The abiotic pattern of biotic CO_2 fixation in early metabolism

Inaugural dissertation

For the attainment of the title of Doctor of Philosophy (Ph.D.)

in the Faculty of Mathematics and Natural Sciences

at the Heinrich Heine University Düsseldorf

presented by

Martina Preiner, M.Sc.

from Burghausen, Germany

2019

From the Institute of Molecular Evolution

at the Heinrich Heine University

Published by permission of the

Faculty of Mathematics and Natural Sciences at

Heinrich Heine University Düsseldorf

Supervisor: Prof. Dr. William F. Martin

Co-supervisor: Prof. Dr. Michael Schmitt

Date of the oral examination: January 28th 2020

Statement of declaration

I hereby declare that this dissertation is the result of my own work. No other person's work has been used without due acknowledgement. This dissertation has not been submitted in the same or similar form to other institutions. I have not previously failed a doctoral examination procedure.

Düsseldorf, December 12th 2019 Martina Preiner

To Dr. Clara Immerwahr

Representing the ignored and inhibited half of scientists in the 20th century and beyond

And to everyone who struggles with injustice and inhumanity as she did

Ja, mach nur einen Plan!
Sei nur ein großes Licht!
Und mach dann noch 'nen zweiten Plan
Gehn tun sie beide nicht.
Denn für dieses Leben
Ist der Mensch nicht schlecht genug.
Doch sein höhres Streben
Ist ein schöner Zug.

Aus: „Das Lied von der Unzulänglichkeit menschlichen Strebens",
Die Dreigroschenoper (1928), Bertholt Brecht.

Ay, make yourself a plan
They need you at the top
Then make yourself a second plan
Then let the whole thing drop
For this bleak existence
Man is never bad enough
Though his sheer persistence
Can be lovely stuff

From: 'Song of the Insufficiency of Human Struggling', The Three Penny Opera (1928), Bertholt Brecht (Translation: Elizabeth Hauptmann).

Acknowledgements

When I decided to leave Science in 2009, it was one of the best decisions I ever made. Being a Science journalist opened my eyes for so many things I had forgotten over the years of studying, amongst them the puzzle that is the origin of life. When I interviewed a certain William F. Martin for an article on the origin of life, I found that 'thing' again, that had made me study Science to begin with and also had made me leave it (in order to actually explore it more). In the interview, Bill told me he was starting a chemistry section in his biology lab and as a joke I offered to help to work there. Well, three years later the joke is on me, I guess. Bill offered me to set up the prebiotic experimental part of the Institute for Molecular Evolution, although he could not be certain if a journalist who was out of the lab for 7 years could do it. For this leap of faith (or better: Vertrauensvorschuss) and the mentoring in the past three years I am very grateful. Together with the always calm and clear advice from Carlo Kleinermanns, I believe we have started a project here that will produce insights into the origin of life research for years to come. Thanks to both of you, and Andrey, who is already doing a great job holding the torch up while I am trying to write this thesis.

This work would not have been possible without our multiple collaboration partners and labs that were willing to host me for shorter and longer times. With this, I want to thank Joseph Moran, Harun Tüysüz, Yoichi Kamagata and Chris Greenwell for letting me stay in Strasbourg, Mülheim, Tsukuba, Sapporo, and Durham, respectively, while sharing your knowledge on the various topics we tried to include in our studies. And thank you, Kamila, Sreejith, Mingquan, Kensuke, Nobu, Alex, Marius, and Valentina for being such knowledgeable and helpful colleagues and friends whenever I bothered you in your labs.

Without Verena, our lab in Düsseldorf would probably implode and leave a black hole behind. Thanks for all the stupid questions you answered, the help and guidance you gave and last but not least the most delicious Quittenlikör I've ever had. Also, I want to thank Michi for helping my incapable self with setting up work computers, installing the right software and being one of the people to stay late on a party as well. Every lab needs someone like Nils. Someone to care about social activities, about actually organizing them and most of all, someone who cares about people, which he does. Thank you for that. Also my thanks to Renate, Doris, and Margarete who always helped with the bureaucratic part of everything (mostly their work was to remind me to do it, for which I am deeply sorry).

When you set up a lab, a lot has to be built from scratch. Waldemar Seidel and his team in the biology workshop are incredibly kind people who always crafted the crazy ideas we had. Without them, many experiments would not have happened. Thank you.

I also want to thank all the students that came through the lab, taking over parts of our work, enduring me, mostly not blowing anything up: Victoria, Oksana, Perli, Janice, Slava, and Delfina.

Almost there!

Michael Schmitt and Peter Jahns supported us very much in the first year of this project, by letting us use their analytical means, thank you. I am grateful that Michael also agreed to be Co-supervisor for this thesis. I also want to thank John Allen and Sean Jordan for proofreading this thesis. I am honoured that you two took the time.

Same goes for Sriram and Joana, who also proofread, but above all, I want to thank you for making our office an oasis of stupid jokes, deep conversations and sometimes also darkness (our old friend), when needed. I cannot imagine a life in which we didn't meet and became friends.

Bill was generous enough to let me stay at least a part-time Science journalist, doing a Podcast with a very talented Science journalist, Franzi, who also happens to be one of my closest friends. Thank you, Bill, for letting me do this, and thank you, Franzi, for helping me to stay connected to my journalistic roots. It kept me sane (or at least sane-adjacent). And if that didn't help, it was friends, close and afar, like Martin, Kerrin, Vicky, Vroni, John, Absi, Alice (x2), Valesca, Nils, Philipp, Flösi, and Daniel who reminded me that there is more to life than being stressed about deadlines, rude colleagues and failure.

Last but most importantly, I want to thank my parents Ana Maria and Gerhard, who have been nothing but generous to me throughout my entire life. And my chosen family, my partner Andrej, who I promised not to mention here so I will just mention your name and won't elaborate on how much your support, our friendship and partnership means to me.

Publications in this thesis

1) Martin, W. F. and Preiner, M. (2017) 'Origin of life, Theories of', *Reference Module in Life Sciences*, Elsevier Inc., Oxford.
doi: 10.1016/B978-0-12-809633-8.02403-1.

2) Preiner, M., Asche, S., *et al.* (2020) 'The future of origin of life research: bridging decades-old divisions', *Life*, 10(3), 20.
doi:10.3390/life10030020.

3) Sousa, F. L., Preiner, M. and Martin, W. F. (2018) 'Native metals, electron bifurcation, and CO_2 reduction in early biochemical evolution', *Current Opinion in Microbiology*, 43, pp. 77–83.
doi: 10.1016/j.mib.2017.12.010.

4) Weiss, M. C., Preiner, M., *et al.* (2018) 'The last universal common ancestor between ancient Earth chemistry and the onset of genetics', *PloS Genetics*, 14(8), e1007518.
doi: 10.1371/journal.pgen.1007518.

5) Xavier, J. C., Preiner, M. and Martin, W. F. (2018) 'Something special about CO-dependent CO_2 fixation', *FEBS Letters*, 285(22), pp. 4181–4195.
doi: 10.1111/febs.14664.

6) Preiner, M., Xavier, J. C., *et al.* (2018) 'Serpentinization: connecting geochemistry, ancient metabolism and industrial hydrogenation', *Life*, 8(4), 41.
doi: 10.3390/life8040041.

7) Preiner, M., Igarashi, K., *et al.* (2020) 'A hydrogen dependent geochemical analogue of primordial carbon and energy metabolism', *Nature Ecology and Evolution*.
doi.org/10.1038/s41559-020-1125-6.

8) Preiner, M., Xavier, J. C., *et al.* (2019) 'Catalysts, autocatalysis and the origin of metabolism', *Interface Focus*, 9, 20190072.
doi: 10.1098/rsfs.2019.0072.

Table of contents

Abstract

Life as we know it here on Earth manifests itself in countless forms, but all of them share three fundamental properties: i) They harness energy from their environment to fuel chemical reactions. ii) These chemical reactions primarily involve carbon-based molecules. iii) Catalysts accelerate the chemical reactions, thus even making them relevant for life.

But how and where did these catalysed reactions begin to run their course? The central molecules of life must have had a tendency to form spontaneously on early Earth and organize towards higher complexity, a process that was driven by the environmental conditions our planet offered when life arose 4 billion years ago. At and before the beginning of life, the fundamental properties listed above probably played a pivotal role, thus the conditions on early Earth had to include an energy source, a carbon source and access to catalysts.

All three can be found in certain geochemical ('serpentinizing') systems: chemical energy in the form of hydrogen (H_2), a carbon source in the form of carbon dioxide (CO_2), and transition metals bearing minerals that can act as catalysts for the reaction between H_2 and CO_2. Some anaerobic prokaryotes also use these two gasses—employing transition metals bearing enzymes—to sustain their metabolic need for complex organic molecules.

In this thesis, the parallels between abiotic geochemical CO_2 fixation and biotic processes observed in cells are examined. The aim is to find a transition point (or points) from geo- to biochemistry at the origin of life.

Zusammenfassung

Das Leben, so wie wir es hier auf der Erde kennen, nimmt unzählige Formen an, aber alle teilen sich folgende Eigenschaften: i) Sie nutzen die Energie ihrer Umwelt um chemische Reaktionen anzutreiben. ii) An diesen chemischen Reaktionen sind maßgeblich Moleküle auf Kohlenstoffbasis beteiligt. iii) Katalysatoren beschleunigen die chemischen Reaktionen und machen diese dadurch erst relevant für das Leben.

Doch wie und wo haben diese katalysierten Reaktionen ihren Lauf genommen? Die zentralen Moleküle des Lebens müssen sich auf der frühen Erde spontan geformt und zu komplexeren Systemen organisiert haben. Dieser Prozess wurde von den Umweltbedingungen angetrieben, die unser Planet vor 4 Milliarden Jahren zu bieten hatte, also dem Zeitpunkt, an dem das Leben nachweislich entstand. Dabei werden die genannten zentralen Eigenschaften des Lebens wahrscheinlich schon eine entscheidende Rolle zu oder sogar vor Beginn des Lebens gespielt haben. Die Konditionen auf der jungen Erden mussten dementsprechend eine Energiequelle, eine Kohlenstoffquelle und Zugang zu Katalysatoren umfassen.

All das lässt sich in bestimmten geochemischen („serpentinisierenden“) Systemen finden: chemische Energie in Form von Wasserstoff (H_2), eine Kohlenstoffquelle in Form von Kohlenstoffdioxid (CO_2) und Übergangsmetall-haltige Mineralien, die als Katalysatoren für die Reaktion zwischen H_2 und CO_2 fungieren können. Auch einige anaerobe Prokaryoten nutzen diese beiden Gase um mithilfe von Übergangsmetall-haltigen Enzymen komplexe organische Moleküle in ihren Stoffwechsel zu bringen.

In dieser Arbeit werden die Parallelen zwischen der abiotischen geochemischen CO_2-Fixierung und den entsprechenden biotischen Prozessen, die man in lebenden Zellen findet, untersucht. Das Ziel ist es, einen Übergangspunkt (oder Übergangspunkte) von Geo- zu Biochemie am Ursprung des Lebens zu finden.

Outline and aim of this thesis

This thesis approaches the question of how geochemical reactions on early Earth eventually led to life as we know it. Many follow-up questions that are important pillars of origin of life (OoL) research start to evolve from this starting point, among them: what geochemical reactions were possible on early Earth? How did complex metabolic systems arise from simple reactions? Are these the first reactions that ultimately led to life still imprinted in life today? Embracing the premise that the answer to this last question is 'yes', this thesis focusses on modern biotic carbon fixation found in anaerobic autotrophs, an important hint towards how metabolism could have developed over 4 billion years ago.

To place the role of carbon fixation and metabolism a position within the framework of the many topics that OoL research entails, an historical and future-oriented perspective is given in Chapter I and Publications 1 and 2 contained therein.

Furthermore, CO_2 fixation will be explored from the standpoints of both biology and prebiotic chemistry. The biological side constitutes a crucial approach, because it is widely ignored within chemical approaches in OoL research. It is important to understand what life does today, in order to build up a concept of how it began. This is elucidated in in Chapter II (Publications 3–5).

The aim here is to find parallels between biological and geochemical H_2-dependent CO_2 fixation, concentrating on the acetyl-CoA pathway on the biological side and serpentinizing systems on the geochemical one. Details of this approach will be found in Publication 6 (Chapter III). The focus lies on the available catalysts, assuming that the 'right' choice would help us understand why anaerobic metabolism is what it is. A fitting setup and an analytical method to analyse the products had to be established in order to conduct simple one-pot experiments based on the conditions found in serpentinizing systems. The results of these experiments are presented in Publication 7, and outlook towards future experiments is given in Publication 8.

By providing insight into the many ways in which nature fixes carbon, it will be shown that both abiotic and biotic pathways share many similarities.

Introduction

1. Abiogenesis: a tale of energy and carbon

1.1 In the beginning

The question 'Where do we come from?' has been fascinating and influential throughout the history of mankind. Modern scientific hypotheses on the origin of life (OoL) on Earth emerged with a deeper understanding of evolution and, to be more precise, of the origin of species (Darwin, 1859). Charles Darwin concludes Chapter IV of his most pivotal publication, *On the origin of species by means of natural selection*, with the following words:

> *"As buds give rise by growth to fresh buds, and these, if vigorous, branch out and overtop on all sides many a feebler branch, so by generation I believe it has been with the great Tree of Life, which fills with its dead and broken branches the crust of the earth, and covers the surface with its ever branching and beautiful ramifications."* (Darwin 1859, p. 119)

In that chapter resides the only illustration of Darwin's treatise, a precise diagram of a rooted phylogenetic tree. Looking at this abstract, schematic diagram, and the almost poetic last sentence of the chapter, one cannot help but wonder what lies beneath the root of this "great Tree of Life".

Darwin asked himself this question later on and, in a correspondence with Joseph D. Hooker (Darwin, 1871), came up with warm little ponds "with all sorts of ammonia & phosphoric salts" which, with the help of external energy, could form "a protein compound (…) ready to undergo still more complex changes". This simple description by Darwin was actually already quite progressive for the time. He described how, in his opinion, life-like processes could have emerged abiotically, so out of chemicals that are incorporated within life today, although he neglected to mention the element that provides the backbone of life: carbon. Many different hypotheses on where, when and how the origin of life occurred have been proposed, developed, debated, and, if possible, tested, since Darwin's letter to his friend was written.

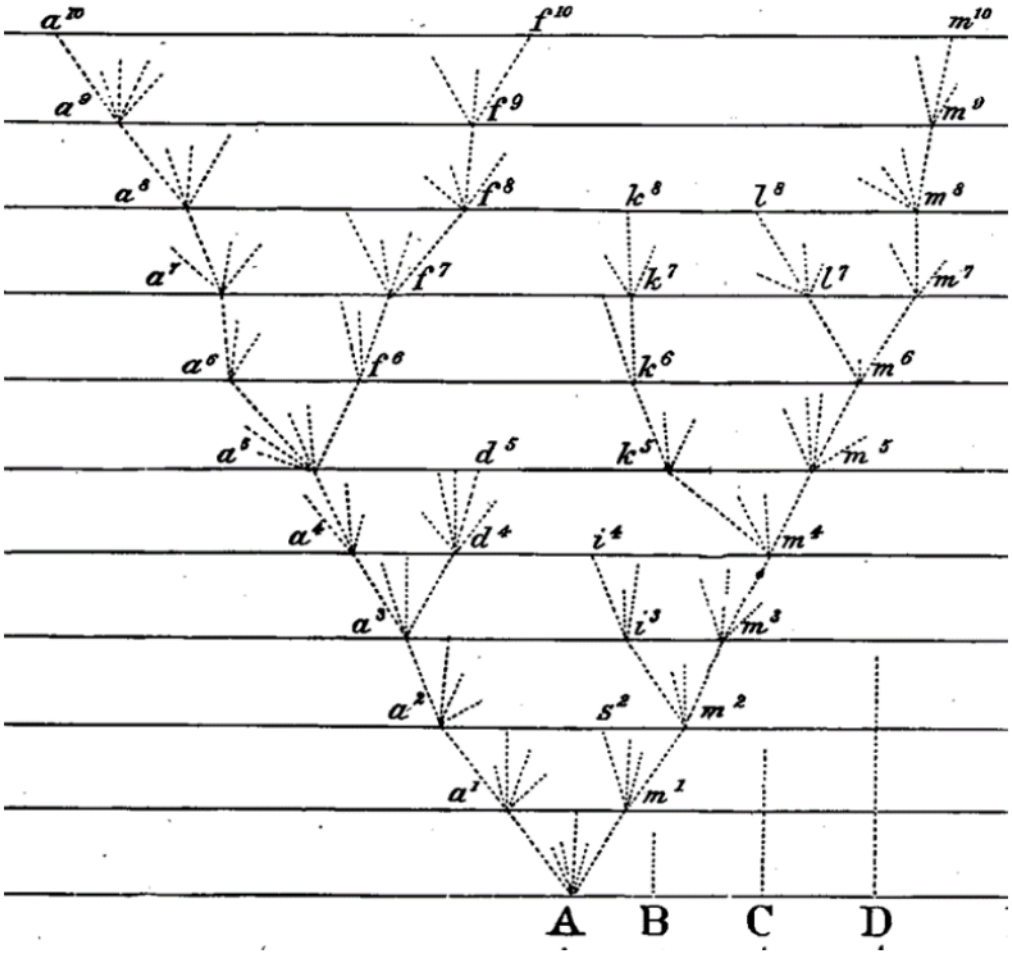


Figure 1: One part of the fold-out diagram in Darwin's *On the origin of species by means of natural selection*. Darwin included the only illustration in his book to facilitate the understanding of "the perplexing subject" of how the "principle of great benefit being derived from divergence of character, combined with the principles of natural selection and extinction, will tend to act." (Darwin, 1859, p. 108)

Since the 1860s, popular OoL theories amongst biologists were those directly connected with the Darwinian view; that of a slow transition from non-living to living matter, a form of 'chemical evolution' (Moore, 1913). This was in contrast with a spontaneous generation of life which was the predominant theory propagated throughout the 18th to 20th century (Farley, 1978; Kamminga, 1988). It was—and for some still is—the conviction that, with enough time, simple forms of life could develop from carbon-based compounds gradually increasing in complexity, which meant that there were no impassable barriers in nature (Kamminga, 1988).

But there was a parallel, rather restrictive ideological development in the 19th century as well: the protoplasmic theory of life (Huxley, 1869; Welch, 1995). Restrictive, because

protoplasm was defined by Thomas H. Huxley as "*the* physical basis and matter of life" (Huxley, 1869, p. 3). In other words, protoplasm or the 'cell substance' was regarded as a single unit and the OoL meant the origin of protoplasm. However, with time and the advent of modern analytical methods came the insight that the cell substance was chemically far too complex to be considered one single entity. It is likely that life could not have started with such a demanding heterogeneous mixture of individual enzymes, proteins and other organic molecules that already was perfectly functioning.

Nevertheless, this 'one entity to solve it all' approach is something that the OoL field circled back to towards the end of the 20th century, when the RNA world hypothesis was postulated. This will be briefly discussed further below (Chapter 1.3). In the beginning of the 20th century, however, it became clear that what had to be solved was how one of these complex systems could have emerged. A lot of biochemists chose to steer clear of opening this Pandora's box, perhaps because the problem could not be solved by one scientific discipline alone (Kamminga, 1988).

1.2 "Boring from the two opposite ends of a tunnel"

For the first few decades of the 20th century, OoL research stagnated as scientists simply avoided this seemingly unsolvable puzzle. But then, the first attempt for an interdisciplinary scientific overview specifically on the OoL was put to paper, written by the Soviet biologist and biochemist Alexander Ivanovich Oparin. His book, *The Origin of Life*, was published in 1936, with the English translation following in 1938 (Oparin, 1938). These and the following extended editions of the book revived the OoL field by presenting new ideas and approaches in organic and inorganic chemistry, comparative biology, biochemistry, geochemistry, and astrophysics. Oparin looked deeply into the biochemical properties of life and, furthermore, stated important considerations about the thermodynamics and kinetics necessarily involved in the transition between abiotic and biotic processes (Oparin, 1957).

Oparin based his assumptions about the conditions of early Earth on geochemical studies of the modern Earth's crust, spectroscopic investigations of stars and planets, and the chemical composition of meteorites. An important insight of these comparative studies was that the Earth's atmosphere was oxygen-depleted and of a reducing nature. This had been independently proposed by J.B.S. Haldane a few years earlier in a short publication (Haldane, 1929). An anaerobic, reducing atmosphere opened the possibility of building up carbon compounds from more oxidized forms like carbon monoxide (CO) or carbon dioxide (CO_2), instead of having the equilibrium position of these reactions lie on the side of CO_2 (as it would in an aerobic

atmosphere). Knowledge of organic chemistry at that time was already advanced enough to predict possible synthetic pathways, from formaldehyde to formate to amino acids and even organic polymers.

Hypothetical access to more complex organic molecules allowed Oparin to consider the role of coacervates after delving into the state of the art literature on colloids, that is, suspensions of two substances that do not share the same phase (Ostwald and Köhler, 1927; Bungenberg de Jong, 1932). In a prebiotic context, Oparin considered coacervates as microdroplets formed from hydrophobic compounds (organic or inorganic) in aqueous solution, like small compartments, which, by most definitions, are a crucial prerequisite for life (Ruiz-Mirazo, Briones and De La Escosura, 2014; Martin and Preiner, 2017). The ability of such droplets to absorb and incorporate organic matter from the solution around them and utilising it to grow and develop reminded Oparin of metabolic structures, for him they constituted a possible link between prebiotic and simple biotic cellular systems (Kamminga, 1988). To this day, the formation of organic (biologically inspired or not) microdroplets under various conditions constitutes a large part of OoL research (Deamer, 1997; Jia *et al.*, 2019; Jordan *et al.*, 2019).

Oparin viewed metabolism as a chain of coupled oxidations and reductions including intermolecular hydrogen transfer based on the biochemistry of his time (Kamminga, 1988). The idea of metabolism being one unit and, moreover, the elementary property of life, was central to his theory. He compared the existing data of the different metabolic systems from existent organisms and mapped a possible order for the evolution of the earliest organisms. With this, he introduced comparative biochemistry, central tool in modern OoL and molecular evolution research (Kamminga, 1988). In later editions of his oeuvre, he also elaborated on the role of CO_2 fixation with H_2 as an important part of building up the necessary carbon compounds, comparing the processes of what he knew from autotrophic carbon fixation, calling it the "abiogenic (…) evolution of carbon compounds" (Oparin, 1957, p. 139).

