mo (or V, or Fe) containing FeS centre with a carbide carbon at the active site [52,53]. Mechanistically, the complex works with dinitrogenase reductase harvesting the energy of ATP hydrolysis and transferring it via conformational changes to dinitrogenase, which then binds the N2 molecule [53,54]. The following steps involve sequential hydrogenations of the nitrogen molecule. There, for CO2 fixation, hydrogenase activity is needed to deliver electrons from H2 to N2. This hydrogenase activity is promoted by the FeS clusters of the nitrogenase complex [53,55], which is the sole entry point of N2 into metabolism. As CODH and hydrogenase, nitrogenase also traces back to LUCA [26,56].

Biological N2 reduction follows very different kinetics from those of the industrial process [57]. For both processes, inorganic catalysts have a central role in the reduction of N2. In industry, the reduction of N2 might resemble prebiotic FeS-based nitrogen fixation [45]. The greatest impediment to N2 reduction is its activation energy. N2 is very stable at normal atmospheric temperatures and pressures. Thus, few processes are capable of activating N2 sufficiently in order to form N-rich molecules. Industrial N2 conversion to NH3 via the Haber–Bosch process (H2-dependent) requires Fe-based catalysts such as Fe3O4, high pressure (200 bar) and temperatures exceeding 400°C [58]. The Haber–Bosch process currently consumes about 1–2% of the World’s total energy production. Biological nitrogen fixation catalysed by the nitrogenase enzyme operates at ambient pressure and room temperature. Accordingly, there is immense commercial interest in the mechanism of biological N2 fixation [57].

Not unlike the stepwise use of Fe atoms found in the active sites of the nitrogenase complex, industrial N2 reduction is extremely dependent on the physico-chemical state of the catalysts. Thus, the yield of ammonia is affected as a result of several factors such as particle size, purity and subsurface dissociation of nitrogen into Fe catalysts, leading to iron nitrides such as Fe3N [59].

Can serpentinization reduce N2? Although there is abundant evidence for abiotic CO2 reduction in serpentinizing systems [60,61], evidence for abiotic N2 reduction is so far lacking. Laboratory simulations suggest that N2 can be reduced to ammonia (NH3) with mineral catalysts under mild hydrothermal conditions [45,62]. Incorporation of N from N2 into organic compounds under hydrothermal conditions presents a more substantial challenge for laboratory simulations. In principle, activated forms of nitrogen chemisorbed to geochemical catalysts (figure 1) might be better starting points for prebiotic synthesis of such compounds than NH3 [25], but this remains to be shown experimentally.

There are nevertheless very curious parallels between industrial hydrogenation processes and geochemical H2-dependent reactions. Serpentinitization not only reduces H2O to H2 and CO2 to formate and CH4, it also generates inorganic catalysts within the Earth’s crust [25]. These include magnetite, Fe3O4, which is the catalyst of choice for the industrial Haber–Bosch process (H2-dependent N2 reduction) and for Fischer–Tropsch (CO2 reduction) applications [59,63] and awaruite, Ni3Fe, which catalyses the H2-dependent reduction of CO2 to methane at high pressures and temperatures [40]. While H2 and CO2 deliver carbon and energy, for an autocatalytic network to emerge, one from which microbial metabolism could unfold, organic cofactors, bases and amino acids are required. All are nitrogenous compounds.

4. What if C, N and H are activated together?

As shown in figure 1, it is possible that mineral surfaces can activate H2, CO2 and N2 simultaneously. If so, amino acids or even bases and cofactors might be obtained via such routes. It has been reported that Fe2+ and Fe0 can catalyse reactions of 2-oxoacids with hydroxylamine to give aspartate, alanine, glycine and glutamine [64]. These should also be the first amino acids to appear in the evolution of metabolism, if metabolism evolved from a pyruvate-fed, incomplete citric acid cycle and if amino acids arose ancestrally as they do in metabolism, namely via reductive amination of the keto group in oxalacetate, pyruvate, glyoxylate and 2-oxoglutarate [9]. Pyruvate is new as a possible prebiotic compound [14]. Using hydrothermal iron minerals instead of enzymes, it is possible to synthesise pyruvate from H2 and CO2 [15]. Pyruvate now appears to be a much more readily synthesized prebiotic compound than previously assumed.