The overall unprecedented interdisciplinary approach of Oparin resulted in what is known today as the 'Heterotrophic origin of life theory' or the 'Oparin–Haldane hypothesis', as it is credited to Oparin and Haldane. The latter was also was the first to mention the word 'soup' in context of the OoL (Haldane, 1929), which is why the expression 'primordial soup' is common parlance today.

Oparin stated his opinion on how to approach research on the origin of life in a booklet that preceded the first edition of his book by 12 years. The booklet did not include the chemical and biochemical detail which made his later works so valuable, nor was it translated from

Russian into English before 1967. Thus, this early work was not recognized widely, although Oparin's idealistic roadmap towards a solution for the OoL puzzle is still relevant today:

> *"A whole army of biologists is studying the structure and organization of living matter, while a no less number of physicists and chemists are daily revealing to us new properties of dead things. Like two parties of workers boring from the two opposite ends of a tunnel, they are working towards the same goal. The work has already gone a long way and very, very soon the last barriers between the living and the dead will crumble under the attack of patient work and powerful scientific thought."* (Oparin, 1967, p. 234)

Oparin was perhaps over-optimistic about the upcoming scientific breakthroughs in OoL research, but his message was clear: all disciplines have to pitch in in order to see the light at the end of the figurative tunnel. His theory reflected this conviction. Formulated openly and including the view of many disciplines, it was predestined to be flexible enough for scientific progress (Farley, 1978). Oparin himself extended and changed his theory until his death in 1980 (Oparin, 1957; Kamminga, 1988). Even if researchers did not agree with parts of Oparin's theory, they were able to build on other aspects of it. Most important of all, parts of Oparin's hypotheses could actually be tested, something that resulted in probably the best known OoL set-up—the Miller-Urey experiment.

1.3 Modern times, divided times

When Stanley Miller, under the supervision of Harold Urey, started setting up an experiment to test Oparin and Haldane's hypothesis of the primordial soup during his PhD thesis, his were not the first chemical experiments central to the field of prebiotic chemistry (Miller, 1953). Reactions like the Strecker-synthesis of amino acids from aldehydes, ammonia, and hydrogen cyanide (Strecker, 1850) or the formose reaction, which is still widely considered the most important prebiotic pathway from aldehydes to sugars (Butlerow, 1861) had already been described a century earlier. Nevertheless, Miller's experiment became a landmark because it was tailored for OoL, it was based on an existing hypothesis, and was performed under conditions that were considered to represent those on early Earth (Bada and Lazcano, 2003). Miller circulated 200 mL water (H_2O) and ca. 20 mbar (Miller used the unit '20 cm H_2O') of methane (CH_4), 20 mbar of ammonia (NH_3), and 10 mbar of hydrogen (H_2) through a glass U-

tube, using electrical discharge to create free radicals. His observations over the one-week long run are described in the *Science* publication of 1954:

> *"During the run the water in the flask became noticeably pink after the first day, and by the end of the week the solution was deep red and turbid. Most of the turbidity was due to colloidal silica from the glass. The red color is due to organic compounds adsorbed on the silica. Also present are yellow organic compounds, of which only a small fraction can be extracted with ether…"*(Miller, 1953, p. 528)

Miller then tested the resulting mixture for amino acids by paper chromatography and was able to identify glycine, α-alanine and β-alanine. Two years later, with improved analytic means he could also detect the amino acids sarcosine, d,l-α-aminobutyric acid, and α-methylalanine. The major part of the acid fraction consisted of carboxylic acids like glycolic, lactic, formic, acetic and propionic acid (Miller, 1955). In the gas phase, apart from the educt gasses, Miller detected nitrogen (N_2), carbon monoxide (CO), and carbon dioxide (CO_2), the latter two probably also reactants *en route* to the amino- and carboxylic acids he found.

These first *de facto* OoL experiments kicked off further experimental investigations of prebiotic chemical reactions. This research provided constraints on the otherwise endless possibilities of organic compounds available in various versions of early Earth environments. Over the years many different OoL locations have been considered in terms of abiotic organic synthesis (McCollom, 2013b). For example, at the time of the Miller–Urey experiments the most favourable environment for the OoL was seen as Oparin and Haldane's primordial soup. In the mid-1970s the discovery of deep-sea hydrothermal vents (Corliss *et al.*, 1979; Baross and Hoffman, 1985; Kelley *et al.*, 2001) led to hypotheses based on these locations and their conditions. The same is true for more recent proposals that built theories around terrestrial hot springs (Milshteyn *et al.*, 2018). These are merely a few of many proposed OoL environments.

A variety of hypotheses followed not only from the physico-chemical constraints of different possible sites for the emergence of life, but they also led to the abandonment of Oparin's concept of looking at metabolism as a unity. When the catalytic properties of RNA were discovered in the early 1980s in the form of ribozymes (Kruger *et al.*, 1982), the concept of the 'RNA world' by Walter Gilbert quickly established a hierarchy of biomolecules (Gilbert, 1986). For many OoL researchers, the synthesis of RNA under 'plausible' prebiotic conditions (Powner, Gerland and Sutherland, 2009) equalled the solution to OoL itself, since those

molecules unified both informational (genetic) and enzymatic (metabolic) properties (Orgel, 2004), both main attributes of life as we know it. Once there was RNA, life could become established, and the role of other molecules and pathways within the metabolic system became very insignificant, not unlike the (by then 100 years old) protoplasm idea, where one substance was also thought to be the solution to all. Unlike the complex, mainly unknown composition of protoplasm, the RNA molecule's structure and chemical behaviour is well investigated and currently seems to be something within reach in terms of laboratory synthesis. But fundamental difficulties with the concept were formulated over the decades that followed. These include the lack of templates for RNA polymerization (Shapiro, 2000) and the improbability of such a complex molecule as RNA forming prebiotically without an established metabolic system given that it is susceptible to destruction via hydrolysis (Oivanen, Kuusela and Lönnberg, 1998; Preiner, Asche, *et al.*, 2019). Also, geochemists point out that most of the 'plausible' conditions that are deployed by typical RNA world experiments are not compatible with the reconstructed environments of early Earth (McCollom, 2013b), something which is addressed in more detail in this thesis' Chapter I (Publication 1).

Often depicted in direct opposition to RNA world are 'metabolism first' theories (Orgel, 2008), which start from simpler molecules building up pathways and networks that can harness energy more or less directly from geochemical sources before reaching genetic complexity (Wächtershäuser, 1997; Dyson, 1999; de Duve, 2003; Smith and Morowitz, 2004). The concept 'from simple to complex' is reflected in all theories for OoL, including the alkaline hydrothermal vent approach described in Chapter 1.4 and, by extension, the publications presented in Chapter II and III of this thesis.

Many divisions within the OoL research community boil down to how carbon enters the system. If carbon enters inorganically (mainly CO_2, sometimes CO), life is considered to have had an autotrophic origin. In heterotrophic origin theories, the source of carbon is a reduced, organic form, (Schönheit, Buckel and Martin, 2016). Heterotrophic origin theories started with Oparin's comparative biochemistry studies and are traditionally favoured by prebiotic chemists (Oparin, 1957; Orgel, 2008), picturing the first cells living from complex organic molecules synthesized abiotically, outside of the cell. Autotrophic theories are normally supported by a biological perspective, where the first cells are usually considered to be autotrophs (Weiss *et al.*, 2016), building up organic carbon molecules from the already mentioned inorganic C1 sources. Evidence for these theories is also discussed in Chapter I (Publication 1) of this thesis. Some have considered the division of autotrophic and heterotrophic OoL theories to be misleading, as ecological trophic levels do not necessarily apply to biosynthetic pathways

(Smith and Morowitz in "Workshop OQOL'09: Open Questions on the Origins of Life 2009"). A chemical distinction between autotrophic and heterotrophic origin models might be whether the mechanisms of a metabolic pathway were conserved or replaced, and "whether the original molecular inventory was similar to the universal core today" (Smith and Morowitz in "Workshop OQOL'09: Open Questions on the Origins of Life 2009", p. 398).

This said, life can be considered as a process harvesting the energy from constantly occurring reactions—and the complex molecules and networks of life as side products of that main exergonic reaction, to be more precise, a redox reaction, an electron transfer. Or, as the biologist Bill Martin puts it: "Life is a chemical reaction" (Martin, 2011, p. 1).

1.4 One reaction to fuel them all?

There are a many facets to the origin of life, but throughout its research history, from Darwin to Oparin and Miller, RNA world to metabolism first, primordial soups and lightning to hydrothermal vents, the fixation (and thus hydrogenation) of carbon is central to them all. Some would even say, that "the hydrogenation of CO_2 [is] life's job", an expression that was coined by geochemist Mike Russell (Russell, 2019, p. 6). The hypothesis behind this expression is that life was a way to harness the free energy stored in CO_2 in Earth's atmosphere and ocean over 4 billion years ago. The chemical potential of CO_2 can be accessed via hydrothermally produced H_2 (Martin and Russell, 2007). In the reaction of H_2 with CO_2, the equilibrium lies on the side of reduced carbon compounds (Shock, 1990), so one could expect them to react spontaneously to produce methane, water and heat (Shock, 1990). But due to kinetic barriers (Maden, 2000), they do not do that, but life makes them react, or to be more precise, life forms like methanogens make them react (Sousa and Martin, 2014; Weiss *et al.*, 2016). With enzymes, cofactors, and a whole plethora of coupled reactions, life draws energy exclusively from the transfer of electrons from donor to acceptor molecules (Morowitz and Smith, 2006). At life's origin, the kinetic barriers to the reaction of H_2 with CO_2 were overcome by abiotic catalysts. Depending on the catalyst composition and on the environmental conditions (temperature, pH, pressure), a complex series of reactions emerged, ultimately leading to the chemical reaction that is life (Martin, 2011).

But how did this series of reactions start exactly? That is where the approach already applied by Oparin—uniting comparative biochemistry and geochemistry—comes into play. When deep-sea hydrothermal vents where discovered in the 1970s, they were quickly established as possible sites for the origin of life (Corliss *et al.*, 1979; Corliss, Baross and Hoffman, 1981; Baross and Hoffman, 1985). The vent hypotheses were adapted with time and

expanded by insights from biology (Decker, Jungermann and Thauer, 1970; Thauer, Jungermann and Decker, 1977) and chemistry (Wächtershäuser 1988, 1990, 1992), leading to a theory built around alkaline, low-temperature hydrothermal vents, rich in iron sulfides (Russell, Hall and Turner, 1989; Russell *et al.*, 1993; Russell and Hall, 1997).

The evolution of the first metabolic cycles with the help of iron sulfides ('iron–sulfur world') was hypothesized by Günther Wächtershäuser (Wächtershäuser, 1992), based on work of Rolf Thauer and his colleagues (Thauer, Jungermann and Decker, 1977). In contrast to the iron–sulfur world, which relied on iron sulfides as reactants/electron donors and catalysts to fix CO_2, the theory around alkaline hydrothermal vent focused on the disequilibrium of two fluids, using iron sulfides as a kind of membrane to separate the alkaline hydrothermal vent fluid from the mildly acidic ocean to drive a pH gradient to electrochemically reduce CO_2 (Huber and Wächtershäuser, 1997; Russell, 2019). At the time this theory was proposed, it was based on experiments with chemical gardens in the laboratory, the existence of hydrothermal systems with alkaline effluents was not confirmed.

The discovery of the vents themselves occurred over 10 years later, when Deborah Kelley and her team discovered the Lost City alkaline hydrothermal field 15 km West of the Mid Atlantic Ridge (Kelley *et al.*, 2001; Kelley, Baross and Delaney, 2002). The alkaline pH is a consequence of a process called serpentinization, a water–rock interaction through which vast amounts of water are reduced to H_2, thus delivering the electron donor for the CO_2 fixation in the effluent. Later on, studies suggested that the hydrogen dependent reduction of CO_2 could be happening in serpentinizing systems abiotically, thus opening the possibility of prebiotic CO_2 fixation (Klein, Bach and McCollom, 2013; Schrenk, Brazelton and Lang, 2013; Früh-Green *et al.*, 2017). This process is explained in detail in Chapter III (Publication 6).

With increasing information on possible geological conditions and progress in microbiology and bioinformatics, autotrophic carbon fixation pathways like the reverse Krebs cycle (rTCA) or the reductive acetyl-CoA pathway became associated with chemical and geochemical reactions (Dayhoff and Eck, 1966; Thauer, Jungermann and Decker, 1977; Wächtershäuser, 1988; Huber and Wächtershäuser, 1997). This led to the idea that these biological pathways could be not only 'biochemical fossils' of the metabolic pathways in already existing cells, but also 'geochemical fossils' from prebiotic times, before enzymes and cofactors arose (Crabtree, 1997). The idea began to mature when geochemist Mike Russell and biologist Bill Martin started to collaborate. Martin's views on the pivotal role of hydrogen in carbon metabolism fit in the geochemical setting alkaline hydrothermal vents provided (Martin and Müller, 1998; Martin and Russell, 2003; Russell, 2019). The geochemical bottom-up

approach of the hydrogenation of CO_2 and the biological top-down approach, looking at molecular evolution in cells, led to the same conclusion: that CO_2 fixation with H_2 is extremely ancient (Martin and Russell, 2003; Russell and Martin, 2004; Martin and Russell, 2007; Russell, Hall and Martin, 2010). The far-from-equilibrium system in alkaline hydrothermal vents not only offered access to biology-like CO_2 reduction, it also offered small, almost cell-like compartments, in which the development of catalytic networks separated from the environment was imaginable, leading to complex structures like RNA and proteins (Martin and Russell, 2007). It should be noted that today, various sub-hypotheses have developed based on alkaline hydrothermal vents, a coherent alkaline hydrothermal vent theory does not exist. While some are proposing methane as a carbon source instead of CO_2, (Russell, Nitschke and Branscomb, 2013; Russell, 2019), others attach less importance to pH gradients or iron sulfides, as will be elaborated on in Chapter III of this dissertation.

Modern bioinformatics makes it possible to reconstruct the biochemical pathways of the first living cell, the last universal common ancestor (LUCA). LUCA links the abiotic phase of Earth with the first evidence of microbial life, which is almost 4 billion years of age (Weiss *et al.*, 2016, 2018; Tashiro *et al.*, 2017). While the methods are refining constantly, it seems that LUCA indeed lived from CO_2 fixation with H_2 via the acetyl-CoA pathway. This topic is covered in detail in Chapter II (Publication 4). What makes the reaction of these two compounds so appealing? To approach this question, one has to do so from different directions. Because in the beginning, this had to be an abiotic reaction, not a biotic one. Abiotic CO_2/H_2 reactions are found both in geology and industry. In order to understand how and why nature could have chosen this reaction as a starting point for life, no catalytic stone can be left unturned.

2 CO_2 fixation across the disciplines

2.1 Well known and still not fully understood: CO_2

Rarely are there single molecules, let alone their empirical formulae, that make it into common parlance. But the most oxidized form of carbon, carbon dioxide, has become so crucial for our understanding of Earth's climate that its chemical formula is known to the broad public.

CO_2 is one of the Earth's most important naturally occurring greenhouse gasses, and as such is directly responsible for making the planet warm enough for life as we know it to develop (Kiehl and Trenberth, 1997). Situated in the atmosphere, greenhouse gasses absorb infrared radiation coming from the Earth's surface and reflect parts of it back to the surface, thereby increasing the temperature to a pleasant 14 °C (Karl and Trenberth, 2003).

CO_2 is also the end product of the primary energy generation process of the 20th and 21st century, the combustion of organic material that was, millions of years ago, fixed by plants and microorganisms via photosynthesis. By burning fossil fuels, large amounts of reduced carbon are oxidized to CO_2 in short periods of time. The reverse process however, CO_2 fixation, is far more complex, needs more resources and consequently takes more time than burning the carbon equivalent of coal, oil or gas. This has led to a rapid increase of atmospheric CO_2 levels, from a pre-industrial value of about 280 ppm to 379 ppm in 2005 (Solomon *et al.*, 2007). Geologically, this is a very short period of time. According to measurements of the Earth System Research Laboratory at Mauna Loa on Hawaii, the current atmospheric global mean concentration of CO_2 is 407.75 ppm (August 2019) and rising (*Trends in atmospheric carbon dioxide*, 2019). This short-term rise of CO_2 concentration entails a concurrent increase of global temperature which, if not constrained, is predicted eventually to have uncontrollable impacts on the World's climate (Pachauri and Meyer, 2014).

Besides drastically reducing the industrial processes that release most CO_2 into the atmosphere, capturing and storing CO_2 in order to keep it out of the atmosphere (Carbon capture and storage, CCS) and transforming CO_2 back into more reduced carbon compounds (Carbon-to-Chem, C2C) are approaches that are being considered to decrease atmospheric CO_2 levels (D'Alessandro, Smit and Long, 2010; Li *et al.*, 2016).

Taking a closer look at this molecule that is so essential and simultaneously threatening to life on Earth as we know it, the chances and caveats of reducing it to organic compounds become clearer. Carbon dioxide is a linear, nonpolar molecule that holds strong polar bonds between carbon and oxygen, resulting in a short distance between the atoms (1.16 Å). The triple point of CO_2, where the three phases coexist in thermodynamic equilibrium, lies at 5.1 bar and –56.6 °C, which means at lower temperatures and higher pressures the gas becomes solid, at

higher temperatures and higher pressures, the gas liquefies. Above temperatures of 31.1.°C in combination with pressures above 73.8 bar, CO_2 turns supercritical (sCO_2), a state in which it exhibits properties between a gas and a liquid, a state that is widely used in industry for extractions and other applications (Cvjetko Bubalo *et al.*, 2015). The electronic structure, $O^{-\delta}$–$C^{+2\delta}$–$O^{-\delta}$, reveals that carbon is susceptible to a nucleophilic attack, meaning electron rich species can form a bond with it (Appel *et al.*, 2013). Oxygen, on the other hand, is open to an electrophilic attack from an electron poor molecule, atom or ion. CO_2 also has a large quadrupole moment that can be described as two electrical dipoles sitting back-to-back and pointing in opposite directions. This makes CO_2 interact well with polar environments—in solution, gas phase or within solid frameworks (Buckingham and Disch, 1963; Choi and Suh, 2009; D'Alessandro, Smit and Long, 2010). Compared to other atmospheric gases (H_2, N_2, O_2), CO_2 is therefore very soluble in water: at 25 °C under 1 bar of CO_2 gas, 0.033 mol/L CO_2 are dissolved in water (Appel *et al.*, 2013). Depending on the pH, either carbonic acid (H_2CO_3), bicarbonate (HCO_3^-) or carbonate (CO_3^{2-}) is the predominant species, although equilibrium for the hydration of CO_2 to H_2CO_3 is only achieved slowly without catalysts. Carbon dioxide acts as a Lewis acid (electron acceptor), while barely reacting with Brønsted (proton donor) or Lewis (electron acceptor) acids.

In order to be able to accept electrons, CO_2 has to be activated, meaning the C–O bonds have to be weakened. This weakening is mostly achieved via the bending of the usually linear molecule, making it easier for nucleophiles and electrophiles to interact with the corresponding orbitals (Fig. 2). The lowest unoccupied molecular orbital (LUMO) of CO_2—the molecular orbital (MO) susceptible to electrons and thus nucleophilic attacks—is an anti-binding π^* MO. Its wave function probabilities are strongly localized at the carbon atom, which enables the transfer of electron density from a nucleophile into the LUMO. Bending enforces this reaction by exposing the LUMO to the nucleophile. The highest occupied molecular orbital (HOMO) of CO_2 is a non-binding π MO (Fig. 2). Here, the electron density and wave function probability are localized as lone electron pairs of oxygen and thus open for reactions with electrophiles (Appel *et al.*, 2013).

The energy needed for activating the structural change that is the bending of the CO_2 molecule is reflected in the very negative electrochemical potentials for one-electron reduction of CO_2 to $CO_2^{-\bullet}$ ($E^{\circ\prime} = -1.9$ V at pH 7) (Frese, 1993), which means that this reaction is highly endergonic. In comparison, the coupled two-electron reduction from CO_2 to CO ($E^{\circ\prime} = -0.52$ V at pH 7) and CO_2 to aqueous formate ($HCOO^-$; $E^{\circ\prime} = -0.53$ V at pH 7) can happen under mild conditions.

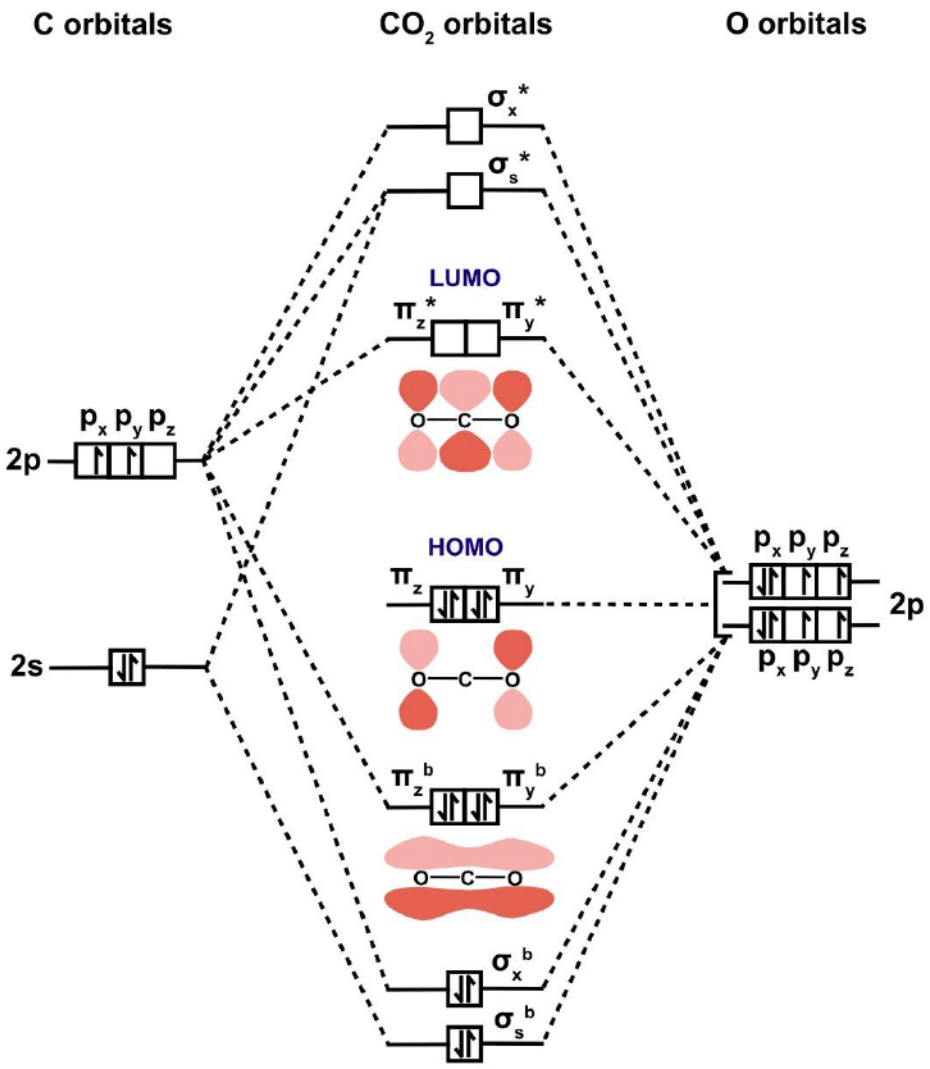


Figure 2: Molecular orbital (MO) energy level diagram for the linear CO_2 molecule. The depicted MOs are the binding π^b MO, the non-binding π MO (the highest occupied MO, HOMO) and the anti-binding π^* MO (the lowest unoccupied MO, LUMO. Redrawn and adapted from Keene, 1993.