If N2 can be activated efficiently under hydrothermal conditions, nucleic acid bases might not be far away. Recent studies show that even aromatic heterocyclic compounds such as tryptophan can be formed abiotically in serpentinizing hydrothermal systems [13]. The connection of simpler amino acids like aspartate and glycine to bases is direct, they sit in the middle of the aromatic pyrimidine (aspartate and glycine) and purine (aspartate) rings. This is shown in figure 2, modified from reference [9]. In metabolism, pyrimidines are made from aspartate and carbamoyl phosphate. Carbamoyl phosphate is made from carbamate and ATP, carbamate forms spontaneously as a colourless precipitate in hot solutions containing CO2 (or carbonate) and ammonium. Four of the atoms in the pyrimidine ring come from aspartate. Purines are more complex, but the components are simple. Glycine comprises the centre of the rings, which are completed by inclusion of C1 units from formyl tetrahydrofolate [65] or from formyl phosphate (in methanogens) [66], by N from the amido group of glutamine, and, as with pyrimidines, by CO2 and N from aspartate.

There is a clear record of geochemical origins preserved in metabolism [26]. This record can be resurrected in the
laboratory, if we find the right conditions. The four amino acids that Muchowska et al. [64] synthesize (Gly, Ala, Asp, Glu) even suggest (reveal, one might say) a connection to the evolution of the genetic code. These are the very same amino acids that are identified as ancient in different theories about the origin and evolution of the genetic code. In some theories, exactly these four (Gly, Ala, Asp, Glu) are the oldest [67]. In other theories, they are the most ancient as members of larger sets [68], while in yet other theories they rank well in order of antiquity, with Gly, Ala and Asp being the oldest, Glu coming in seventh [69]. A look at the biosynthetic families of amino acids reveals that the Asp and Glu families stand out as central.

5. Autocatalytic networks

If we assume that simultaneous activation of N₂, H₂ and CO₂ can lead to thermodynamically stable products that include...
amino acids, nucleic acid bases and cofactors (that is currently a big assumption, we admit), then small chemical networks on a laboratory scale become possible. Central to various schools of thought on chemical origins are constructs called autocatalytic networks [70]. These can represent abstract mathematical constructs or they can describe interactions in real sets of molecules. As applied to molecular interactions, autocatalytic networks contain molecules that promote the synthesis of copies of themselves [71]. According to this very general definition, autocatalytic networks can provide theoretical frameworks for both the genetics first and the metabolism first approaches to prebiotic evolution. In the former, they can be sets of nucleic acids that ligate to form specific products [72], in the latter, they can be sets of metabolites that interact in such a way as to generate self-sustaining metabolic networks [24].

When describing molecular interactions, autocatalytic sets require input molecules in order to promote the synthesis of their constituent elements. This condition draws attention to a particular class of autocatalytic networks called reflexively autocatalytic food-generated networks—RAFs [73]—in which each reaction is catalysed by a molecule from within the network, and all molecules can be produced from a set of food molecules by the network itself. RAFs are particularly interesting in the context of early evolution, because they do not require a pre-existing catalyst for a reaction before it is required. The reaction can proceed uncatalysed, or rather catalysed by an unknown molecule, as long as the known catalyst is produced at some point by the network and assumes the role of catalysis in that reaction of the RAF. Moreover, when it comes to the concrete modelling of early evolution, the nature and source of the food molecules [74] that generate a given RAF or other autocatalytic set are of particular interest, because in order for the reactions in the set to take place, the overall thermodynamics of the network must be exergonic. In other words, in order for RAFs (or other autocatalytic networks) to serve as a useful model for early evolution, the set of reactants (educts) needs to release energy en route to the products (adducts), as is always the case in metabolism [18].

Of course, in cellular metabolism, the overall energetics are given by the sum of the changes in free energy for the core bioenergetic reactions [18]. For individual reactions of metabolism, the change in free energy from substrate to product is often endergonic, which is why such reactions are usually coupled to energy-releasing reactions involving exergonic electron transfer, ion gradients across the plasma membrane, or hydrolysis of high-energy bonds, such as ATP, acyl phosphates or thioesters [18,37]. Energetic coupling can also occur within RAFs, which makes them more interesting models of cellular metabolism.