CO_2 can be reduced either with strong nucleophiles or when it is activated/bent well enough. In biology, in order to avoid the necessity of strong nucleo- or electrophiles, nucleophilic and electrophilic reactions are often combined within the active centres of the involved enzymes to fix CO_2 (Appel *et al.*, 2013).

Biological pathways for CO_2 fixation use a variety of mechanisms and enzymes. From making C–H and C–C bonds to cleaving C–O bonds, nature has had billions of years to develop six (known) ways to process CO_2 (Fuchs, 2011). The first discovered and best known of these six is the reductive pentose phosphate (Calvin-Benson) cycle, found in various autotrophic prokaryotes and plants. It converts CO_2, NADPH and phosphate (P_i) from ATP into sugar

phosphates (Bassham, 1979). The other five autotrophic pathways that can turn inorganic into organic material found in prokaryotes are: the reductive citric acid cycle (rTCA), the dicarboxylate/4-hydroxybutyrate cycle, the 3-hydroxypropionate/4-hydroxybutyrate cycle, the 3-hydroxypropionate bi-cycle, and the reductive acetyl-CoA (or Wood-Ljungdahl) pathway (Berg *et al.*, 2007; Fuchs, 2011). If these five fixation pathways are compared, it becomes evident that the reductive acetyl-CoA pathway requires minimal energy and links anabolism and catabolism via its product, acetyl-CoA. The latter can be hydrolysed to acetate, a step that releases ATP (Berg *et al.*, 2010; Fuchs, 2011). This means, the acetyl-CoA pathway can be linked directly to an energy providing process, which is crucial for a metabolic cycle. The possible role of the pathway in building up a protometabolism is discussed in detail in the publications included in Chapters II and III. The intermediate steps of the acetyl-CoA pathway can be identified, and, as in industrial processes, transition metals play a key role, being embedded in the active centres of the enzymes involved (Volbeda *et al.*, 1995; Chabrière *et al.*, 1999; Dobbek *et al.*, 2001; Jeoung and Dobbek, 2007; Martin, 2019).

While industrial CO_2 fixation processes are far less refined than those in prokaryotic cells, research on homogeneous (catalyst in the same phase as reactants) and heterogeneous (catalyst and reactant phases differ) catalysts for CO_2 reduction can help to deepen the understanding of how C–C and C–H bonds form or C–O bonds are cleaved in biology. At the same time, much knowledge can be retrieved from biology for industrial conversion of CO_2. Suitable catalysts would constitute a turning point in energy production (Bell, Gates and Ray, 2007). A good starting point for the search of such catalysts could be the metal-bearing active centres of the enzymes that have been catalysing metabolic reactions of carbon for billions of years. The activation of the CO_2 molecule seems to be the crucial point for its reduction. And it can be promoted by catalysts. Recent calculations have shown that the CO_2 molecule interacts strongly with magnetite (Fe_3O_4) and other surfaces, causing the linear molecule to bend and thus activating it (Álvarez *et al.*, 2017; Santos-Carballal *et al.*, 2018). However, in order for reduction to happen, not only does CO_2 need to be in the right state, but so do the electrons. These electrons to a large extent—both in biology and chemical industry—come from a second gas: hydrogen (H_2).

2.2 Source of electrons and energy: H_2

The main source of electrons for chemoautotrophic, industrial and geochemical CO_2 fixation is dihydrogen (Sleep *et al.*, 2004; Fuchs, 2011; Porosoff, Yan and Chen, 2016). There are two main sources of naturally occurring H_2: abiotic geochemical production (serpentinization) or biotic biochemical production (hydrogenases, fermentation). Today, anaerobic autotrophs such as methanogens grow mainly on H_2 of biotic origin, approximately 150 million tons per year are produced by microorganisms and consumed by methanogens (Thauer *et al.*, 2008, 2010). Only a small fraction sustains itself from geochemically generated H_2, for instance in hydrothermal vents (Thauer *et al.*, 2010). When the first microbial lineages evolved, however, the abiotic H_2 had to play a bigger role to supply the energy and electrons of ancient metabolism (Thauer *et al.*, 2010).

Hydrogen is a crucial electron source in chemical industry. It is central for all hydrogenation processes, and for any kind of chemical reaction in which hydrogen is added to another compound thereby reducing it. Catalytic hydrogenations are considered to produce the largest product volume of all man-made chemical reactions (Kubas, 2007). For instance, all crude oil is treated with hydrogen to remove non-carbon contaminations like sulfur or nitrogen. Nitrogen is hydrogenated in large volumes as well. 141 million tonnes of nitrogen (N_2) were reduced to ammonia for production of fertilizers, plastics, fibres and explosives in 2015 (Kelly and Matos, 2015). In general, H_2 on its own does not react very well, the two electron H–H bond is very strong. Hydrogen can only become chemically useful when this bond is broken and the two H-atoms are separated (Kubas, 2007). The homolytic cleavage of gaseous H_2, in which the H–H bond is broken in the middle (Fig. 3a), is very endergonic by +436 kJ mol^{-1}. The heterolytic cleavage—meaning one H getting both electrons of the bond, becoming a hydride (H^-), and leaving a proton (H^+) behind (Fig. 3b)—is less endergonic, but still requires +200 kJ mol^{-1}. The mechanisms of H–H bonds splitting to form a metal dihydride complex have been studied for a long time (Kubas, 2007).

The activation of H_2 is challenging and the mechanistic details are of great interest for the development of suitable catalysts. H_2 was long assumed to be inert to further chemical interaction, except in the case of weaker versions such as physisorption, where mainly van der Waals force is acting between the adsorbed molecule and the surface (Kubas, 2007). Only in the 1980s was an intact H_2 molecule shown to be coordinating 'side-on' to a metal complex by donating its two σ-bond electrons to a vacant d-orbital of the metal (Kubas *et al.*, 1984), leading to a stable, isolable σ-complex. Based on this discovery, further investigations made it possible to sketch a mechanism from metal bound H_2 molecules to metal dihydrides (Fig. 3a).

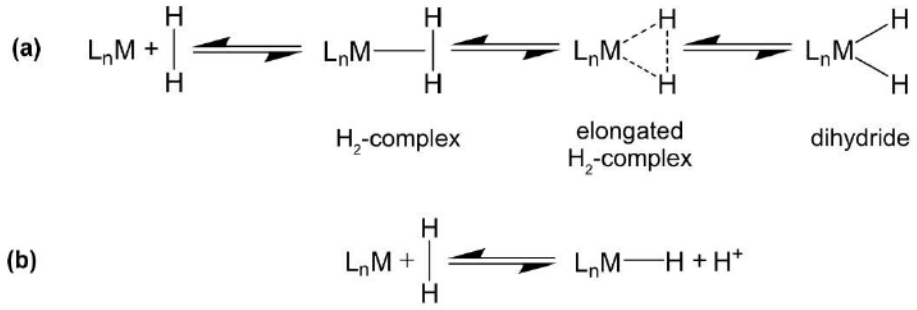


Figure 3: (a) Binding and splitting of H_2 on a transition metal complex (homolytic cleavage); (b) Heterolytic cleavage of H_2 on a transition metal complex. L = Ligand; M = Metal. Redrawn and adapted from Kubas 2007.

Metal dihydrides as a result of oxidative addition of H_2 to a metal centre have long been known to be part of catalytic cycles of hydrogenation (Chalk and Halpern, 1959; Halpern, 1980). Investigation of H_2-complexes shows that the H_2 molecule is not physisorbed but chemisorbed (involving the formation of chemical bonds between adsorbed atoms or molecules and the metal), facilitating the splitting of the H–H bond. The formation of side-on H_2-complexes are the first step of the H–H bond cleavage (Kubas *et al.*, 1984; Kubas, 2007).

Biology, of course, figured out how to activate and cleave H_2 about 4 billion years ago. Autotrophic prokaryotes use hydrogenases, redox enzymes that were already present in LUCA (Weiss *et al.*, 2016). H_2 was the source of electrons for primary production before photosynthesis emerged (Martin, Bryant and Beatty, 2018). Hydrogenases can either use H_2 as an energy source or dispose of leftover electrons as H_2. There are three different kinds of hydrogenases, all of them holding transition metals coordinated by varying ligands in their active centres: [NiFe] (tracing back to LUCA), [FeFe], and [Fe] (Fontecilla-Camps and Ragsdale, 1999; Thauer *et al.*, 2008; Thauer, 2011; Lubitz *et al.*, 2014). Here, similar side-on coordination and bond cleavage is expected as observed in inorganic homogeneous catalytic processes (Kubas, 2007).

Most of the intermediate constellations of the active centres of the three hydrogenase-types have been determined via crystallographic studies, Fourier-transform infrared (FTIR), and electron paramagnetic resonance (EPR) spectroscopy (a method using magnetic fields to study materials with unpaired electrons), radionuclide labelling, and density functional theory (DFT) calculations (Lubitz *et al.*, 2014). All three hydrogenases perform a heterolytic cleavage of H_2, but do so in mechanistically different ways; the same applies to the reverse reaction, the formation of H_2. The first step in heterolytic cleavage is to polarize the bond at an open metal side, so while the proton (H^+) is accepted by a nearby base ligand, the hydride (H^-) remains

bound to the metal. In the case of [FeFe] hydrogenases, H^- is bound end-on to one of the two Fe atoms (Kubas, 2007; Siegbahn, Tye and Hall, 2007; Lubitz *et al.*, 2014). In [NiFe] hydrogenases, H^- binds to both Ni and Fe (with help of an extra electron coming from Ni), making the hydride more stable than the one in [FeFe] hydrogenases (Kubas, 2007; Siegbahn, Tye and Hall, 2007; Lubitz *et al.*, 2014). The mechanism of [Fe] hydrogenases is less well resolved (Lubitz *et al.*, 2014). In all hydrogenases, as in inorganic metal complexes, ligands such as cyanide (CN^-) or carbon monoxide (CO) have great impact on the catalytic activity, as they can draw electrons in and out of the complex, stabilizing or destabilizing certain spin states of the transition metals (Kubas, 2007; Lubitz *et al.*, 2014).

Carbon–metal bonds, just like the ones formed between ligands and metals, are involved in central enzymes and cofactors found in LUCA, hinting towards their possible role in ancient chemical evolution (Martin, 2019). Electrons from H_2 are transferred to electron acceptors like NAD^+, F_{420} or oxidized ferredoxins (proteins containing iron–sulfur clusters) for entry into metabolism. While NADH and F_{420} can be considered H^--donors (two electrons), ferredoxin donates single electrons. There is no known example of direct hydrogen transfer in biology similar to synthetic hydrogenation. The path from activated H_2 to an ultimate electron acceptor such as CO_2 traverses intermediate electron carriers. Investigating the mechanisms of hydrogenases is of great interest for biotechnological hydrogen production (Friedrich, Fritsch and Lenz, 2011) or biofuel cells for hydrogen oxidation (Vincent *et al.*, 2005), and to provide a blueprint for designing new catalytic systems (Kubas, 2007; Lubitz *et al.*, 2014).

In current industrial practice, hydrogenation is the main large-scale hydrogen reaction. While homogeneous catalysis is a new development for hydrogenation, heterogeneous catalysts have been around since the late 19th and early 20th century. The Sabatier process, for instance, the synthesis of CH_4 from CO and H_2 over nickel catalysts, was first described in 1897. In the subsequent decade, a large-scale process for the hydrogenation of nitrogen over iron catalysts was developed (also known as the Haber–Bosch process). The mechanisms have been investigated ever since (Ertl, 1983; Leigh, 2004; Kandemir *et al.*, 2013), including of course, how hydrogen reacts on and interacts with catalytic surfaces. As in homogeneous catalysis, hydrogen can be adsorbed both as a molecule by physisorption and as H-atoms by dissociative chemisorption. Physisorption of H_2 on surfaces only takes place under low temperature. For chemisorption, hydrogen has to overcome the activation barrier and thus higher temperatures are required. If the kinetic energy of the molecule is high enough, H_2 can be dissociated directly on the surface of the catalyst. But also indirect chemisorption starting from the transient

physisorbed state is possible; physisorbed H_2 molecules diffuse quickly on the surface and then dissociate when the right catalytic site is met (Pisarev, 2012).

In summary, H_2 has to be activated to donate its electrons, CO_2 needs to be activated to accept them. Both activations have to occur on the same catalytic surface in order to enable CO_2 reduction. This applies to industrial hydrogenation, to biological CO_2 fixation, and probably also applied to the origin of life.

2.3. From surfaces to enzymes

As outlined in the foregoing section, CO_2 and H_2 can hardly react without being activated, their intermolecular bonds are too strong and the activation energies are too high, unless they meet on a compatible catalyst. Once the kinetic barrier is surmounted, the subsequent reactions can run their course, depending on the conditions and the catalysts. The heterogeneously catalysed reaction between CO_2 by H_2 has been studied extensively (Riedel *et al.*, 2001, 2003; Porosoff *et al.*, 2015; Rodriguez *et al.*, 2015; Porosoff, Yan and Chen, 2016; Wang, Feng and Bao, 2018). The development of catalysts suitable for large-scale processes is still a challenge, mainly because so little is known about why some catalysts work better than others.

Three types of reactions are established in industrial heterogeneous hydrogenation of CO_2: i) the reverse water gas shift (RWGS) reaction to form CO; ii) selective hydrogenation via formate, adsorbed to the surface (*HCOO), which through several subsequent carbon and oxygen hydrogenations and C−O cleavage leads to the formation of methanol (CH_3OH) and water (Appel *et al.*, 2013); iii) combination of CO_2 reduction with Fischer–Tropsch (FT) reactions, the latter being a hydrogenation reaction starting from CO. There, the initial hydrogenation of CO_2 forms adsorbed carboxylate (*HOCO) species that dissociate to form CO and OH. The intermediate CO then either desorbs from the catalyst or undergoes further hydrogenation reactions to form CH_3OH, CH_4 or other hydrocarbons ($(-CH_2-)_x$).

In all these pathways, CO_2 reduction is initiated by a hydrogen transfer when both compounds are bound to the surface (Porosoff, Yan and Chen, 2016; Böller, Durner and Wintterlin, 2019). The interactions between intermediate species and catalytic surfaces influence the product distribution. For example, catalysts that bind carbonyls (*CO) strongly favour methanol and hydrocarbon formation while weaker binding leads to free CO in the gas phase (Porosoff, Yan and Chen, 2016). Additionally, the type of catalyst surface, and even just small amounts of impurities, make a difference in product distribution. The literature on which metals, additives, and surface compositions suit a certain catalytic process is vast, and illustrates that catalyst development still relies on trial-and-error approaches because the efficiency of a

catalyst is hard to anticipate before it is tested (Zambelli *et al.*, 1996; Wintterlin *et al.*, 1997; Porosoff *et al.*, 2015; Rodriguez *et al.*, 2015; Böller, Durner and Wintterlin, 2019).

Magnetite (Fe_3O_4) is one example among many industrial catalysts. It is also one of the main components of industrial heterogeneous catalysts because it is stable, available and cheap. In theory, the mixed valence state of the Fe-ions (Fe^{2+}/Fe^{3+}) enables the iron oxide to catalyse both oxidation and reduction reactions (Cornell and Schwertmann, 2003; Santos-Carballal *et al.*, 2018). Fe_3O_4 can accept and donate electrons, theoretically making it ideal for reactions with both electron acceptors and donors such as CO_2 and H_2 (or N_2 and H_2). In practice, the situation is more complicated. In the Fischer–Tropsch process, for instance, Fe_3O_4 is used, but it is not the working catalyst. The latter is a mixture of α-Fe (iron with a body-centred cubic crystal structure) and the iron carbide Fe_5C_2, both products of the reaction between the FT-reactants CO and H_2 (syngas) and Fe_3O_4 (Satterfield *et al.*, 1986; Hou *et al.*, 2012; Spreitzer and Schenk, 2019).

Apart from the catalyst, there are other parameters that can influence the outcome of the H_2/CO_2 reactions. These include the composition of the gas mixture, the reaction temperatures, pressures, and, if aqueous, the pH. Each of these factors significantly affects the reaction intermediates and products. That means that the same sets of reactions performed under less well-defined conditions, as found in geological settings, lead to a broader product spectrum than a streamlined industrial synthesis.

These 'messier' or 'less yield oriented' geochemical conditions for CO_2 fixation have been explored in many directions. H_2 was not necessarily the primary electron source in such experiments, but native metals, electrodes coated in metal sulfides, and metal sulfides (and oxides) have been used in aqueous solution (Chen and Bahnemann, 2000; Guan *et al.*, 2003; He *et al.*, 2010; McCollom, 2013a, 2016; Roldan *et al.*, 2015; Miller *et al.*, 2017; Varma *et al.*, 2018). In many cases in which H_2 was not added, its presence as an intermediate, however, cannot be excluded. Such studies constitute an important bridge between catalytic chemistry and geochemistry and also focus on the important role of minerals as prospective catalysts for CO_2 fixation (Camprubi *et al.*, 2017). Geochemists have investigated the chemical reactions that occur naturally in serpentinizing systems, detecting low molecular weight carbon compounds such as formate and methane (Kelley and Früh-Green, 1999; Holm and Charlou, 2001; Lang *et al.*, 2010; Schrenk, Brazelton and Lang, 2013; Etiope and Schoell, 2014; Etiope and Ionescu, 2015; Konn *et al.*, 2015). Recently, even complex carbon compounds, including the aromatic amino acid tryptophan, have been found preserved in rock samples from serpentinizing systems (Ménez *et al.*, 2018). While it is possible to distinguish carbon

compounds from abiotic and biotic origin via carbon and hydrogen isotope signatures, the mechanisms of the carbon compound synthesis inside such geochemical systems are still unclear. Nevertheless, hydrogenation-like reactions between CO_2 (dissolved or bound as carbonate) and the constantly synthesized H_2 are a very probably source, especially looking at laboratory simulations under comparable conditions (Horita and Berndt, 1999; Miller *et al.*, 2017; Varma *et al.*, 2018). This thesis will pick up these approaches and elaborate on them in Chapter III.

The question is, how does a geochemical process transform into something as elaborate as the reductive acetyl-CoA pathway? The answer, whatever it is, almost certainly has to do with catalysts. If the answer involves transition metal-containing surfaces, transition metal complexes or transition metal-incorporating enzymes, then there would be continuity in evolution from rocks as catalysts to proteins that incorporated the catalytic properties. This basic idea has been common currency among biologists at least since the discovery of FeS centres in ferredoxin (Dayhoff and Eck, 1966; Thauer, Jungermann and Decker, 1977; Wächtershäuser, 1988). If we look at the formate route of CO_2 hydrogenation to methanol, it has definite parallels with the methyl branch of the acetyl-CoA pathway (Fig. 4).

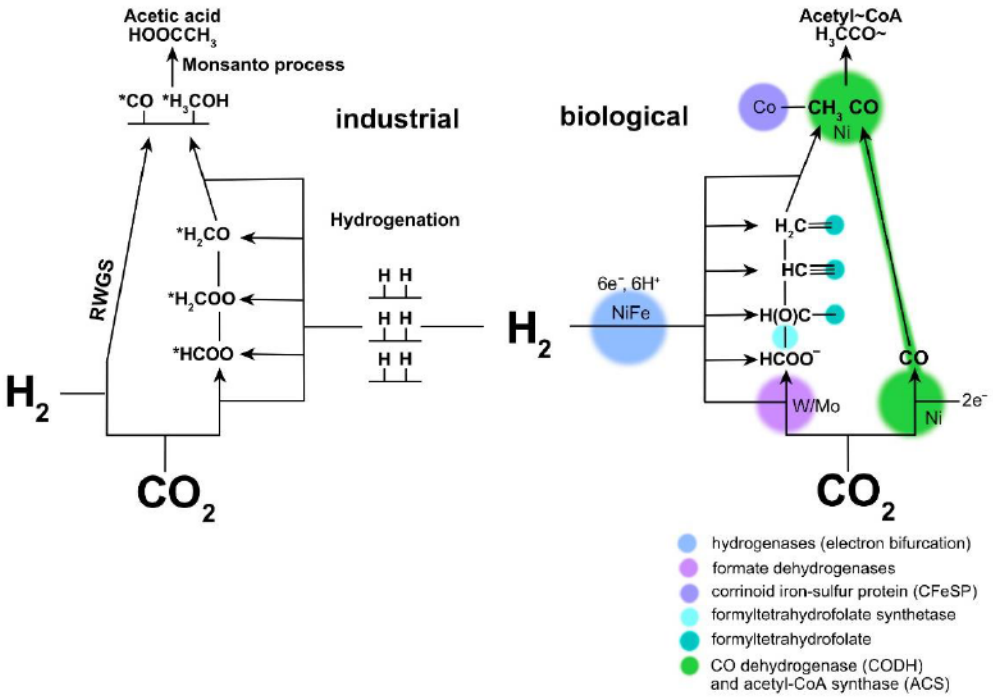


Figure 4: The formate route in industrial hydrogenation processes for methanol synthesis (left) and the methyl branch of the acetyl-CoA pathway (right) share mechanistic parallels. The last step of the acetyl-CoA pathway, the carbonyl insertion between the methyl group and the metal centre of the according

catalyst, also has an equivalent in industrial processes. Drawn based on information in Fuchs (2011) and Porosoff, Yan and Chen (2016).

In the industrial hydrogenation process, another H-atom adds to the adsorbed formate (*HCOO), forming adsorbed dioxomethylene (*H_2COO), and from there, one H atom at a time, methanol (*CH_3OH) is formed on the surface (Behrens *et al.*, 2012; Appel *et al.*, 2013; Porosoff, Yan and Chen, 2016). In the methyl branch of the acetyl-CoA pathway, similar steps can be observed, the intermediate groups being bound to tetrahydrofolate and tetrahydromethanopterin, respectively (Fuchs, 2011; Preiner, Igarashi, *et al.*, 2019). Also the CO branch, which in the acetyl-CoA pathway is united with the methyl branch by the acetyl-CoA synthase (ACS) to form acetyl-CoA, has a parallel in the abiotic approach: the RWGS reaction (Ragsdale and Kumar, 1996).

The reaction between CO and the methyl group, biotically catalysed by the acetyl-CoA synthase (ACS), also has an industrial equivalent: the carbonylation of methanol. The industrial approach to synthesize acetic acid, now commonly known as the Monsanto process, was described in the 1960s by BASF and Monsanto, both using homogeneous catalysts, namely iodide-promoted metal complexes. The catalyst presented by BASF was based on cobalt, the Monsanto catalyst on rhodium (Fuchs, 1986). Later, an iridium based version, the Cativa process, was developed (Sunley and Watson, 2000). All three metal complexes catalyse the insertion of CO into the carbon metal bond that is formed between the methyl group of methanol and the chosen transition metal (Forster, 1979). A very similar mechanism is proposed for the nickel bearing active site of ACS (Can *et al.*, 2017).

The backbone of CO_2 fixation with H_2 is possibly a highly conserved reaction in biology, as suggested by both biological and chemical data (Weiss *et al.*, 2016; Varma *et al.*, 2018). This observation is interesting in an OoL context because not only is the biological pathway ancient, it is compatible with geochemical observations in serpentinizing systems such as alkaline hydrothermal vents, constantly producing a source of chemical energy (H_2) and the prospective catalysts to utilize that energy. Iron oxides, iron sulphides, and zero valent intermetallic compounds can be found inside these systems, being renewed steadily by the far-from-equilibrium system of which they form part (Klein and Bach, 2009). There is still little known about the processes going on inside of these geochemical settings, hence they offer opportunity for investigation and constitute a good starting point for the experimental part of this thesis, the results of which are presented in Chapter III (Publication 7).

I Past and current concepts of origin of life research

As outlined in Chapter 1, chemical research on the OoL has about a century of history. It has led to a diversity of ideas and has produced literature that became exponentially vast with time. But one must start somewhere. This attempt is the subject of the two publications presented in this chapter. Publication 1, an encyclopaedia article, presents a research starting point, a stepping stone into OoL research. Publication 2 shows discussions, revelations, debates and progress concerning the OoL research and aims to forge connectivity between disciplines and different premises.

.

Publication 1

Title:	Origin of life, Theories of
Year:	2017
Authors:	William F. Martin and Martina Preiner
Published in:	*Reference Module in Life Sciences*, Elsevier Inc., Oxford. doi:10.1016/B978-0-12-809633-8.02403-1.
Contribution:	Second author. Medium: provided extensive literature research and contributed towards writing parts of the text.
Summary:	This encyclopaedia chapter provides an overview of the most pivotal issues of origin of life (OoL) research, while giving a historical context to the different approaches.

Origin of Life, Theories of

William F Martin and Martina Preiner, Heinrich Heine University, Düsseldorf, Germany

There are many facets to the problem of understanding life's origin and equally many ways to address it. The origin of life can be viewed from a variety of different standpoints: information theory (Yockey, 2005), RNA replication (Eigen and Schuster, 1977), meteorite impacts (Brack, 2009), physics (Smith and Morowitz, 2016), specific chemical synthesis (Powner *et al.*, 2009), geochemistry (Martin and Russell, 2003), or entropy (Russell *et al.*, 2013), to name a few. Common to all current theories for the origin of life is the view that some kinds of molecules, probably similar to nucleic acids, were able to replicate and undergo selection for the ability to replicate prior to the advent of fully fledged free living cells.

When it comes to addressing the origin of life as we know it, as opposed to considering forms of life that are imaginable, it is helpful to connect ideas on the origin of life with the biology of modern microbes. This constrains the thoughts on the topic from becoming too theoretical or otherwise decoupled from life (that which is to be explained). The properties of life and the nature of living things focus thoughts about how or where life might have been arisen. Understanding which forms of life among the many that we know are likely to be the most primitive is also important in understanding life's origin, because it helps to narrow down the many possibilities when it comes to trying to narrow the gap between spontaneous chemical processes on the early Earth and biological processes in microbes with primitive metabolism. In the literature on the origin of life, four attributes of life recurrently come up as central to the issue: compartmentation, information, replication, and energy.

Compartmentation

Because all life that we know is organized as cells, there is every reason to assume that the first forms of life were also cellular. It also means that in the time before there were fully fledged free-living cells, there were simpler entities that were not capable of independent replication in the wild (Gantí, 1975). Compartmentation is important for two reasons. First, whatever chemical synthesis one has in mind for the synthesis of life's building blocks, some kind of concentrating mechanism or barrier has to exist such that the products of synthesis do not diffuse off into the oceans (Kuhn, 1972). Second, for any form of evolution to take place among replicating molecules, the population of molecules has to be structured in some manner in order to permit selection to act. That is, selection for new populations of molecules with new and different properties, for example, a catalytic activity or efficiency of replication, cannot occur unless the molecular populations are distinct. In an unstructured population of molecules, the fastest replicators, also called parasites, will prevail (Branciamore *et al.*, 2009). There are various thoughts about how early molecular systems might have been compartmentalized from the environment to permit concentration and selection, these include evaporation on land, freezing in ice (eutectic mixtures), or naturally forming inorganic compartments at hydrothermal vents (Martin and Russell, 2003).

Information

All forms of life have the universal genetic code. It is therefore reasonable to assume that a form of information processing that involves the modern genetic code was also present in the very first forms of life as well. That is consistent with the observation that the genetic code is one of the oldest and least modified traits of life as we know it (Wong, 1975). The genetic code is quite complex. It requires the accurate loading of coding amino acids onto tRNA, tRNA–mRNA interactions mediated by the ribosome (and ribosomal RNA). It is likely that earlier forms of the code were simpler than the modern code, using a two letter code, for example, rather than three letter triplets. The issue of how the code arose is an unresolved problem (Koonin and Novozhilov, 2009), although theories strive to link the origin of the code to metabolism by suggesting that the amino and oxygen moieties in the bases of the tRNA anticodon provided catalytic activity that synthesized the cognate amino acid on the tRNA acceptor stem, as opposed to linking preformed amino acids with tRNA (Copley *et al.*, 2007).

Replication

Replication has long been viewed as one of the most important aspects at life's origin. For anything to evolve in a Darwinian sense, it must generate copies of itself that differ to some degree from the original (natural variation) and exhibit properties that would permit their differential success in a proliferation sense (natural selection). Since the pioneering work of Sol Spiegelman and Manfred Eigen with a viral enzyme called Qβ replicase, replication of RNA has stood in the foreground of thinking on the issue (Spiegelman, 1965; Eigen, 1971). RNA can be replicated (like DNA) and can exhibit catalytic activity (proteins) making it attractive as a precursor to both in evolution. Eigen's work in particular demonstrated the differential success of different RNA templates to replicate in vitro. That established the concept of natural variation and natural selection among molecules prior to the origin of variation and selection among cells. This aspect, in addition to replication plus catalysis, gave rise to the concept of

evolution in an RNA world that preceded the evolution of cells (Gilbert, 1986). RNA evolution experiments are widely conducted under laboratory conditions (Lincoln and Joyce, 2009). However, on the early earth, the chemical environment was harsh. There were only rocks and water, there was the constant threat of diffusion into the ocean, UV light was abundant, as was constant meterorite impact and widespread volcanic activity. Such conditions have little if any similarity with modern laboratory conditions, a circumstance that acts as a kind of watershed in origin of life theories as concerns the role of RNA. Proponents of the RNA world point to the experimental accessibility of RNA replication as a model for early evolution, while geochemists point to the lack of known environmental conditions that would allow an RNA world to arise or evolve.

Part of the replication aspect of origins concerns speed. The fastest replicators will incorporate all environmentally available monomers into polymeric likenesses of themselves, thereby greedily consuming all available resources. Hence they will survive much better and dominate under molecular selection schemes than other sequence variants (stability to hydrolysis being equal among different sequences). But a fast replicating sequence variant might mutate so as to generate even faster variants, in which case they would outcompete the slower ancestor and follow an evolutionary trajectory toward the fastest replicating sequence. Once that locally optimal, fastest replicating sequence has evolved, mutational variants drift away from the optimum and generate slower replicating progeny while cluttering the sequence space (the sum of all possible sequence variants around the fastest replicator). That concept of a locally optimal replicator and its less optimal variants was developed by Eigen in the context of replication; such a replicator together with its variants is called a quasispecies (Eigen and Schuster, 1977).

Energy

Perhaps most essentially, all forms of life harness one or more sources of energy that are available in the environment. No form of replication or selection or evolution among either molecules or cells can take place without energy release hence energy supply, as stipulated by the second law of thermodynamics. As a consequence of that constraint, we can be rather certain that prior to the advent of biological cells, prebiotic chemical reactions had to occur spontaneously. All spontaneous chemical reactions entail, in turn, an overall exergonic reaction that supplies the free energy change needed for the reaction to take place (Thauer *et al.*, 1977). This does not mean that every single reaction in prebiotic evolution had to be exergonic. Many reactions in modern metabolism are endergonic. But endergonic reactions are necessarily coupled to a main energy releasing reaction that the cell harnesses in order to drive metabolism and the life process forward. In modern metabolism, the most widespread form of coupling is through adenosine triphosphate (ATP). ATP hydrolysis releases energy, such that endergonic reactions can be coupled (enzymatically) to ATP hydrolysis such that the overall reaction is energetically downhill and will thus go forward (Sousa and Martin, 2014). There are other forms of energy currencies in metabolism besides ATP, these include thioesters (Goldford *et al.*, 2017), acyl phosphates (Schönheit *et al.*, 2016), and reduced ferredoxin, a small soluble iron–sulfur cluster containing protein that transfers low potential electrons in metabolism (Eck and Dayhoff, 1966). These currencies are synthesized in what biologists call energy metabolism (the main bioenergetic reaction of the cell), and are used to drive other reactions forward. At the origin of life, there must have been some sort of energy supply that drove the synthesis of organic compounds forward to a level where complex polymers (peptides and nucleic acids) were formed. Among the possible sources of environmental energy most often discussed are ultraviolet light, lightning, heat, and chemical energy (Kaufmann, 2009). If we look around among anaerobic chemoautotrophic microbes, only chemical energy is harnessed for energetic coupling.

As seen from the standpoint of biology, life is a side product or a byproduct of the main energy releasing (exergonic) reaction that serves to drive all other reactions in the cell forward. Put more succinctly, life is an exergonic chemical reaction. The synthesis of the very first chemical components of which cells are composed, often called the building blocks of life, must also have been the products of exergonic reactions, because the second law of thermodynamics stipulates that chemical reactions can only go forward if the overall reaction releases energy.

Molecular Networks

Living cells are side products of a main energy releasing reaction. Stated another way, metabolism generates life, not vice versa; the main flux of carbon, energy, and nutrients through the cells generates the substance and organization of cell mass. For example, acetogens channel 24 molecules of CO_2 through the cell as waste product (acetate) for every molecule of CO_2 that they incorporate into cell carbon (Daniel *et al.*, 1990). One of the longstanding puzzles in the origin of metabolism is the emergence of biochemical networks. Recent progress on that topic has come from the investigation of autocatalytic networks (Hordijk and Steel, 2004; Hordijk *et al.*, 2012), which have shown that the elements of a chemical network need not provide much catalysis for other steps in the network in order for flux through the system, hence net synthesis of the system itself, to increase. In the laboratory, chemical networks with oscillatory properties have been constructed, using thioesters as the chemical energy source (Semenov *et al.*, 2016).

Two Main Theories for Carbon

As recently summarized elsewhere (Schönheit *et al.*, 2016), a convenient structure to the problem is obtained if we sort theories on the origin of life into two categories based upon the assumptions they entail regarding the nature of carbon metabolism in the

earliest cells. By that criterion, theories about the earliest phases of evolution fall into two main categories: autotrophic origins versus heterotrophic origins. Theories for autotrophic origins posit that the first cells fulfilled their carbon needs from CO_2. Heterotrophic origin theories are based on the idea that the first cells lived from the fermentations of reduced organic compounds present in some kind of rich organic soup. The main contours of the two theories as they interface with biology can be summarized briefly.

Heterotrophic Origins

Heterotrophic origin theories are traditionally favored by chemists (Bada and Lazcano, 2008; Miller and Orgel, 1974; Orgel, 2008). Seen from a biological perspective, heterotrophic origin theories have two main drawbacks. First, they do not connect with the chemistry of modern cells. They start with cyanide (Levy *et al.*, 1999), formamide (Saladino *et al.*, 2012), or UV light-dependent (Ritson and Sutherland, 2013) condensations. The conditions underlying heterotrophic origin theories are sometimes viewed by geochemists as unlikely to ever have existed on early Earth (McCollom, 2013). Organic compounds delivered from space are seen as a source of carbon for the first cells in some heterotrophic origin theories, however, carbon from space is an unfermentable substrate and is furthermore chemically too heterogeneous to have served as a carbon source for the first cells (Schönheit *et al.*, 2016).

Autotrophic Origins

Autotrophic origin theories, by contrast, generally operate with chemicals that undoubtedly did exist on the anaerobic early Earth: H_2, CO_2, N_2 and H_2S. Transition metals and transition metal sulfide (FeS and FeNiS) centers play a crucial role in autotrophic origin theories for several reasons. First, FeS and FeNiS centers serve as essential catalysts in the modern day core carbon and energy metabolism of anaerobic autotrophs, as such they provide direct links between the origin of life and modern cells (Eck and Dayhoff, 1966). Second, FeS and FeNiS minerals abound in anaerobic geological settings both today and on the early Earth (Wächtershäuser, 1992). Third FeS and FeNiS centers are naturally catalytic by virtue of their unfilled d and f electron orbitals, which can readily hybridize to generate metastable bonds with carbon and nitrogen. Under the theory of autotrophic origins the first cells satisfied their carbon needs from CO_2 (Schönheit *et al.*, 2016).

Serpentinization

In recent years, geochemical processes in modern hydrothermal vents have been characterized that resemble biological processes germane to autotrophic origin theories. This development can perhaps be best summarized with the main keyword "serpentinization" (Russell *et al.*, 2010). During serpentinization, water is drawn by the force of gravity into cracks in the Earth's crust. At depths of one to several kilometers and temperatures around 200°C, water reacts with Fe^{2+} in the iron–magnesium silicates that comprise the crust. Water is thereby reduced to generate molecular H_2. The effluent of hydrothermal systems can contain up to 10 vol% or more H_2 (Etiope *et al.*, 2011). H_2 is a useful currency of chemical energy for many modern microbes, in particular for anaerobic autotrophs.

Geochemists have also discovered a modest variety of chemical reactions that are occurring naturally at hydrothermal vents today that entail the synthesis of methane (CH_4) and other small molecular weight carbon compounds such as formate and short chain hydrocarbons within hydrothermal systems (Schrenk *et al.*, 2013; McDermott *et al.*, 2015; McCollom, 2016). Methane can be present in the effluent of hydrothermal vents at up to 2 mmol kg^{-1} (Schrenk *et al.*, 2013) and can constitute up to 87 vol% of the vent gas emission (Etiope *et al.*, 2011), hydrocarbons in much lower amounts, formate has only been reported for some systems (Schrenk *et al.*, 2013). The mechanism of methane synthesis in hydrothermal systems is still uncertain (McCollom, 2016). It might entail reduction of dissolved CO_2 or of rock-bound carbonate. Whether H_2 or reduced transition metal species in the crust are the source of electrons for CO_2 reduction is not known.

Some modern microbes called methanogens belonging to the domain archaea synthesize methane from H_2 and CO_2 or C_1 compounds (Thauer *et al.*, 2008; Mayumi *et al.*, 2013). To determine whether the methane produced in hydrothermal vents is of biological (methanogenesis) or of geochemical (serpentinization) origin, geochemists use stable isotope ratios that involve measuring the $^{13}C/^{12}C$ and $^{2}H/^{1}H$ ratios of the respective compounds. Such measurements cannot be fully explained here; let it suffice to say that the $^{13}C/^{12}C$ and $^{2}H/^{1}H$ isotope signatures of methane produced by biological and geological processes can be distinguished, geochemically produced methane being typically enriched in the heavier carbon isotope and depleted in the heavier hydrogen isotope relative to biogenically produced methane (Etiope *et al.*, 2011; Pedrera *et al.*, 2016). In laboratory scale serpentinization reactions, acetate and formate synthesis has been reported (Miller *et al.*, 2017).

Ancient Microbial Lineages

Various lines of evidence suggest that anaerobic autotrophs are the most ancient lineages among microbes known today. Modern anaerobic autotrophs thrive upon H_2, which is continuously generated in serpentinizing geological settings at activities of the

order of 10 mmol kg^{-1} or more via disequilibria driven by rock–water interactions in hydrothermal systems. At such high H_2 activities, and under strictly anaerobic conditions, the synthesis of cell mass from CO_2 is thermodynamically favorable. The core pathway of carbon and energy metabolism in anaerobic autotrophs that inhabit such hydrothermal and deep crust environments is the acetyl-coenzyme A (CoA) pathway, the most ancient of the six pathways of CO_2 fixation known (Fuchs, 2011) and the only one present in archaea and bacteria.

The observation that spontaneous exergonic organic syntheses from H_2 and CO_2 occur today at hydrothermal vents suggests that these processes are similar, if not homologous (Martin, 2012), to core energy-releasing reactions of carbon and energy metabolism in methanogens (Thauer *et al.*, 2008) and acetogens (Schuchmann and Müller, 2014), which live from the reduction of CO_2 by H_2. Chemists are less taken by the similarity of the overall exergonic processes and tend to doubt in rather general terms that there are any traces of ancestral metabolic processes retained in any modern metabolic pathways (Orgel, 2008). However, the evidence is fairly solid that microbial communities have been thriving in hydrothermal vents for over 3.3 billion years (Westall *et al.*, 2012). From the biological standpoint, there is no strong a priori reason to assume that the energy releasing reactions from which those ancient microbes harnessed energy were fundamentally different from those that modern anaerobes living in the same environment use (Martin and Sousa, 2016). From the geochemical standpoint, there is no clear indication that early biological processes were fundamentally different from those still existing today in anaerobic environments (Arndt and Nisbet, 2012). Investigations of microbial genomes support the view that the first cells lived from H_2 and CO_2 and that acetogens (bacteria) and methanogens (archaea) are among the most primitive prokaryotic lineages currently known (Weiss *et al.*, 2016).

References

Arndt, N.T., Nisbet, E.G., 2012. Processes on the young Earth and the habitats of early life. Annual Review of Earth and Planetary Sciences 40, 521–549.

Bada, J.L., Lazcano, A., 2008. Prebiotic soup – Revisiting the Miller experiment. Science 300, 745–746.

Brack, A., 2009. Impacts and origins of life. Nature Geoscience 2, 8–9.

Branciamore, S., Gallori, E., Szathmary, E., Czárán, T., 2009. The origin of life: Chemical evolution of a metabolic system in a mineral honeycomb? Journal of Molecular Evolution 69, 458–469.

Copley, S.D., Smith, E., Morowitz, H.J., 2007. The origin of the RNA world: Co-evolution of genes and metabolism. Bioorganic Chemistry 35, 430–443.

Daniel, S.L., Hsu, T., Dean, S.I., Drake, H.L., 1990. Characterization of the H_2- and CO-dependent chemolithotrophic potentials of the acetogens *Clostridium thermoaceticum* and *Acetogenium kivuit*. Journal of Bacteriology 172, 4464–4471.

Eck, R.V., Dayhoff, M.O., 1966. Evolution of the structure of ferredoxin based on living relics of primitive amino acid sequences. Science 152, 363–366.

Eigen, M., 1971. Self-organization of matter and the evolution of biological macromolecules. Die Naturwissenschaften 58, 465–523.

Eigen, M., Schuster, P., 1977. The hypercycle. A principle of natural self-organization. Die Naturwissenschaften 64, 541–565.

Etiope, G., Schoell, M., Hosgörmez, H., 2011. Abiotic methane flux from the *Chimaera seep* and *Tekirova ophiolites* (Turkey): Understanding gas exhalation from low temperature serpentinization and implications for Mars. Earth and Planetary Science Letters 310, 96–104.

Fuchs, G., 2011. Alternative pathways of carbon dioxide fixation: Insights into the early evolution of life? Annual Review of Microbiology 65, 631–658.

Gantí, T., 1975. Organization of chemical reactions into dividing and metabolizing units: The Chemotons. BioSystems 7, 15–21.

Gilbert, W., 1986. The RNA world. Nature 319, 618.

Goldford, J.E., Hartman, H., Smith, T.F., Segré, D., 2017. Remnants of an ancient metabolism without phosphate. Cell 168, 1126–1134.

Hordijk, W., Steel, M., 2004. Detecting autocatalytic, self-sustaining sets in chemical reaction systems. Journal of Theoretical Biology 227, 451–461.

Hordijk, W., Steel, M., Kauffman, S., 2012. The structure of autocatalytic sets: Evolvability, enablement, and emergence. Acta Biotheoretica 60, 379–392.

Kaufmann, M., 2009. On the free energy that drove primordial anabolism. International Journal of Molecular Sciences 10, 1853–1871.

Koonin, E.V., Novozhilov, A.S., 2009. Origin and evolution of the genetic code: The universal enigma. IUBMB Life 61, 99–111.

Kuhn, H., 1972. Self-organization of molecular systems and evolution of the genetic apparatus. Angewandte Chemie International Edition 11, 798–820.

Levy, M., Miller, S.L., Oró, J., 1999. Production of guanine from NH_4CN polymerizations. Journal of Molecular Evolution 49, 165–168.

Lincoln, T.A., Joyce, G.F., 2009. Self-sustained replication of an RNA enzyme. Science 323, 1229–1232.

Martin, W.F., 2012. Hydrogen, metals, bifurcating electrons, and proton gradients: The early evolution of biological energy conservation. FEBS Letters 9, 485–493.

Martin, W., Russell, M.J., 2003. On the origins of cells: A hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. Philosophical Transactions of the Royal Society B: Biological Sciences 358, 59–85.

Martin, W.F., Sousa, F.L., 2016. Early microbial evolution: The age of anaerobes. Cold Spring Harbor Perspectives in Biology 8, a018127.

Mayumi, D., Dolfing, J., Sakata, S., *et al.*, 2013. Carbon dioxide concentration dictates alternative methanogenic pathways in oil reservoirs. Nature Communications 4, 1–6.

McCollom, T.M., 2013. Miller-Urey and beyond: What have we learned about prebiotic organic synthesis reactions in the past 60 years? Annual Review of Earth and Planetary Sciences 41, 207–229.

McCollom, T.M., 2016. Abiotic methane formation during experimental serpentinization of olivine. Proceedings of the National Academy of Sciences of the United States of America 113, 13965–13970.

McDermott, J.M., Seewald, J.S., German, C.R., Sylva, S.P., 2015. Pathways for abiotic organic synthesis at submarine hydrothermal fields. Proceedings of the National Academy of Sciences of the United States of America 112, 7668–7672.

Miller, H.M., Mayhew, L.E., Ellison, E.T., *et al.*, 2017. Low temperature hydrogen production during experimental hydration of partially serpentinized dunite. Geochimica et Cosmochimica Acta. doi:10.1016/j.gca.2017.04.022 (in press).

Miller, S., Orgel, L.E., 1974. The Origins of Life on Earth. Englewood Cliffs, NJ: Prentice-Hall.

Orgel, L.E., 2008. The implausibility of metabolic cycles on the prebiotic earth. PLOS Biology 6, e18.

Pedrera, A., Galindo-Zaldívar, J., Acosta-Vigil, A., *et al.*, 2016. Serpentinization-driven extension in the Ronda mantle slab (Betic Cordillera, S. Spain). Gondwana Research 37, 205–215.

Powner, M.W., Gerland, B., Sutherland, J.D., 2009. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. Nature 459, 239–242.

Ritson, D.J., Sutherland, J.D., 2013. Synthesis of aldehydic ribonucleotide and amino acid precursors by photoredox chemistry. Angewandte Chemie International Edition 52, 5845–5847.

Russell, M.J., Hall, A.J., Martin, W., 2010. Serpentinization as a source of energy at the origin of life. Geobiology 8, 355–371.

Russell, M.J., Nitschke, W., Branscomb, E., 2013. The inevitable journey to being. Philosophical Transactions of the Royal Society B: Biological Sciences 368. doi:10.1098/rstb.2012.0254.

Saladino, R., Crestini, C., Pino, S., Costanzo, G., Di Mauro, E., 2012. Formamide and the origin of life. Physics of Life Reviews 9, 84–104.
Schönheit, P., Buckel, W., Martin, W.F., 2016. On the origin of heterotrophy. Trends in Microbiology 24, 12–25.
Schrenk, M.O., Brazelton, W.J., Lang, S.Q., 2013. Serpentinization, carbon, and deep life. Reviews in Mineralogy and Geochemistry 75, 575–606.
Schuchmann, K., Müller, V., 2014. Autotrophy at the thermodynamic limit of life: A model for energy conservation in acetogenic bacteria. Nature Reviews Microbiology 12, 809–821.
Semenov, S.N., Kraft, L.J., Ainla, A., *et al.*, 2016. Autocatalytic, bistable, oscillatory networks of biologically relevant organic reactions. Nature 537, 656–660.
Smith, E., Morowitz, H.J., 2016. The Origin and Nature of Life on Earth: The Emergence of the Fourth Geosphere. Cambridge: Cambridge University Press.
Sousa, F.L., Martin, W.F., 2014. Biochemical fossils of the ancient transition from geoenergetics to bioenergetics in prokaryotic one carbon compound metabolism. Biochimica et Biophysica Acta (BBA) – Bioenergetics 1837, 964–981.
Spiegelman, S., 1965. The synthesis of a self-propagating and infectious nucleic acid with a purified enzyme. Proceedings of the National Academy of Sciences of the United States of America 54, 919–927.
Thauer, R.K., Jungermann, K., Decker, K., 1977. Energy conservation in chemotropic anaerobic bacteria. Bacteriological Reviews 41, 100–180.
Thauer, R.K., Kaster, A.-K., Seedorf, H., Buckel, W., Hedderich, R., 2008. Methanogenic archaea: Ecologically relevant differences in energy conservation. Nature Reviews Microbiology 6, 579–591.
Wächtershäuser, G., 1992. Groundworks for an evolutionary biochemistry: The iron-sulfur world. Progress in Biophysics and Molecular Biology 58, 85–201.
Weiss, M.C., Sousa, F.L., Mrnjavac, N., *et al.*, 2016. The physiology and habitat of the last universal common ancestor. Nature Microbiology 1, 1–8.
Westall, F., 2012. Early earth. In: Lunine, J., *et al.* (Eds.), Astrobiology. Cambridge: Cambridge University Press, pp. 89–114.
Wong, T.-F.J., Impey, C., Lunine, J., Funes, J., 1975. The co-evolution theory of the genetic code. Proceedings of the National Academy of Sciences of the United States of America 72, 1909–1912.
Yockey, H.P., 2005. Information Theory, Evolution, and the Origin of Life. Cambridge: Cambridge University Press.

Publication 2

Title: The future of origin of life research: bridging decades-old divisions

Year: 2020

Authors: Martina Preiner, Silke Asche, Sidney Becker, Holly C. Betts, Adrien Boniface, Eloi Camprubi, Kuhan Chandru, Valentina Erastova, Sriram G. Garg, Nozair Khawaja, Gladys Kostyrka, Rainer Machné, Giacomo Moggioli, Kamila B. Muchowska, Sinje Neukirchen, Benedikt Peter, Edith Pichlhöfer, Ádám Radványi, Daniele Rossetto, Annalena Salditt, Nicolas M. Schmelling, Filipa L. Sousa Fernando D. K. Tria, Dániel Vörös, and Joana C. Xavier

Published in: *Life*, 10(3), 20.
10.3390/life10030020.

Contribution: First and shared corresponding author.
Major: wrote first manuscript draft together with last author and edited heavily throughout the writing process. Illustrated and designed figures together with last author.

Summary: This perspective originated in a meeting (Interdisciplinary Origin of Life Meeting, 2018) organized by the first and last author. The goal is to provide a forward-looking view of OoL research with a focus on building bridges between unreconciled approaches and hypotheses.

life

MDPI

Perspective

The Future of Origin of Life Research: Bridging Decades-Old Divisions

Martina Preiner [1,*], Silke Asche [2], Sidney Becker [3], Holly C. Betts [4], Adrien Boniface [5], Eloi Camprubi [6], Kuhan Chandru [7,8], Valentina Erastova [9], Sriram G. Garg [1], Nozair Khawaja [10], Gladys Kostyrka [11], Rainer Machné [12,13], Giacomo Moggioli [14], Kamila B. Muchowska [15], Sinje Neukirchen [16], Benedikt Peter [17], Edith Pichlhöfer [16], Ádám Radványi [18,19], Daniele Rossetto [20], Annalena Salditt [21], Nicolas M. Schmelling [12,22], Filipa L. Sousa [16], Fernando D. K. Tria [1], Dániel Vörös [18,19] and Joana C. Xavier [1,*]

1 Institute of Molecular Evolution, University of Düsseldorf, 40225 Düsseldorf, Germany; gargs@hhu.de (S.G.G.); tria@hhu.de (F.D.K.T.)
2 School of Chemistry, University of Glasgow, Glasgow G128QQ, UK; 2304694A@student.gla.ac.uk
3 Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK; sb2391@cam.ac.uk
4 School of Earth Sciences, University of Bristol, Bristol BS8 1RL, UK; hb1864@bristol.ac.uk
5 Environmental Microbial Genomics, Laboratoire Ampère, Ecole Centrale de Lyon, Université de Lyon, 69130 Ecully, France; adrien.boniface@univ-lyon1.fr
6 Origins Center, Department of Earth Sciences, Utrecht University, 3584 CB Utrecht, The Netherlands; e.camprubicasas@uu.nl
7 Space Science Center (ANGKASA), Institute of Climate Change, Level 3, Research Complex, National University of Malaysia, 43600, UKM Bangi, Selangor, Malaysia; kuhan@ukm.edu.my
8 Department of Physical Chemistry, University of Chemistry and Technology, Prague, Technicka 5, 16628 Prague 6–Dejvice, Czech Republic
9 UK Centre for Astrobiology, School of Chemistry, University of Edinburgh, Edinburgh EH9 3FJ, UK; valentina.erastova@ed.ac.uk
10 Institut für Geologische Wissenschaften, Freie Universität Berlin, 12249 Berlin, Germany; nozair.khawaja@fu-berlin.de
11 Lycée Colbert, BP 50620 59208 Tourcoing Cedex, France; gladys.kostyrka@wanadoo.fr
12 Institute of Synthetic Microbiology, University of Düsseldorf, 40225 Düsseldorf, Germany; machne@hhu.de (R.M.); nicolas.schmelling@hhu.de (N.M.S.)
13 Quantitative and Theoretical Biology, University of Düsseldorf, 40225 Düsseldorf, Germany
14 School of Biological and Chemical Sciences, Queen Mary University of London, London E1 4DQ, UK; giacomo.moggioli@qmul.ac.uk
15 Université de Strasbourg, CNRS, ISIS, 8 allée Gaspard Monge, 67000 Strasbourg, France; muchowska@unistra.fr
16 Archaea Biology and Ecogenomics Division, University of Vienna, 1090 Vienna, Austria; sinje.neukirchen@univie.ac.at (S.N.); a00807114@unet.univie.ac.at (E.P.); filipa.sousa@univie.ac.at (F.L.S.)
17 Cellular and Molecular Biophysics, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany; peter@biochem.mpg.de
18 Department of Plant Systematics, Ecology and Theoretical Biology, Eötvös Loránd University, Pázmány Péter sétány 1/C, 1117 Budapest, Hungary; adamradvanyi117@gmail.com (A.R.); mrdanielred@gmail.com (D.V.)
19 Institute of Evolution, MTA Centre for Ecological Research, Klebelsberg Kuno u. 3., H-8237 Tihany, Hungary
20 Department of Cellular, Computational and Integrative Biology (CIBIO), University of Trento, 38123 Trento, Italy; daniele.rossetto@unitn.it
21 Systems Biophysics, Physics Department, Ludwig-Maximilians-Universität München, 80799 Munich, Germany; a.salditt@physik.uni-muenchen.de
22 Cluster of Excellence on Plant Sciences (CEPLAS), University of Cologne, 50674 Cologne, Germany
*Correspondence: xavier@hhu.de (J.C.X.); preiner@hhu.de (M.P.)

Received: 31 January 2020; Accepted: 21 February 2020; Published: 26 February 2020

Abstract: Research on the origin of life is highly heterogeneous. After a peculiar historical development, it still includes strongly opposed views which potentially hinder progress. In the 1st Interdisciplinary Origin of Life Meeting, early-career researchers gathered to explore the commonalities between theories and approaches, critical divergence points, and expectations for the future. We find that even though classical approaches and theories—e.g. bottom-up and top-down, RNA world vs. metabolism-first—have been prevalent in origin of life research, they are ceasing to be mutually exclusive and they can and should feed integrating approaches. Here we focus on pressing questions and recent developments that bridge the classical disciplines and approaches, and highlight expectations for future endeavours in origin of life research.

Keywords: origins of life; prebiotic chemistry; early life; LUCA; abiogenesis; top-down; bottom-up; emergence

1. Introduction

Understanding the origin of life (OoL) is one of the major unsolved scientific problems of the century. It starts with the lack of a commonly accepted definition of the phenomenon of life itself [1], but difficulties go far beyond merely that obstacle. OoL research involves a large number of diffuse concepts cornering several natural sciences and philosophy, such as entropy, information and complexity. Despite evidence that untangling this knot will require a concerted and collaborative effort between different disciplines, technologies, individuals and groups [2], division in OoL research is still marked, concerning both theories (e.g. RNA world vs. metabolism-first) and approaches (e.g. bottom-up vs. top-down). What causes these on-going divisions, and how can heated debates be moderated?

There is some consensus on a few points. First, the earliest undisputed fossil evidence places life on Earth prior to 3.35 Ga [3] and molecular clocks suggest an origin prior to the late heavy bombardment >3.9 Ga [4]. Second, the origin of life must have resulted from a long process or a series of processes, not a sudden event, for the complexity of a cell could not have appeared instantaneously. The OoL must have started from simple abiotic processes, involving one or more sources of energy and matter (particularly CHNOPS: carbon, hydrogen, nitrogen, oxygen, phosphorus and sulfur, the six most prevalent elements in life on Earth) forming protometabolism, compartmentalization and inheritance. But strikingly, the list of agreements does not expand much further than this. Several researchers have speculated on life forms different than the ones we know and see on Earth—not based on cells [5] or even on CHNOPS [6,7]. Even if evidence for those alternative life forms has yet to be found, here we do not exclude their possibility. Throughout this text we will focus on life as we know it, which insofar as has been demonstrated is all based on individual cells with metabolism, genetic inheritance and compartmentalization.

The list of individual theories, different lines of experimental and theoretical research and diverse views on the OoL is extensive and eclectic. It is not the purpose of this article to be a comprehensive review of all of those. Rather, based on discussions that took part in the 1st Interdisciplinary Origin of Life meeting for early-career researchers (IOoL), we present a forward-looking perspective on how discontinued discourses on the OoL can be (re)united in a new mosaic with resolution and meaning. We reflect purposely on individual topics causing the most distressing divisions in OoL research, most of which result from classical separations between disciplines and theories that date to decades ago. We then portray examples of bridges being built between classically opposed views and finish by providing a roadmap for future dialogue and evidence-based research in OoL.

2. Classical Divisions in Origin of Life (OoL) Research

2.1. Top-Down versus Bottom-Up: Where To?

Classical OoL research has gained from synthetic or bottom-up approaches, aiming at synthesizing de novo the building blocks of life (and assemblies of them), and top-down or analytic approaches, which start from living systems, deconstruct them into their parts and survey their properties. But while most origins questions (e.g. that of the universe, atoms, continents or social structures) are usually led by 'their' disciplines with small contributions from others, the OoL has been peculiar in calling multiple subjects to play major roles.

Chemistry focused on the heavy lifting at the bottom-up, synthetic side of OoL research. Chemical reactions potentially relevant to life's beginnings date back as far as 1828 with Wöhler's urea synthesis [8], a landmark for the birth of organic chemistry. The year of 1861 brought Butlerow's formose reaction [9], which generates a mixture of carbohydrates from formaldehyde through an autocatalytic mechanism that came to be appreciated by the OoL community nearly a century later [10]. The launch of the chemical origins field occurred with Miller's experiment in 1953, based on Oparin and Haldane's theoretical work (see "Section 3. Building Bridges"), synthesizing a mixture of organic compounds (including essential amino acids) from water, methane, ammonia and hydrogen using electric discharges [11,12]. Later, significant advances were made with the synthesis of phospholipid membranes under prebiotic conditions [13]. The synthetic quest continued with efforts focusing on the synthesis of amino acids (usually involving the Strecker or the Bücherer–Bergs synthesis), sugars and nucleobases (usually based on the formose reaction and the chemistry of nitriles and their derivatives); for thorough recent reviews see [14,15] and references therein. Progress is scarcer in understanding the polymerization of proto-biological monomers into oligomers. Recent achievements include RNA synthesis assisted by ribozymes [16,17] and lipids [18]; polymerization of carbohydrate monomers [19] and peptide-bond formation at the water-air interface [20], under dehydration-hydration conditions [21] and in mineral interlayers [22]. Synthetic biology worked on the assembly of more complex features, as the synthesis of a whole bacterial genome [23] and of a viral cDNA genome that was transcribed, translated and replicated without cells [24].

Given that virtually all of the origins hypotheses are non-falsifiable, the diversity of the bottom-up approaches mentioned here has led to an important question: to what extent do the experimental synthetic advances resemble what actually happened at the OoL? With the complete set of inorganics and ready-synthesized organics at its disposal (plus a wide range of experimental conditions), the bottom-up approach parallels an artistic endeavour that can paint both abstract and realistic pictures of the first biomolecules and their assemblies. For this reason, to constrain the experimental space, clearer pictures of i) the fundamental features of life, ii) ancestral life forms and iii) the environmental conditions at their origin are urgently required.

Surprisingly, biology has played a modest part in the search for the OoL [14,25]. A closer look reveals possible reasons for this (quantitatively) sparse involvement. Biology is complex and its numbers are vast: the queue of living species is long, and only recently did technology to start exploring their workings within new disciplines (e.g. biochemistry, molecular biology and genetics) become available. Looking at microbes only, recent estimates point to 1 trillion different species on Earth [26] and orders of magnitude more bacterial cells on the planet than the number of stars in the known universe. A single one of those cells can have at any given moment 2–4 million expressed proteins [27]. By looking at whole living systems, biology has much in its hands, and a gap still remains between i) experimental biology looking at specific mechanistic aspects in a cell or organism (e.g. metabolism, particular enzymes or pathways) and ii) the larger (yet microscopic) picture of how life as a whole started, which falls out of the traditional interpretation of the theory of evolution [28]. Were there mechanisms that drove variation and selection before genomes?

One could say that the top-down approach kick-started when a small number of biological species were compared, the mechanism of speciation was uncovered and the theory of evolution formulated, in parallel with the discarding of spontaneous generation [29]. Yet, the bona fide dawn of top-down approaches in the OoL question would not come to be until the genomics era. Modelling

early evolution and the OoL required a precise and holistic way to trace species back in time, and that came first in the form of genome sequences. Prokaryotes, the simplest forms of cellular life, are increasingly supported by evolutionary studies as the oldest lifeforms, and thus of utmost importance for OoL research [30]. The first comparison of prokaryotic genomes revealed a conserved set of 240 genes [31]. Later, the exponential growth of the number of sequenced genomes came with a daunting realization: the prokaryote world is highly diverse, full of redundancy with non-orthologous gene displacements (unrelated sequences encoding the same biochemical function) and lateral gene transfer [32], and the intersection of genomes shrank to a mere handful of ribosomal genes [33]. The search for the genetic content of LUCA—the Last Universal Common Ancestor—using the classical comparative top-down approach had thus stagnated. Some argue that the gap between the OoL and the biochemical wiring of LUCA is so vast that whatever constituted the chemistry leading to the earliest life could have been rewritten multiple times, and changed beyond recognition to then become LUCA [34]. Others argue that even if a good reconstruction of LUCA cannot determine the chemistry at the OoL, universal biochemical features can meet with OoL chemistry by advancing experimental and analytical methods [35]. Parallel but relevant to this line of thought, a new biological concept has been devised—the IDA (initial Darwinian ancestor) [36] also called FUCA (first universal common ancestor) [37], the first entity that could be considered capable of evolution.

The bridging of all the critical advances made by chemistry's bottom-up and biology's top-down approaches seems to be an indelible requirement for the advancement of the OoL field [38–40]. It is not impossible to envision these seemingly non-overlapping approaches to be reconciled by incorporating expertise from other fields. Indeed, disciplines beyond classical biology and chemistry have much to contribute. Before exploring those contributions, we elaborate further on disputes in OoL research by briefly focusing on another particular format of OoL theories: the prebiotic 'worlds'.

2.2. One Origin, Abundant Worlds

The relative abundance and distribution of building blocks of the main classes of biomolecules ~4 Ga ago on Earth is still elusive and a matter of strong debate [15]. Based on earlier suggestions [41], RNA emerged among all biomolecules in a privileged metaphor that would cement and permeate OoL research for decades to come—the 'RNA world' [42]. Ribozymes constituted a shocking and promising discovery: there was now genetic material that was catalytic as well [43,44]. Many found the metaphor appealing: a world with a jack-of-all-trades RNA molecule, catalyzing the formation of indispensable cellular scaffolds, from which somehow then cells emerged [45,46]. Others were quick to notice several difficulties with that scenario. These included the lack of templates enabling the polymerization of RNA in the prebiotic complex mixture [47] and RNA's extreme lability at moderate to high temperatures and susceptibility to base-catalyzed hydrolysis [48]. The 'metabolism-first' theory emerged independently [49–52], standing in favour of simpler molecular networks harnessing energy from geological disequilibrium leading to the emergence of genetic complexity. The RNA world and the metabolism-first theories have often been portrayed in stark opposition to each other [53,54], leading the OoL field to an unprecedented division. Ever since, other prebiotic worlds came to light with their own preferred class of biomolecules and significant insights, e.g. the protein (or peptide) [39,55,56], lipid [57,58], coenzyme [59], and even virus [60] worlds are some of the most popular theories for the order and/or relevance of appearance of biomolecules on Earth.

Each prebiotic world generated invaluable insight on its own class of biomolecules, but also proposes a privileged, precursory function for them. This privileged position imposes a sequential, stepwise hierarchy on the OoL timeline: subsequent takeovers by the next world(s) until a living cell originates. Still, the question of which class of biomolecules initiated the OoL is a loaded question [61]. All known living cells contain DNA, RNA, proteins, lipids, coenzymes, and other metabolites—and the earliest cells as those known on Earth would have had to fulfil these minimal cell requirements [62]. There is a strong argument to be made for the emergence of essential biomolecules to have been (at least to some extent) contemporaneous and interdependent. More importantly, the

origin of biomolecules needs to be distinguished from the origin of cells, and life. Cells are not mere collections of their chemical components, but highly dynamic, complex systems with multiple interlocked processes involving those components. For this reason, the emergence of life cannot be distilled down to biomolecular retrosynthesis only. This appeal is not new, Kamminga asked biochemists to focus on processes rather than on pure retrosynthesis more than 30 years ago [54]. If common requirements between the origin of each biomolecule are found, they may provide necessary constraints for the OoL problem. Ultimately, for the emergence of biomolecules to be relevant for the OoL, it should be tightly linked with their participation in dynamic processes characteristic of life (e.g. replication, energy coupling and compartmentalization). These are some of the reasons why bridging different prebiotic worlds, with all their conceptual advances, is an urgent endeavour.

3. Building Bridges

3.1. Pressing Questions in OoL are Interdisciplinary

Insights from classical approaches, hypotheses and 'worlds' have led to many advances, but they have also resulted in an ideological isolation that is possibly hindering progress in the OoL field [35,53,63–67]. A need to link different disciplines and approaches becomes evident looking at, in our view, the most central questions that should be addressed (Figure 1). These and other questions were identified before [68]. Only cooperation can push their answers forward—in both bottom-up and top-down directions. This is truly not an innovative line of thought. Aleksandr Oparin formulated it in 1924 in a comprehensive way:

"A whole army of biologists is studying the structure and organization of living matter, while a no less number of physicists and chemists are daily revealing to us new properties of dead things. Like two parties of workers boring from the two opposite ends of a tunnel, they are working towards the same goal. The work has already gone a long way and very, very soon the last barriers between the living and the dead will crumble under the attack of patient and powerful scientific thought." [69]

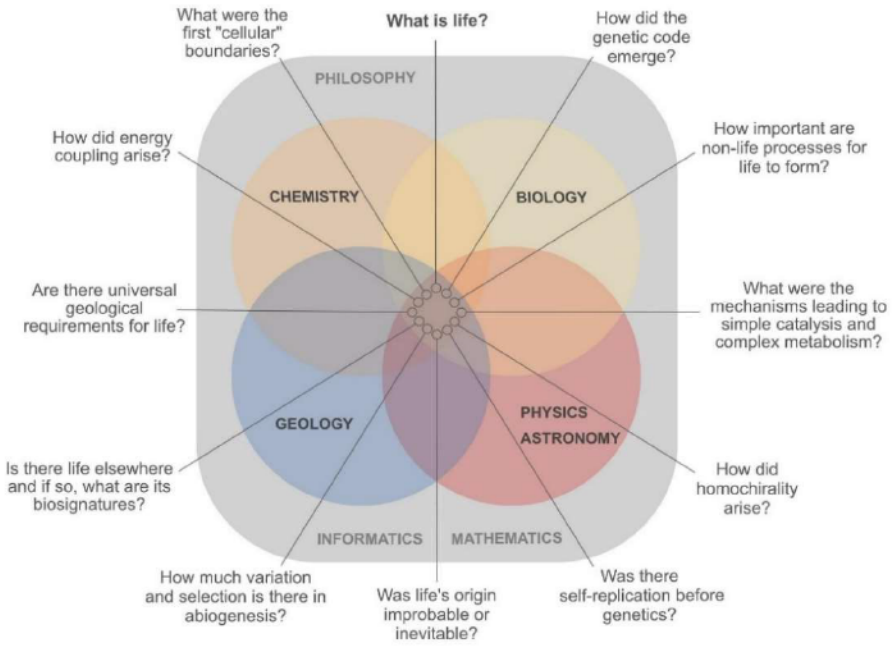


Figure 1. Current central questions in origin of life (OoL) research. In order to find answers, each of the natural sciences illustrated has to play a role by converting these questions into hypotheses and theories, while constantly testing them experimentally. The fundamental roles of philosophy, mathematics and informatics are portrayed in the background.

Both Oparin and Haldane proposed independently integrative theories for the OoL in the 1920s [70]. Their 'Heterotrophic origin of life theory' was one of the first approaches to describe a slow transition between a non-living primordial 'soup' (Haldane) and living cells. Oparin was the first to integrate data from geochemistry, chemistry, planetary sciences and biochemistry in his seminal work proposing different prebiotic and biological stages [54,69]. His words, quoted above, now almost 100 years old, were optimistic. The "very, very soon" has not happened to this day—there is still a pressing need to bridge (or, in Oparin's words "tunnel") top-down and bottom-up approaches. Every attempt to reach over to "the opposite end of the tunnel" opens ways to new data and/or new data interpretations. Cross-disciplinary collaborations can lead to innovative discoveries when new information from discipline A challenges the status quo in discipline B [71]. This forces researchers to formulate new approaches and leads to progress. 'Outsider information' like that may help to find a middle ground for irreconcilable hypotheses or theories (see also "Section 4. Towards the Future") [72].

The OoL question benefits greatly from crossovers of scientific disciplines, each of which brings its specific skills to the table. Chemistry supplies the knowledge of the construction kit, and physics, for instance, the energetic boundaries for the assembly of life [73,74]. Biology is confronted with bigger pictures, providing the roadmap from extant life to its origins via data collection and analysis (from multiple 'omics' technologies [75] to the study of elaborate microbial communities [76]). Geology and other geosciences help unfold the possible timeline and set the environmental constraints for the OoL (terrestrial and/or extraterrestrial) [77,78]. All of the above disciplines rely heavily on engineering, mathematics and informatics. Finally, philosophy provides a structure that allows scientists to ask the right questions following the rules of reasoning and argumentation, avoiding the traps of non-sequiturs and other logical fallacies [2,49].

An important first step in connecting disciplines is making the borders between them fully permeable. In the past 50 years, several interdisciplinary fields developed and paved the way for new research. For OoL in particular, astrochemistry, astrobiology and their search for biosignatures in the universe had a significant impact. These disciplines focused on physical and chemical processes that could lead to biotic or prebiotic systems on other planets [79] (Figure 1). One example is the discovery of subsurface oceans on icy moons, which extended the concept of habitable zone from the inner to the darker, colder places of the outer solar system [80]. Both complex, large insoluble organic material and low mass, reactive and soluble molecules, have been found in the subsurface ocean of Enceladus [81,82]. Large molecules have also been found in the atmosphere of Titan [83].

How does the need for discipline crossovers and permeability translate into practice? To illustrate this, we will use three central questions—all concerning essential cellular processes—as examples: (i) how did energy coupling arise, (ii) what mechanisms led to metabolism, and (iii) how did the genetic code emerge?

Life on Earth couples energy-releasing (spontaneous) reactions to energy-demanding (non-spontaneous) ones, capturing energy from its environment and eventually dissipating it as heat. This enables cellular processes such as growth and division. But how did this sophisticated system develop? Today, energy-coupling is mediated by enzymes which, acting as engines, funnel energy released from the cell's diet into chemical energy. This energy is stored in a thioester linkage (as in acetyl-CoA), a phosphate-ester bond to carbon like in acetyl phosphate or a phosphate bond in the adenosine triphosphate (ATP) molecule [84]. These molecules are commonly known as energetic currencies in cells and mediate energy coupling by transferring energy between non-related biochemical processes. Are energetic currencies indispensable or exchangeable within existing metabolisms? To answer this biological insight on different species and pathways is crucial [85]. In order to find out if energy coupling was necessary for life's first steps (still a subject of debate) several alternative energy currencies have been proposed and tested, for instance high-energy phosphates [86] or thioesters [87]—as it is known that cells use both, often sequentially (e.g. the production of ATP is sometimes mediated by thioesters). From there, further questions emerge: how did these energy-storing molecules become coupled and implemented into life as we know it (or: what was the order of entry of such heteroatoms as nitrogen, sulfur and phosphorus into proto-life's chemistry)? Here, it would be helpful to look at different geochemical conditions to explore under which

conditions these couplings could be possible. Although the answers are still unclear [88], discovering them will be pivotal to understanding the transition between non-living and living systems [84,89].

The question of energy coupling is directly linked to the question of what geochemical mechanisms led to complex cellular metabolism. A systematic search for metabolism's beginnings (i.e. protometabolism) requires a clear and agreed-upon definition of metabolism, yet to be developed. This definition should start from the molecular level and the fundamental physical driving forces (kinetics vs. thermodynamics) [90], leading to combinations of chemical reactions, ultimately giving rise to a complex network [91,92], governed by a set of rules and boundaries [93,94]. Starting from known biochemical pathways one can deduce a conserved network that could sustain itself without enzymes [95–97]. The feasibility of such top-down analysis reducing the complexity of biological metabolic networks has to be tested in a bottom-up laboratory setting [98–100]. Conditions for these laboratory experiments can either be inspired by known geological environments, or a set of favourable conditions can be searched among existing geological settings and/or models. Such comparisons and convergences between biochemical and geochemical conditions, reactions and products can help to unveil environments where metabolic pathways and networks could emerge [101–103]. This applies not only to the early Earth but also other locations elsewhere in the universe like Enceladus [81,104]. Certainly, the same methodological rules will apply to other angles from which the OoL question is tackled: researching compartmentalisation, reconstructing the characteristics of LUCA, studying molecular replicators and precursors to genetics —hardly can we envision any of these succeeding in isolation.

Along with metabolism, life is based on another equally-important fundamental principle-inheritance, also described as "information that copies itself" [105]. As discussed (see "Section 2.2. One Origin, Abundant Worlds"), a better question than "which came first, information or metabolism?" would be: does the emergence of one have common properties with the other? In other words, perhaps we will know that acetate originated before ATP, but the emergence of metabolism and inherited information as complex systems was most likely interdependent and simultaneous. This directly leads to the fundamental question of how and when metabolism and information storage became linked. Nature's elegant solution is the genetic code, the origin of which remains a true enigma.

Different hypotheses exist for the origin of the genetic code, often depending on assumptions of the different hotly debated prebiotic conditions. Ribozymes show that RNA can harbour both catalysis and genetics, and there is support for the capability of RNAs to aminoacylate [106]. However, the code today is self-referential, that is, the mapping between amino acids and codons heavily relies on encoded proteins. When, why and how did non-coded peptides become involved? For the origin of the involvement of peptides in the genetic code, the conjecture that they are a superior catalyst is not enough, especially because they only have a supporting but not catalytic role in the ribosome [107]. Therefore, we must justify likely proximal selective advantages of using amino acids and simple peptides in the context of RNAs. These so-called 'exaptations' may bridge the RNA and peptide worlds, offering the basis on which later elements of the translational apparatus were built. Possible advantages of peptides include providing catalytic aid and expanding the catalytic repertoire of RNAs [107], membrane transport [108], scaffolding [109] and energy storage [110]. It has indeed been shown experimentally that non-coded peptides can potentiate the functions of RNA, which supports the coevolution of RNA and peptides [111]. Both the RNA and protein worlds ask whether a bipartite polymer setup (with nucleic acids and proteins) and the current genetic code is an inherent requirement of life, or if it is possible to envision one dominant polymer carrying information in prebiotic stages. Neither world alone can provide a clear explanation for the interlacing of the two, though, and there are relatively few endeavours to try out 'messy emergence', where some initial cooperation between non-coded proteins and RNAs was vital. However, regardless of the polymers at the origin of the code, another question urges: how did the code freeze to the current codon table? This question seemed for some time to be a mere statistical or even cryptographic problem [112], and a variety of explanations emerged to solve it, most popularly: stereochemical basis for the assignment between nucleic acids and amino acids [113], the

development of the code guided by the biosynthetic pathways of amino acids [114], and optimization in order to reduce the severity of mutations [115]. These are not mutually exclusive hypotheses, and the origin of the code might have simultaneously involved several of them [116]. Most likely, the development of the genetic code took place in a continuous expansion [117], a hypothesis supported by functional proteins with reduced amino acid repertoires [118]. The later expansion of the repertoire was possibly governed by a physicochemical optimization that fits well with water-based biochemistry [119].

Answers on the origin of the code still seem very distant [115,116,120–122]. In order to finally solve this puzzle it will be necessary to investigate not only the emergence of biological information [120], but also the evolution of interactions between the molecules involved in translation [117,123].

3.2. *On the Right Track? Looking at the Past Decade*

In the past 10 years, many have worked and asked for OoL research to unite [25,39,124,125]. Here we look at experimental examples from the last decade that connect different disciplines, theories or interpretations (Figure 2).

One of the first barriers to come down stood between the decades-old views of different single-biomolecule worlds (see "Section 2.2. One Origin, Abundant Worlds"). Studies merging the lipid world with others are pioneers. The requirement for compartmentalization to keep genomic molecules and their products spatially together, as well as to allow for vectorial (bio)chemistry, suggests that the potential of lipids and other amphiphilic structures to self-assemble into micelles and bilayer vesicles within an aqueous phase constituted a critical step in the emergence of life [57,126,127]. Vesicles were thought to be stable only in salt-poor aqueous environments, such as surficial freshwater ponds or hydrothermal springs, but recent work showed that these become much more resilient to extreme salinity and pH if they are composed of mixtures of amphiphiles [128], a feature which better reflects the naturally messy aspect of prebiotic chemistry. Such prebiotic compartments, also referred to as 'protocells', are defined as primitive precursors of modern cells which, although not yet alive, exhibited essential cellular characteristics [129]. Efforts of OoL researchers aim to establish various in vitro protocell models that mimic key features of life as we know it, including simple metabolic pathways, replication or vesicle growth and division [130]. A protometabolism leading to sugar synthesis has been assembled within lipid vesicles, with the final products diffusing through the lipid barrier and being detected by living bacteria [131]. Not long after, DNA amplification was shown to induce growth and division of lipid vesicles, linking the reproduction of an informational substance with that of a compartment [132]. These demonstrate that simpler systems than cells can model fundamental interactions between membranes and their contents. Another example is the bridge between the metabolism-first and the RNA world theories [114]. It is now clear that the building blocks for RNA and DNA are intermediates of metabolic networks; they are never directly uptaken from the environment in their ready-for-polymerization forms (i.e. as phosphorylated nucleosides), but as unphosphorylated biogenic nucleosides [133], and are also in fact moieties of several essential cofactors [134]. The idea of the simultaneous and interdependent origins for RNA and DNA's building blocks has been given experimental evidence [135]. But bridges between the RNA world and metabolism-first can be built beyond their typical molecules. Classical approaches in OoL have often been constrained to biomolecules due to their ubiquity in biology, however, most of these biomolecules were not necessarily available at early prebiotic stages [136]. Prebiotic environments most likely included compounds not central to modern biopolymers, for example, alpha hydroxy acids (aHA) [137] among numerous others [138,139]. Recent work has shown that aHAs easily form combinatorial polyesters under wet-dry conditions that may have played a role in the catalytic landscape within which they were formed [140]. These polyesters can form membraneless compartments mimicking a cell, capable of differentially segregating various kinds of dyes, hosting a protein and even accumulating lipids in their exterior [141]. If one considers that the accumulation of biomolecules is a biological invention, optimizing for certain geochemical and/or biochemical properties [119], then compounds such as aHAs may have played a role in some form of nascent biology not conserved in modern biology [142]. Considering the

'messy' nature (meaning the inherent diversity) of prebiotic chemistry [143,144], this may have been the case, and the utility of central biomolecules such as RNA and ubiquitous metabolites to elucidate the OoL may be partial, despite their omnipresence in modern life.

Other striking examples for integration in OoL research come from geological studies providing a better picture of the Hadean world, including the likely atmospheric composition and the depths of oceans and tectonic activity [145,146]. These studies constrained and brought closer the work of both biologists and chemists. Experimentation with origins in hydrothermal vents [147–149] and geothermal fields [150–152] have taken into consideration geological insights in their OoL scenarios. In particular, the importance of relevant metals and metal clusters in the Hadean has settled in experimental work that recreates the origin of biochemistry in vitro, including carbon cycling [98,103,153], nitrogen fixation [154], ribosomal translation [155], and even the generation of pH gradients [156]. Simulated hydrothermal conditions, in particular pores, select for the replication of longer oligonucleotides [157]. Mineral surfaces have shown promising features in promoting biochemistry, including the selection of longer RNA molecules [100], and a variety of organic reactions including nitrogen reduction, lipid self-organization, condensation-polymerization reactions, selection and concentration of amino acids and sugars and chiral selection (see [158] and references therein). Surfaces can also help to tame the combinatorial space in a complex system of organic molecules [100,159].

By providing new ways to handle vast amounts of data in a systematic and quantitative manner, advances in computer sciences and technologies help the much-required aforementioned integrations that are starting to reflect on OoL research. Larger chemical libraries can now be monitored over a manageable timeline, and the diversification of self-replicating molecules has been observed in such systems [160]. The self-replication of small organic molecules has also been observed in an autocatalytic process displaying complex, non-linear responses to changes in environmental conditions [161]. Computational models allow us to test hypotheses unattainable in a laboratory alone, be it chemical conditions [22], timescales [4] or other routine perturbations to the model. Computational biology has now moved to mathematical simulations at increasing levels of complexity, from topological, interaction-based to constraint- and mechanism-based [162]. Recently, a database aggregating a variety of bioinformatic approaches to LUCA was compiled, allowing for testing hypotheses in investigations of ancient biochemistry [163]. Simulations can now be undertaken with geology and biology in mind, both in molecular dynamics [22] and network biology [95,96], identifying constraints and parameters that can provide direction for experimental work [98,159,164].

The progress of top-down approaches from traditional comparative genomics to more integrative approaches revives optimism in the quest for LUCA [165]. Integrating fossil data with molecular evolutionary clocks has solidified evidence for the age of LUCA at >3.9 Ga [4]. Taking into account morphological characteristics together with genetic signatures is another promising direction. Recent analyses of membranes and cell walls of prokaryotes suggest that LUCA was able to sporulate, that is, to reproduce into a dormant, non-metabolising cell that could survive long periods of harsh conditions such as the late heavy bombardment [146,166]. Modern phylogenetic considerations minimize the effect of lateral gene transfer in reconstructing the first genomes, pointing to a thermophilic autotrophic LUCA [149].

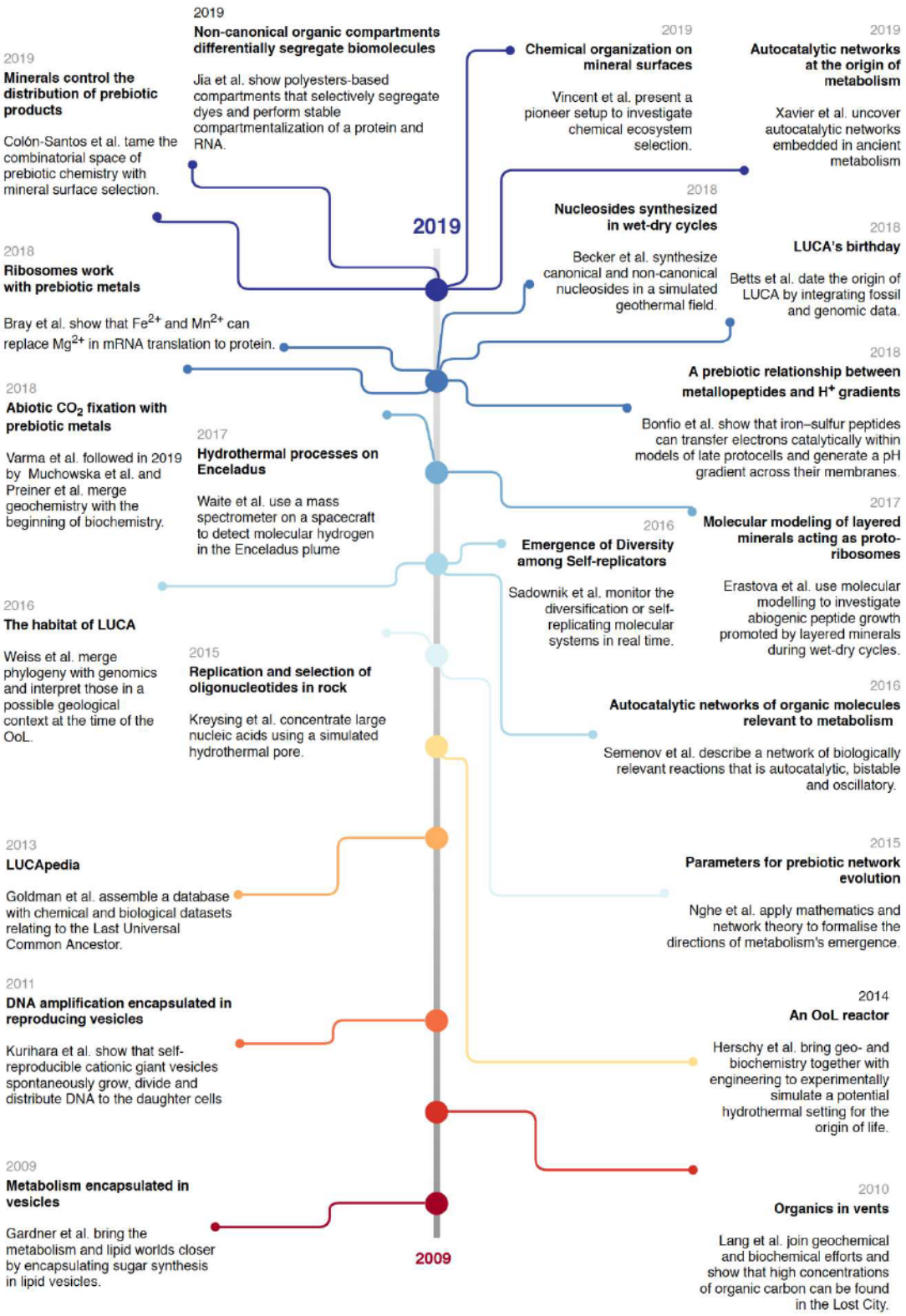


Figure 2. Timeline of recent multidisciplinary achievements that build bridges in OoL research. Included are examples from the past 10 years of OoL research that bridge disciplines, approaches and/or methods, biomolecules/single-world scenarios, simulations and experiments. The choice of studies does not aim to cover (exclusively) novel findings, but those that build bridges.

4. Towards the Future

4.1. General Remarks

We expect by now to have presented a convincing argument for the necessity of bridges in answering OoL's most prominent questions (Figure 1), together with examples of successful studies (Figure 2). How can the future hold more interconnections between approaches and theories? Some tools are general for all types of linking between seemingly disconnected research (Box 1). These include interdisciplinary conferences, where the focus lies not only on one molecule or one theory, but either on the whole field or on the question (Figure 1) sought to be answered. Interdisciplinary higher education and scientific collaborations should progress in the same direction in the coming decades, with more focus on questions and holistic pictures.

Box 1: Tools for future integration of OoL research

General → **Specific**

• Interdisciplinary conferences • Consistent nomenclature • Sharing experimental designs • Public communication • Collaborations • Interdisciplinary higher education • Debate/communication training • Incentives for early-career researchers • Careful prevention of conflicts of interest	Experiments *in vitro, in vivo, in silico*	• User-friendly informatic tools • Better databases • Promote reproducibility • Create interface spaces for easier data exchange (between computational and experimental) • Writing with broad audience in mind
	Theories	• Explore both commonalities and differences • Explore particular parts of an alternative theory, specially the testable ones • Promote either double-blind or double-open peer-review of articles and grants, to reduce personal and institutional bias

Regarding the limits of experimentation, computational tools can extrapolate from experimental input, and thus will play a major role in untangling complexity in OoL. However, research in vitro, in vivo or in silico (Box 1) needs to be attainable to everyone. This requires more user-friendly computational tools, the communication of tools and results in an accessible manner, and the creation of curated and moderated spaces where data and tools can be exchanged, faster than by the traditional scientific paper (as forums, curated databases, websites and discussion channels). Standards that can be used by different disciplines are heavily required, particularly in chemical, biochemical and genetic nomenclature—we recommend Chebi [167], KEGG [168] or BiGG [169]. Nomenclature in theoretical biology also requires attention. In particular, nomenclature regarding the first biological entities includes a great amount of confusing variability, e.g. an initial Darwinian ancestor (IDA) [36] akin to a first universal common ancestor, FUCA [37], and the more commonly used LUCA, akin to the urancestor [170], to the cenancestor [171], and to the (simpler) universal common ancestor [172]. Here, we recommend, merely based on the greater usage to this day, the use of IDA and LUCA. Semantic variability without clear conceptual justification does not ease literature review and theoretical studies, and a strong community effort is required to make different names clearly distinguishable or for standards to be adopted.

Sharing and combining experimental setups (both computational and laboratorial), can help to increase reproducibility but also to connect disciplines and approaches. For example, to narrow the current gap between prebiotic (geo)chemistry and ancient biochemistry, simulation efforts could focus on the least contested facts on the early Earth. These include a CO_2-rich atmosphere and hydrosphere from volcanism, and an H_2-rich hydrothermal outflow from the crust. Other parameters which are harder to characterize, as pH ranges, pressure, access to ultraviolet (UV) radiation, water activity, etc. could then be varied. There are means to attempt such simulations in the laboratory today, where collecting vast amounts of data is possible [173]. This data can be analysed and incorporated into the computational models which then, in turn, can shed more light on what laboratory experiments should be attempted. Hereby, the steady bilateral exchange between

disciplines is crucial. Not only should geology provide conditions for experiments [174], those experiments should also provide the attributes that geochemical studies could search for.

4.2. Commonalities between Opposing Theories

Bridges between methodologies will allow the OoL field to advance, but ultimately theories will have to become more interconnected as well. We foresee that this can happen with an intentional focus on commonalities rather than solely on differences, and plenty of examples can be mentioned. Here we choose to focus on four of the most heated debates, heavily interrelated: (i) the geological setting for the OoL, (ii) the source of food/starting molecules, (iii) the source of energy and (iv) the RNA world vs. metabolism-first division.

4.2.1. The Geological Setting

The proposed and heavily discussed locations for the OoL, including terrestrial hot springs [150], submarine hydrothermal vents [175,176], volcanic landscapes [177], hydrothermally-percolated sediments [178] and warm little ponds [179] create some of the strongest disagreements in OoL research, but have in common several requirements. These include a constant energy and CHNOPS source, heterogeneous or homogeneous catalysts, non-equilibrium conditions and conditions that enable aggregation of molecules to form life's building blocks and, eventually, polymerization of said building blocks [22,103,152,180]. One example is the requirement for wet–dry cycles for polymerization, mostly discussed in the context of geothermal fields [180], even though wet-dry cycles can also be found in hydrothermal pores [181] and even at the water-air interface [20]. On another note, some propose that different geological settings could have created different biomolecules that were then able to meet in an unchanged chemical state in the location where cells ultimately emerged [182].

4.2.2. The Food Source

A historically important division drawing from the analogy with different types of metabolism in prokaryotes [125] is the heterotrophy vs. autotrophy conundrum. Each side has envisioned their abiogenic scenario from a trophic point of view: if the carbon source for organosynthesis was eminently inorganic (usually CO_2), this represented an autotrophic origin (e.g. [183]), whereas if the carbon source were reduced organic compounds, a heterotrophic one (e.g. [184]). This division usually links to whether the earliest cells are considered autotrophic (e.g. [149]) or heterotrophic (e.g. [185]). It is important to note that for the first cells, the question is indeed meaningful—those cells either imported C_1 compounds, subsequently reducing them and forming C–C bonds (autotrophy) or imported C_n compounds and used those directly in biosynthesis [186,187]. But others are wary that this division may be unhelpful before the cellular stage (see Smith and Morowitz in [68]), since trophic types apply poorly to biosynthetic pathways. All organic molecules synthesised non-enzymatically ultimately derive from an inorganic carbon source (C, CO or CO_2) which was subsequently reduced. The endogenous (i.e. terrestrial origin) vs. exogenous (i.e. extraterrestrial origin) debate also entails limitations, by simply referring to the physical origin of the organic molecules, but not informing on the chemical relationship between these and the nascent proto-biochemistry. For the origin of the first biochemical networks, potentially acellular IDAs, we propose referring instead to whether all or some of the reactions leading to organic molecules occurred in situ (i.e. within, or physically contiguous to the nascent proto-biochemistry) or ex situ (i.e. physically separated from it). Note that this is not just a geographical distinction. Instead, it aims to find out whether the nascent chemical network (eventually leading to living cells) could affect (positively and/or negatively) the chemical reactions feeding it. This, we believe, is a meaningful change in language aiming to mend the limitations associated with the autotrophic vs. heterotrophic dilemma. Therefore, the in situ synthesis implies the reactions yielding biomolecules as part of the nascent protometabolism. In contrast, in the ex situ subtype the synthetic reactions rest unaffected by the

nascent chemical network, and the latter needs to invent alternative chemical pathways to eventually become independent from the former's unidirectional supply.

4.2.3. The Energy Source

Non-equilibrium conditions suggest an important bridge that can be established between most OoL hypotheses [14,188]. The fact that life is a process that is by definition not in equilibrium is not debatable in physics, biology or any other of the involved disciplines [189,190]. One example of how 'disequilibration' can be observed in life as we know it, is that life builds up complexity and then breaks it down again—in biology, this is, in broad strokes, anabolism and catabolism. These complex processes in general parallel energy spending and gaining in the cell, respectively. However, some organisms (e.g. methanogens) are known to conserve energy only through anabolism [191]. The earliest life would have had to couple both the constructive and destructive regime in some way [188,192]. To fulfil their energetic requirements, prebiotic systems would need both electron donors and acceptors supplied by their environment. The source of energy for the earliest life has been strongly disputed, with plausible hypotheses ranging from pH and redox gradients to thermal energy and UV light (for a detailed discussion see [74,193,194]. It is, however, not unlikely that several energy sources played a role in different prebiotic stages. Nevertheless, other constraints than the primary energy sources are imposed as soon as more complex prebiotic systems arise. A cell-like energy-coupling system could only be persistent over time if it can be inherited, thus forging a necessary link between the offspring hypotheses of the classical 'metabolism-first' and 'genetic-first' worlds [194].

4.2.4. RNA World versus Metabolism-First

Modern thinking on the OoL highlights the need for not only the synthesis of life's building blocks (themselves crucial goals), but also the reenactment of the processes they participate in [195]. Thankfully, what follows from this realization is that metabolism- and genetics-oriented hypotheses are gradually ceasing to be seen as mutually exclusive, because cellular processes involve both metabolites and genetic molecules [25,124]. A particular example regards evolution. Looking for the origin of natural selection and evolution is to look for the origin of an inheritance system that must maintain the capacity to produce new combinations, while keeping fidelity in information transfer [196,197]. But what *is* this information? Biological information is more than a collection of bits; it has contextuality, translated into functionality and is prone to evolve, as languages do [198]. In other words, genetic molecules only hold information in the right context: when they can be recognized and translated to functions that keep the system going. This definition opens different avenues to look for information storage and transmission during chemical evolution.

The earliest information could have been made from the types and quantities of molecules within a self-sustaining chemical assembly, also called 'composome' [199] or autocatalytic network [200]. This relates to the emphasis by some recent OoL research on the co-dependence of metabolism, containment and replication/information (see "Section 3.2. On the Right Track: Looking at the Last Decade"). This emphasis is central to bridging the RNA world and metabolism-first theories and points to the origin of molecular coordination or cooperation. Following this line of thought, some authors propose early 'selfish cooperators', akin to viruses, which formed stable ensembles of co-inherited genetic elements [60]. These ensembles should have been able to perform both a kind of proto-replication and a kind of proto-metabolism [201]. Still, there is a 'cooperation barrier' in the transition of non-life to life, caused by (i) molecules that could cooperatively contribute to the success of an ensemble but which are often not supported by the ensemble, and (ii) side reactions or processes that undermine cooperation [202]. To overcome this barrier, the management of those otherwise unconstrained ensembles is required [202]. This hypothesis is particularly interesting because it exposes the necessity of a digitally-coded (genetic, on/off) management of the analog (continuous) reactions of metabolism in order to overcome the cooperation barrier efficiently, reflecting the two-tiered structure of all known living cells [198,202].

If such analog information storage mechanisms as autocatalytic networks are able to undergo Darwinian evolution is a matter under debate [93,203]. Several have suggested fitness criteria for

natural chemical selection, including the rate of entropy production [204] and other kinetic or thermodynamic features of chemical reactions, such as stoichiometric catalysis, autocatalysis, and cooperativity [205]. A specific example is the greater reactivity of proteinaceous amino acids when compared to their non-proteinaceous counterparts, which naturally selects from oligomers of the former [206]. At some point, most likely quite early during the development of prebiotic systems, genetic molecules did crystallize, and information storage became based on molecular recognition. Every letter of the 'genetic alphabet', A, C, G, T and U, given a recognition/translation system, contains information. Early molecular recognition could have been significantly different to contemporary DNA and RNA [207,208] and even minerals have been proposed as information storage molecules [209].

What we know now, is that if we wish to search for answers on the origin of life as we know it—meaning the origin of cells—we are faced with an increasing conceptual [2] and experimental complexity [210] that will require the integration of RNA molecules with peptides, lipids, and protometabolism. Better models of complex biological systems [100], as well as better techniques to characterise them [211], will help to tackle such complexity. However, this complexity can no longer be denied. In other words, both the RNA-world and metabolism-first theories aim at the origin of life as we know it. However, we do not know it without each and both of them.

5. Conclusions

The main reason the OoL research field is still divided on so many issues is that it seems virtually impossible to find definite answers for all of our questions—we cannot wait hundreds of thousands of years to observe in real-time if a certain geological setting trumps another or which are the first enzymes to develop. We can only approach solutions asymptotically and, returning to the image of an OoL mosaic, add one pebble at a time, one insight that brings us closer to a more complete picture. In this article, we tried to convey how this is already being done—and hopefully, can be done even better in the future. Just as life itself is a synergistic process, in the sense that life's biological, physical and chemical properties are intertwined, the research that tries to explain its origin has to be synergistic as well. No discipline or approach should be dismissed, as long as its claims are evidence-based. However, we should keep in mind that while single results have dominated the field in the past, these can never constitute a solution to such a complex problem. Synthesising biomolecules, for example, although informative, cannot be enough to tackle the OoL, because life is not a mere collection of biomolecules, but rather a dynamic process involving biomolecules within boundary conditions interacting with its environment. Modern thinking on origins will thrive from recreating or imitating processes, rather than focusing only on the prebiotic synthesis of biomolecules.

Finally, we must highlight bias in OoL experiments, starting from how to define 'plausible' prebiotic conditions to which biomolecules are more central than others [212]. It is imperative to overcome personal biases in order to make progress. To get there, we need to increase the exchange, openness and respect between all those involved despite the inherent competition and inherited bias towards different hypotheses or approaches in research these days. The bridges we propose are essential, not because scientific research must be homogeneous, but because heterogeneity must be articulated (for examples see "Section 3. Building Bridges"). This articulation will change both the way research is done but also communication (in published articles, conferences and others). At this point, it is important to emphasize how crucial classical approaches in OoL research were and still are. However, while pivotal insights were achieved in classical studies, a more complete OoL narrative can only unfold by building bridges between them. This might all appear far too obvious to mention, but the reality in origin of life research shows that it cannot be mentioned often enough.

Author Contributions: Conceptualization, M.P. and J.C.X.; investigation, M.P., S.A., S.B., H.C.B., A.B., E.C., K.C., V.E., S.G.G., N.K., G.K., R.M., G.M., K.B.M., S.N., B.P., E.P., Á.R., D.R., A.S., N.M.S., F.L.S., F.D.K.T., D.V., J.C.X.; writing—original draft preparation, M.P. and J.C.X.; writing—review and editing, M.P., S.A., S.B., H.C.B., A.B., E.C., K.C., V.E., S.G.G., N.K., G.K., R.M., G.M., K.B.M., S.N., B.P., E.P., Á.R., D.R., A.S., N.M.S., F.L.S., F.D.K.T., D.V., J.C.X., in particular K.B.M. and E.C. on metabolism and energy, S.B. and A.R. on the genetic code and information, K.C. on messy chemistry, B.P. on protocells, G.K. on the philosophy of OoL, F.L.S. and F.T. on the

top-down approach, H.B. on the time frame of the OoL and N.K. on astrobiology; visualization, M.P. and J.C.X.; supervision, J.C.X.; project administration, M.P. and J.C.X.; funding acquisition, M.P. and J.C.X. All authors have read and agreed to the published version of the manuscript.

Funding: M.P. and J.C.X. are funded by grants from the Deutsche Forschungsgemeinschaft (MA-1426/21-1), the European Research Council (666053) and the Volkswagen Foundation (93046) to William F. Martin. E.C. is thankful to the support from the Origins Center through NWA StartImpuls. K.C. is supported by the European Structural and Investment Funds Operational Programme "Research, Development and Education"-funded project "ChemJets" (C.Z. 02.2.69/0.0/0.0/16_027/0008351) and Research Encouragement Fund UKM (GGP-2019-029). F.L.S. gratefully acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement 803768). F.L.S., S.N. and E.P. acknowledge the Wiener Wissenschafts-, Forschungs- und Technologiefonds (grant agreement VRG15-007, awarded to F.L.S.). D.V. and Á.R. were supported by the National Research, Development and Innovation Office (NKFIH) under the grant number GINOP 2.3.2-15-2016-00057. A.S. and B.P. were supported by the CRC 235 Emergence of Life granted to Dieter Braun and Petra Schwille, respectively. H.C.B. was supported by the NERC GW4 PhD studentship. N.K. was supported by European Research Council (ERC) Consolidator Grant 724908-Habitat OASIS.

Acknowledgements: All authors wish to thank William F. Martin (Heinrich Heine University, Düsseldorf) for generously sponsoring the entirety of the 1st Interdisciplinary Origin of Life Meeting for early career researchers in Düsseldorf, where many insightful discussions motivated this paper. The authors are thankful to Caroline Ring for the graphic recording of the entire meeting, available in https://www.molevol.hhu.de/iool-meeting-bilder.html. M.P. and J.C.X. thank William F. Martin for mentoring and support. S.B. thanks Thomas Carell for mentoring and support. K.B.M. gratefully acknowledges mentoring and support from Joseph Moran (Institut de Science et d'Ingénierie Supramoléculaires, Strasbourg). S.A. acknowledges mentoring and support by Leroy Cronin (School of Chemistry, University of Glasgow). S.N. and E.P. thank Filipa L. Sousa for mentoring and support. D.R. acknowledges Sheref S. Mansy for mentoring and support. E.C. acknowledges mentoring and support by Inge Loes ten Kate (Origins Center, Department of Earth Sciences, Utrecht University).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gayon, J.; Malaterre, C.; Morange, M.; Raulin-Cerceau, F.; Tirard, S. Defining Life: conference proceedings. *Origins. Life Evol. B.* **2010**, *40*, 119–120.
2. Mariscal, C.; Barahona, A.; Aubert-Kato, N.; Aydinoglu, A.U.; Bartlett, S.; Cárdenas, M.L.; Chandru, K.; Cleland, C.; Cocanougher, B.T.; Comfort, N.; et al. Hidden Concepts in the History and Philosophy of Origins-of-Life Studies: A Workshop Report. *Origins. Life Evol. B.* **2019**, *49*, 111–145.
3. Javaux, E.J. Challenges in evidencing the earliest traces of life. *Nature* **2019**, *572*, 451–460.
4. Betts, H.C.; Puttick, M.N.; Clark, J.W.; Williams, T.A.; Donoghue, P.C.J.; Pisani, D. Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin. *Nat. Ecol. Evol.* **2018**, *2*, 1556–1562.
5. Czárán, T.; Könnyű, B.; Szathmáry, E. Metabolically Coupled Replicator Systems: Overview of an RNA-world model concept of prebiotic evolution on mineral surfaces. *J. Theor. Biol.* **2015**, *381*, 39–54.
6. Davies, P.C.W.; Benner, S.A.; Cleland, C.E.; Lineweaver, C.H.; McKay, C.P.; Wolfe-Simon, F. Signatures of a shadow biosphere. *Astrobiology* **2009**, *9*, 241–249.
7. McLendon, C.; Opalko, F.J.; Illangkoon, H.I.; Benner, S.A. Solubility of polyethers in hydrocarbons at low temperatures. A model for potential genetic backbones on warm titans. *Astrobiology* **2015**, *15*, 200–206.
8. Wöhler, F. Ueber künstliche Bildung des Harnstoffs. *Ann. Phys. Chem.* **1828**, *87*, 253–256.
9. Butlerow, A. Bildung einer zuckerartigen Substanz durch Synthese. *Liebigs. Ann Chem.* **1861**, *120*, 295–298.
10. Breslow, R. On the mechanism of the formose reaction. *Tetrahedron Lett.* **1959**, *1*, 22–26.
11. Miller, S.L. A production of amino acids under possible primitive earth conditions. *Science* **1953**, *117*, 528–529.
12. Miller, S.L.; Urey, H.C. Organic compound synthesis on the primitive earth. *Science* **1959**, *130*, 245–251.

13. Hargreaves, W.R.; Mulvihill, S.J.; Deamer, D.W. Synthesis of phospholipids and membranes in prebiotic conditions. *Nature* **1977**, *266*, 78–80.
14. Ruiz-Mirazo, K.; Briones, C.; de la Escosura, A. Prebiotic systems chemistry: new perspectives for the origins of life. *Chem. Rev.* **2014**, *114*, 285–366.
15. Kitadai, N.; Maruyama, S. Origins of building blocks of life: A review. *Geosci. Front.* **2018**, *9*, 1117–1153.
16. Attwater, J.; Raguram, A.; Morgunov, A.S.; Gianni, E.; Holliger, P. Ribozyme-catalysed RNA synthesis using triplet building blocks. *eLife* **2018**, *7*.
17. Horning, D.P.; Joyce, G.F. Amplification of RNA by an RNA polymerase ribozyme. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 9786–9791.
18. Rajamani, S.; Vlassov, A.; Benner, S.; Coombs, A.; Olasagasti, F.; Deamer, D. Lipid-assisted synthesis of RNA-like polymers from mononucleotides. *Origins. Life Evol. B.* **2008**, *38*, 57–74.
19. Mamajanov, I.; MacDonald, P.J.; Ying, J.; Duncanson, D.M.; Dowdy, G.R.; Walker, C.A.; Engelhart, A.E.; Fernández, F.M.; Grover, M.A.; Hud, N.V.; et al. Ester Formation and Hydrolysis during Wet–Dry Cycles: Generation of Far-from-Equilibrium Polymers in a Model Prebiotic Reaction. *Macromolecules* **2014**, *47*, 1334–1343.
20. Griffith, E.C.; Vaida, V. In situ observation of peptide bond formation at the water-air interface. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 15697–15701.
21. Rodriguez-Garcia, M.; Surman, A.J.; Cooper, G.J.T.; Suárez-Marina, I.; Hosni, Z.; Lee, M.P.; Cronin, L. Formation of oligopeptides in high yield under simple programmable conditions. *Nat. Commun.* **2015**, *6*, 8385.
22. Erastova, V.; Degiacomi, M.T.; G Fraser, D.; Greenwell, H.C. Mineral surface chemistry control for origin of prebiotic peptides. *Nat. Commun.* **2017**, *8*, 2033.
23. Gibson, D.G.; Benders, G.A.; Andrews-Pfannkoch, C.; Denisova, E.A.; Baden-Tillson, H.; Zaveri, J.; Stockwell, T.B.; Brownley, A.; Thomas, D.W.; Algire, M.A.; et al. Complete chemical synthesis, assembly, and cloning of a Mycoplasma genitalium genome. *Science* **2008**, *319*, 1215–1220.
24. Cello, J.; Paul, A.V.; Wimmer, E. Chemical synthesis of poliovirus cDNA: Generation of infectious virus in the absence of natural template. *Science* **2002**, *297*, 1016–1018.
25. Walker, S.I.; Packard, N.; Cody, G.D. Re-conceptualizing the origins of life. *Philos. Trans. A Math. Phys. Eng. Sci.* **2017**, *375*.
26. Locey, K.J.; Lennon, J.T. Scaling laws predict global microbial diversity. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 5970–5975.
27. Milo, R. What is the total number of protein molecules per cell volume? A call to rethink some published values. *BioEssays* **2013**, *35*, 1050–1055.
28. Mayr, E. Cause and effect in biology. *Science* **1961**, *134*, 1501–1506.
29. Fry, I. The origins of research into the origins of life. *Endeavour* **2006**, *30*, 24–28.
30. Cleland, C.E. Pluralism or unity in biology: could microbes hold the secret to life? *Biol. Philos.* **2013**, *28*, 189–204.
31. Mushegian, A.R.; Koonin, E.V. A minimal gene set for cellular life derived by comparison of complete bacterial genomes. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 10268–10273.
32. Koonin, E.V. Comparative genomics, minimal gene-sets and the last universal common ancestor. *Nat. Rev. Microbiol.* **2003**, *1*, 127–136.
33. Lagesen, K.; Ussery, D.W.; Wassenaar, T.M. Genome update: the 1000th genome-a cautionary tale. *Microbiology* **2010**, *156*, 603–608.

34. Sutherland, J.D. The Origin of Life--Out of the Blue. *Angew. Chem. Int. Ed Engl.* **2016**, *55*, 104–121.
35. Ralser, M. An appeal to magic? The discovery of a non-enzymatic metabolism and its role in the origins of life. *Biochem. J* **2018**, *475*, 2577–2592.
36. Yarus, M. Getting past the RNA world: the initial Darwinian ancestor. *Cold Spring Harb. Perspect. Biol.* **2011**, *3*.
37. Prosdocimi, F.; José, M.V.; Farias, S.T. de Be Introduced to the First Universal Common Ancestor (FUCA): The Great-Grandmother of LUCA (Last Universal Common Ancestor). **2018**. DOI: 10.20944/preprints201806.0035.v1. Available online: https://europepmc.org/article/ppr/ppr49297 (accessed on: 25th February 2020)
38. Lahav, N.; Nir, S.; Elitzur, A.C. The emergence of life on Earth. *Prog. Biophys. Mol. Biol.* **2001**, *75*, 75–120.
39. Ikehara, K. Evolutionary Steps in the Emergence of Life Deduced from the Bottom-Up Approach and GADV Hypothesis (Top-Down Approach). *Life* **2016**, *6*.
40. Vicens, J.; Vicens, Q. Emergences of supramolecular chemistry: from supramolecular chemistry to supramolecular science. *J. Incl. Phenom. Macrocycl. Chem.* **2011**, *71*, 251–274.
41. Rich, A. On the problems of evolution and biochemical information transfer. In *Horizons In Biochemistry*; Kash, M., Pullman, B., Eds.; Academic Press: New York, NY, USA, 1962; pp. 103–126.
42. Gilbert, W. Origin of life: The RNA world. *Nature* **1986**, *319*, 618–618.
43. Kruger, K.; Grabowski, P.J.; Zaug, A.J.; Sands, J.; Gottschling, D.E.; Cech, T.R. Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* **1982**, *31*, 147–157.
44. Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **1983**, *35*, 849–857.
45. Robertson, M.P.; Joyce, G.F. The origins of the RNA world. *Cold Spring Harb. Perspect. Biol.* **2012**, *4*.
46. Orgel, L.E. Prebiotic chemistry and the origin of the RNA world. *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 99–123.
47. Shapiro, R. A Replicator Was Not Involved in the Origin of Life. *IUBMB: Life* **2000**, *49*, 173–176.
48. Oivanen, M.; Kuusela, S.; Lönnberg, H. Kinetics and Mechanisms for the Cleavage and Isomerization of the Phosphodiester Bonds of RNA by Brønsted Acids and Bases. *Chem. Rev.* **1998**, *98*, 961–990.
49. Wächtershäuser, G. The origin of life and its methodological challenge. *J. Theor. Biol.* **1997**, *187*, 483–494.
50. Dyson, F. *Origins of Life,* Cambridge University Press: Cambridge, UK, 1999.
51. de Duve, C. A Research Proposal on the Origin Of Life. Closing Lecture given at the ISSOL Congress in Oaxaca, Mexico, on July 4, 2002. *Origins Life Evol. B.* **2003**, *33*, 559–574.
52. Smith, E.; Morowitz, H.J. Universality in intermediary metabolism. *Proc. Natl. Acad. Sci. U.S.A* 2004, *101*, 13168–13173.
53. Orgel, L.E. The implausibility of metabolic cycles on the prebiotic Earth. *PLoS Biol.* **2008**, *6*, e18.
54. Kamminga, H. Historical perspective: the problem of the origin of life in the context of developments in biology. *Origins. Life Evol. B.* **1988**, *18*, 1–11.
55. Caetano-Anollés, G.; Wang, M.; Caetano-Anollés, D.; Mittenthal, J.E. The origin, evolution and structure of the protein world. *Biochem. J.* **2009**, *417*, 621–637.
56. Plankensteiner, K.; Reiner, H.; Rode, B. Prebiotic Chemistry: The Amino Acid and Peptide World. *Curr. Org. Chem.* **2005**, *9*, 1107–1114.
57. Segré, D.; Ben-Eli, D.; Deamer, D.W.; Lancet, D. The lipid world. *Origins Life Evol. B.* **2001**, *31*, 119–145.
58. Tessera, M. Origin of evolution versus origin of life: A shift of paradigm. *Int. J. Mol. Sci.* **2011**, *12*, 3445–3458.

59. Sharov, A.A. Coenzyme world model of the origin of life. *Biosystems.* **2016**, *144*, 8–17.
60. Koonin, E.V.; Senkevich, T.G.; Dolja, V.V. The ancient Virus World and evolution of cells. *Biol. Direct.* **2006**, *1*, 29.
61. Lanier, K.A.; Williams, L.D. The Origin of Life: Models and Data. *J. Mol. Evol.* **2017**, *84*, 85–92.
62. Xavier, J.C.; Patil, K.R.; Rocha, I. Systems biology perspectives on minimal and simpler cells. *Microbiol. Mol. Biol. Rev.* **2014**, *78*, 487–509.
63. Pross, A. Causation and the Origin of Life. Metabolism or Replication First? *Origins Life Evol. B.* **2004**, *34*, 307–321.
64. Wächtershäuser, G. In Praise of Error. *J. Mol. Evol.* **2016**, *82*, 75–80.
65. Sutherland, J.D. Opinion: Studies on the origin of life — the end of the beginning. *Nat. Rev. Chem.* 2017, *1*.
66. Ross, D.S. It is Neither Frankenstein Nor a Submarine Alkaline Vent, It is Just the Second Law. *BioEssays* **2018**, *40*, 1800149.
67. Branscomb, E.; Russell, M.J. Frankenstein or a Submarine Alkaline Vent: Who is Responsible for Abiogenesis? *BioEssays* **2018**, *40*, 1700182.
68. Emergence, O. Workshop OQOL'09: Open Questions on the Origins of Life 2009. *Origins Life Evol. B.* **2010**, *40*, 347–497.
69. Oparin, A. I. The origin of life (A. Synge, transl) In *The Origin of Life*, Bernal, J. D., Weidenfeld & Nicolson: London, U.K. 1967 pp. 197–234.
70. Haldane, J.B.S. The Origin of Life. *Ration. Annu.* **1929**, 3–10.
71. Pennington, D.D.; Simpson, G.L.; McConnell, M.S.; Fair, J.M.; Baker, R.J. Transdisciplinary Research, Transformative Learning, and Transformative Science. *BioScience* **2013**, *63*, 564–573.
72. Polanco, C, Why interdisciplinary research matters. *Nature* **2015**, *525*, 305.
73. Morowitz, H.; Smith, E. Energy flow and the organization of life. *Complexity* **2007**, *13*, 51–59.
74. Kaufmann, M. On the free energy that drove primordial anabolism. *Int. J. Mol. Sci.* **2009**, *10*, 1853–1871.
75. Zhang, W.; Li, F.; Nie, L. Integrating multiple "omics" analysis for microbial biology: application and methodologies. *Microbiology* **2010**, *156*, 287–301.
76. Larsen, P.; Hamada, Y.; Gilbert, J. Modeling microbial communities: current, developing, and future technologies for predicting microbial community interaction. *J. Biotechnol.* **2012**, *160*, 17–24.
77. Waite, J.H.; Glein, C.R.; Perryman, R.S.; Teolis, B.D.; Magee, B.A.; Miller, G.; Grimes, J.; Perry, M.E.; Miller, K.E.; Bouquet, A.; et al. Cassini finds molecular hydrogen in the Enceladus plume: Evidence for hydrothermal processes. *Science* **2017**, *356*, 155–159.
78. Zahnle, K.; Arndt, N.; Cockell, C.; Halliday, A.; Nisbet, E.; Selsis, F.; Sleep, N.H. Emergence of a Habitable Planet. *Space Sci. Rev.* **2007**, 35–78.
79. Ehrenfreund, P.; Irvine, W.; Becker, L.; Blank, J.; Brucato, J.R.; Colangeli, L.; Derenne, S.; Despois, D.; Dutrey, A.; Fraaije, H.; et al. Astrophysical and astrochemical insights into the origin of life. *Rep. Prog. Phys.* **2002**, *65*, 1427–1487.
80. Vance, S.D. The Habitability of Icy Ocean Worlds in the Solar System. In *Handbook of Exoplanets*; Springer: Berlin, Germany, 2018; pp. 1–23.
81. Postberg, F.; Khawaja, N.; Abel, B.; Choblet, G.; Glein, C.R.; Gudipati, M.S.; Henderson, B.L.; Hsu, H.-W.; Kempf, S.; Klenner, F.; et al. Macromolecular organic compounds from the depths of Enceladus. *Nature* **2018**, *558*, 564–568.
82. Khawaja, N.; Postberg, F.; Hillier, J.; Klenner, F.; Kempf, S.; Nölle, L.; Reviol, R.; Zou, Z.; Srama, R. Low-mass nitrogen-, oxygen-bearing, and aromatic compounds in Enceladean ice grains. *Mon. Not. R. Astron.*

Soc. **2019**, *489*, 5231–5243.

83. Hörst, S.M.; Yelle, R.V.; Buch, A.; Carrasco, N.; Cernogora, G.; Dutuit, O.; Quirico, E.; Sciamma-O'Brien, E.; Smith, M.A.; Somogyi, Á.; et al. Formation of Amino Acids and Nucleotide Bases in a Titan Atmosphere Simulation Experiment. *Astrobiology* **2012**, *12*, 809–817.
84. Sousa, F.L.; Thiergart, T.; Landan, G.; Nelson-Sathi, S.; Pereira, I.A.C.; Allen, J.F.; Lane, N.; Martin, W.F. Early bioenergetic evolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2013**, *368*, 20130088.
85. Müller, V.; Chowdhury, N.P.; Basen, M. Electron Bifurcation: A Long-Hidden Energy-Coupling Mechanism. *Annu. Rev. Microbiol.* **2018**, *72*, 331–353.
86. Whicher, A.; Camprubi, E.; Pinna, S.; Herschy, B.; Lane, N. Acetyl Phosphate as a Primordial Energy Currency at the Origin of Life. *Origins. Life Evol. B.* **2018**, *48*, 159–179.
87. Duve, C.D. Clues from present-day biology: the thioester world. In *The Molecular Origins of Life*; Brack, A., Ed.; Cambridge University Press: Cambridge, UK, 1998; pp. 219–236.
88. 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*, 29883.
89. Schoepp-Cothenet, B.; van Lis, R.; Atteia, A.; Baymann, F.; Capowiez, L.; Ducluzeau, A.-L.; Duval, S.; ten Brink, F.; Russell, M.J.; Nitschke, W. On the universal core of bioenergetics. *Biochim. Biophys. Acta* **2013**, *1827*, 79–93.
90. Jinich, A.; Sanchez-Lengeling, B.; Ren, H.; Goldford, J.E.; Noor, E.; Sanders, J.N.; Segrè, D.; Aspuru-Guzik, A. A thermodynamic atlas of carbon redox chemical space. *BioRxiv* **2019**, 245811
91. Mattia, E.; Otto, S. Supramolecular systems chemistry. *Nat. Nanotechnol.* **2015**, *10*, 111–119.
92. Kim, H.; Smith, H.B.; Mathis, C.; Raymond, J.; Walker, S.I. Universal scaling across biochemical networks on Earth. *Sci Adv* **2019**, *5*, eaau0149.
93. Markovitch, O.; Lancet, D. Excess mutual catalysis is required for effective evolvability. *Artif. Life* **2012**, *18*, 243–266.
94. Pascal, R.; Pross, A. Stability and its manifestation in the chemical and biological worlds. *Chem. Commun.* **2015**, *51*, 16160–16165.
95. Xavier, J.C.; Hordijk, W.; Kauffman, S.; Steel, M.; Martin, W.F. Autocatalytic chemical networks preceded proteins and RNA in evolution. *bioRxiv* doi: 10.1101/693879 . Available online: https://www.biorxiv.org/content/10.1101/693879v1.abstract. (accessed on: on: 25th February 2020)
96. Goldford, J.E.; Hartman, H.; Smith, T.F.; Segrè, D. Remnants of an Ancient Metabolism without Phosphate. *Cell* **2017**, *168*, 1126–1134.e9.
97. Goldford, J.E.; Segrè, D. Modern views of ancient metabolic networks. *Curr. Opin. Syst. Biol.* **2018**, *8*, 117–124.
98. Muchowska, K.B.; Varma, S.J.; Moran, J. Synthesis and breakdown of universal metabolic precursors promoted by iron. *Nature* **2019**, *569*, 104–107.
99. Keller, M.A.; Turchyn, A.V.; Ralser, M. Non-enzymatic glycolysis and pentose phosphate pathway-like reactions in a plausible Archean ocean. *Mol. Syst. Biol.* **2014**, *10*, 725–725.
100. Vincent, L.; Berg, M.; Krismer, M.; Saghafi, S.S.; Cosby, J.; Sankari, T.; Vetsigian, K.; Ii, H.J.C.; Baum, D.A. Chemical Ecosystem Selection on Mineral Surfaces Reveals Long-Term Dynamics Consistent with the Spontaneous Emergence of Mutual Catalysis. *Life* **2019**, *9*.
101. Amend, J.P.; LaRowe, D.E.; McCollom, T.M.; Shock, E.L. The energetics of organic synthesis inside and outside the cell. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2013**, *368*, 20120255.
102. Russell, M.J.; Martin, W. The rocky roots of the acetyl-CoA pathway. *Trends Biochem. Sci.* **2004**, *29*, 358–363.

103. Preiner, M.; Igarashi, K.; Muchowska, K.B.; Yu, M.; Varma, S.J.; Kleinermanns, K.; Nobu, M.K.; Kamagata, Y.; Tüysüz, H.; Moran, J.; Martin, W.F. A hydrogen-dependent geochemical analogue of primordial carbon and energy metabolism. *Nat. Ecol. Evol.* **2020**, in press; doi: 10.1038/s41559-020-1125-6
104. Barge, L.M.; White, L.M. Experimentally Testing Hydrothermal Vent Origin of Life on Enceladus and Other Icy/Ocean Worlds. *Astrobiology* **2017**, *17*, 820–833.
105. Adami, C.; Labar, T. From Entropy to Information: Biased Typewriters and the Origin of Life. *arXiv* preprint arXiv:1506.06988. Available online: https://arxiv.org/abs/1506.06988 (accessed on: 25th February 2020)
106. Turk, R.M.; Chumachenko, N.V.; Yarus, M. Multiple translational products from a five-nucleotide ribozyme. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 4585–4589.
107. Szathmáry, E. The origin of the genetic code: amino acids as cofactors in an RNA world. *Trends Genet.* **1999**, *15*, 223–229.
108. Morris, C.E. How did cells get their size? *Anat. Rec.* **2002**, *268*, 239–251.
109. Noller, H.F. The driving force for molecular evolution of translation. *RNA* **2004**, *10*, 1833–1837.
110. de Vladar, H.P. Amino acid fermentation at the origin of the genetic code. *Biol. Direct* **2012**, *7*, 6.
111. Tagami, S.; Attwater, J.; Holliger, P. Simple peptides derived from the ribosomal core potentiate RNA polymerase ribozyme function. *Nat. Chem.* **2017**, *9*, 325–332.
112. Maddox, J. The genetic code by numbers. *Nature* **1994**, *367*, 111–111.
113. Yarus, M. The Genetic Code and RNA-Amino Acid Affinities. *Life* **2017**, *7*, 13.
114. Wong, J.T. A co-evolution theory of the genetic code. *Proc. Natl. Acad. Sci. U. S. A.* **1975**, *72*, 1909–1912.
115. Knight, R.D.; Freeland, S.J.; Landweber, L.F. Selection, history and chemistry: the three faces of the genetic code. *Trends Biochem. Sci.* **1999**, *24*, 241–247.
116. Kun, Á.; Radványi, Á. The evolution of the genetic code: Impasses and challenges. *Biosystems.* **2018**, *164*, 217–225.
117. Higgs, P.G.; Pudritz, R.E. A thermodynamic basis for prebiotic amino acid synthesis and the nature of the first genetic code. *Astrobiology* **2009**, *9*, 483–490.
118. Müller, M.M.; Allison, J.R.; Hongdilokkul, N.; Gaillon, L.; Kast, P.; van Gunsteren, W.F.; Marlière, P.; Hilvert, D. Directed evolution of a model primordial enzyme provides insights into the development of the genetic code. *PLoS Genet.* **2013**, *9*, e1003187.
119. Ilardo, M.; Bose, R.; Meringer, M.; Rasulev, B.; Grefenstette, N.; Stephenson, J.; Freeland, S.; Gillams, R.J.; Butch, C.J.; Cleaves, H.J., 2nd Adaptive Properties of the Genetically Encoded Amino Acid Alphabet Are Inherited from Its Subsets. *Sci. Rep.* **2019**, *9*, 12468.
120. Cartwright, J.H.E.; Giannerini, S.; González, D.L. DNA as information: at the crossroads between biology, mathematics, physics and chemistry. *Philos. Trans. A Math. Phys. Eng. Sci.* **2016**, *374*.
121. Woese, C.R.; Dugre, D.H.; Saxinger, W.C.; Dugre, S.A. The molecular basis for the genetic code. *Proc. Natl. Acad. Sci. U. S. A.* **1966**, *55*, 966–974.
122. Copley, S.D.; Smith, E.; Morowitz, H.J. A mechanism for the association of amino acids with their codons and the origin of the genetic code. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 4442–4447.
123. Koonin, E.V.; Novozhilov, A.S. Origin and Evolution of the Universal Genetic Code. *Annu. Rev. Genet.* **2017**, *51*, 45–62.
124. Scharf, C.; Virgo, N.; Cleaves, H.J., 2nd; Aono, M.; Aubert-Kato, N.; Aydinoglu, A.; Barahona, A.; Barge, L.M.; Benner, S.A.; Biehl, M.; et al. A Strategy for Origins of Life Research. *Astrobiology* **2015**, *15*, 1031–1042.
125. Gutekunst, K. Hypothesis on the Synchronistic Evolution of Autotrophy and Heterotrophy. *Trends Biochem. Sci.* **2018**, *43*, 402–411.

126. Mansy, S.S.; Szostak, J.W. Reconstructing the emergence of cellular life through the synthesis of model protocells. *Cold Spring Harb. Symp. Quant. Biol.* **2009**, *74*, 47–54.
127. Cornell, C.E.; Black, R.A.; Xue, M.; Litz, H.E.; Ramsay, A.; Gordon, M.; Mileant, A.; Cohen, Z.R.; Williams, J.A.; Lee, K.K.; et al. Prebiotic amino acids bind to and stabilize prebiotic fatty acid membranes. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 17239–17244.
128. Jordan, S.F.; Rammu, 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*, 1705–1714.
129. Monnard, P.-A.; Walde, P. Current Ideas about Prebiological Compartmentalization. *Life* **2015**, *5*, 1239–1263.
130. Dzieciol, A.J.; Mann, S. Designs for life: protocell models in the laboratory. *Chem. Soc. Rev.* **2012**, *41*, 79–85.
131. Gardner, P.M.; Winzer, K.; Davis, B.G. Sugar synthesis in a protocellular model leads to a cell signalling response in bacteria. *Nat. Chem.* **2009**, *1*, 377–383.
132. Kurihara, K.; Tamura, M.; Shohda, K.-I.; Toyota, T.; Suzuki, K.; Sugawara, T. Self-reproduction of supramolecular giant vesicles combined with the amplification of encapsulated DNA. *Nat. Chem.* **2011**, *3*, 775–781.
133. Gumpenberger, T.; Vorkapic, D.; Zingl, F.G.; Pressler, K.; Lackner, S.; Seper, A.; Reidl, J.; Schild, S. Nucleoside uptake in Vibrio cholerae and its role in the transition fitness from host to environment. *Mol. Microbiol.* **2016**, *99*, 470–483.
134. White, H.B., 3rd Coenzymes as fossils of an earlier metabolic state. *J. Mol. Evol.* **1976**, *7*, 101–104.
135. Bhowmik, S.; Krishnamurthy, R. The role of sugar-backbone heterogeneity and chimeras in the simultaneous emergence of RNA and DNA. *Nat. Chem.* **2019**, *11*, 1009–1018.
136. Trifonov, E.N. Consensus temporal order of amino acids and evolution of the triplet code. *Gene* **2000**, *261*, 139–151.
137. Parker, E.T.; Cleaves, H.J.; Bada, J.L. Quantitation of α-hydroxy acids in complex prebiotic mixtures via liquid chromatography/tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **2016**. 30, 2043–2051.
138. Sephton, M.A. Organic compounds in carbonaceous meteorites. *Nat. Prod. Rep.* **2002**, *19*, 292–311.
139. Johnson, A.P.; Cleaves, H.J.; Dworkin, J.P.; Glavin, D.P.; Lazcano, A.; Bada, J.L. The Miller volcanic spark discharge experiment. *Science* **2008**, *322*, 404.
140. Chandru, K.; Guttenberg, N.; Giri, C.; Hongo, Y.; Butch, C.; Mamajanov, I.; Cleaves, H.J. Simple prebiotic synthesis of high diversity dynamic combinatorial polyester libraries. *Commun. Chem.* **2018**, *1*, 30.
141. Jia, T.Z.; Chandru, K.; Hongo, Y.; Afrin, R.; Usui, T.; Myojo, K.; Cleaves, H.J. Membraneless polyester microdroplets as primordial compartments at the origins of life. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 15830–15835.
142. Chandru, K.; Mamajanov, I.; Cleaves, H.J., 2nd; Jia, T.Z. Polyesters as a Model System for Building Primitive Biologies from Non-Biological Prebiotic Chemistry. *Life* **2020**, *10*.
143. Guttenberg, N.; Virgo, N.; Chandru, K.; Scharf, C.; Mamajanov, I. Bulk measurements of messy chemistries are needed for a theory of the origins of life. *Philos. Trans. Royal Soc. A.* **2017**, *375*.
144. Schmitt-Kopplin, P.; Gabelica, Z.; Gougeon, R.D.; Fekete, A.; Kanawati, B.; Harir, M.; Gebefuegi, I.; Eckel, G.; Hertkorn, N. High molecular diversity of extraterrestrial organic matter in Murchison meteorite revealed 40 years after its fall. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 2763–2768.
145. Arndt, N.T.; Nisbet, E.G. Processes on the Young Earth and the Habitats of Early Life. *Annu. Rev. Earth Planet. Sci.* **2012**, *40*, 521–549.
146. Nisbet, E.G.; Sleep, N.H. The habitat and nature of early life. *Nature* **2001**, *409*, 1083–1091.

147. Lang, S.Q.; Butterfield, D.A.; Schulte, M.; Kelley, D.S.; Lilley, M.D. Elevated concentrations of formate, acetate and dissolved organic carbon found at the Lost City hydrothermal field. *Geochim. Cosmochim. Acta* **2010**, *74*, 941–952.
148. Herschy, B.; Whicher, A.; Camprubi, E.; Watson, C.; Dartnell, L.; Ward, J.; Evans, J.R.G.; Lane, N. An origin-of-life reactor to simulate alkaline hydrothermal vents. *J. Mol. Evol.* **2014**, *79*, 213–227.
149. Weiss, M.C.; Sousa, F.L.; Mrnjavac, N.; Neukirchen, S.; Roettger, M.; Nelson-Sathi, S.; Martin, W.F. The physiology and habitat of the last universal common ancestor. *Nat Microbiol* **2016**, *1*, 16116.
150. Mulkidjanian, A.Y.; Bychkov, A.Y.; Dibrova, D.V.; Galperin, M.Y.; Koonin, E.V. Origin of first cells at terrestrial, anoxic geothermal fields. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, E821–30.
151. Milshteyn, D.; Damer, B.; Havig, J.; Deamer, D. Amphiphilic Compounds Assemble into Membranous Vesicles in Hydrothermal Hot Spring Water but Not in Seawater. *Life* **2018**, *8*, 11.
152. Becker, S.; Schneider, C.; Okamura, H.; Crisp, A.; Amatov, T.; Dejmek, M.; Carell, T. Wet-dry cycles enable the parallel origin of canonical and non-canonical nucleosides by continuous synthesis. *