Figure 3. Purely geochemical reactions such as CO₂ fixation with H₂ can give rise to autocatalytic networks and protometabolism, as long as energy is released. Kinetically controlled reactions build up a specific set of products which interact further to form an autocatalytic network that serves as a basis for higher complexity. C₁, C₂, C₃ represent carbon compounds with 1, 2 or 3 carbon atoms such as formyl groups, acetyl groups, and pyruvate. (Online version in colour.)
It seems likely that at least a subset of the catalysts, high-energy bonds and energetic currencies that occur in modern metabolism were generally present and functional in prebiotic chemistry. Sources and transduction of modern metabolic catalysis and energy should then have analogues or homologues in geochemical settings. Regarding catalysis, there are now good indications that metals and simple organic cofactors could have promoted the emergence of cell-sized autocatalytic networks [15,64,75,76]. In physiology, the term energy metabolism generally means ATP synthesis. There are two sources of ATP in cells: chemiosmotic coupling and substrate level phosphorylation (SLP). Chemiosmotic coupling needs ion gradients as an energetic intermediate and proteins, without exception. SLP does not require ion gradients, its energy source is the Gibbs free energy of chemical reactions, and SLP reactions can take place without enzymes [77–79]. Although vents harbour natural ion gradients, ATP synthesis via chemiosmotic coupling always involves the ATPase, for which there is no known geochemical homologue or mechanistic analogue. The energy for SLP stems from the redox chemistry of carbon whereby both carbon oxidation to CO$_2$ and H$_2$-dependent CO$_2$ reduction can be coupled to SLP [80]. Because the H$_2$-dependent CO$_2$ reducing reaction that drives SLP in acetogens (acetate synthesis) operates in the laboratory under simulated hydrothermal vent conditions with only metals and metal ions as catalysts [15], it is currently the only candidate for a primordial (geochemical) source of energy conservation (acyl phosphates via SLP) that is mechanistically linked to naturally occurring carbon redox reactions at vents.

A set of molecules that is generated by kinetically controlled reactions (the most rapidly formed products accumulate) will contain chemical energy that permits members of the set to interact further and to form an autocatalytic network that can serve as a basis for higher complexity [76]. Such a process is sketched in figure 3. The energetic input is necessarily centralized because thermodynamically stable metabolites and end products are synthesized from the core exergonic reaction, in our example the reduction of CO$_2$ with H$_2$ via the acetyl-CoA pathway [9,15,31].

6. Conclusion

Hydrothermal vents contain catalysts and chemical disequilibria that resemble life and metabolism in many ways. However, the natural chemical environment at vents does not strongly resemble metabolism in many forms of life, because metabolism is extremely diverse. Rather, it very specifically resembles the physiology of acetogens and methanogens, even down to the catalysts involved. The connections between the origin of microbial life and the chemical elements seem more tangible than ever before. Current genomic analyses indicate that the last universal common ancestor of all life, LUCA, lived from gasses: H$_2$, CO$_2$ and N$_2$ [23,56]. Although our main focus is on these three gasses, it is evident that the incorporation of sulfur (S) and phosphorus (P) into early metabolism was also essential. Sulfur enters metabolism as HS$^-_{\text{a}}$ at cysteine synthesis from O-acetyl serine or O-phospho serine [81], while phosphorus enters metabolism via thioesters as acyl phosphates [9]. Under reducing conditions, H$_2$S (HS$^-_{\text{a}}$ in alkaline vents) would be the likely sulfur source, phosphorus could enter the geochemical setting as phosphate dissolved in seawater or leached from the primordial crust, but data on phosphate under early Earth conditions is scarce [82–84]. Focusing on the enzymes that channel H$_2$, CO$_2$ and N$_2$ into metabolism might uncover clues about the environment within which life arose and about the catalysts that activated these gasses at origins. The presence of carbon metal bonds in the active sites of hydrogenases, nitrogenase and carbon monoxide dehydrogenase suggest that these might be ancient relicts of the catalytic realm that led to the autocatalytic synthesis of the first organic compounds. We propose that the biology of methanogens and acetogens, anaerobic autotrophs that inhabit vents today, holds clues about the primordial catalysts that enzymes ultimately came to replace.

Data accessibility. This article has no additional data.

Authors’ contributions. M.P. edited and revised the manuscript and contributed to the paragraphs on catalysis and activation of N$_2$, CO$_2$ and H$_2$; J.C.X. edited and contributed to the paragraphs on autocatalysis and, together with W.F.M., drew figure 3; A.N.V. edited and contributed to paragraphs on nitrogen activation and fixation. K.K. and J.F.A. edited and advised on the content throughout the manuscript; K.K., W.F.M. and M.P. discussed and drew figure 1. W.F.M. wrote the first draft of the manuscript and sketched figure 2.

Competing interests. We declare we have no competing interests.

Funding. This work was supported by the Deutsche Forschungsgemeinschaft (grant no. MA-1426/21-1); the European Research Council (grant no. 666053); and the Volkswagen Foundation (grant no. 93046).

Acknowledgements. All authors thank the organizers of the 30/80 Meeting for the opportunity to submit to this special issue and Mike Russell for his inspirational work on the role of hydrothermal vents in the origin of life.

References


18. Martin WF. 2012 Hydrogenases from the ancient transition from geoenergetics to endothermic reactions. IUBMB Life 64, 1038–1049. (doi:10.1002/iub.1010)


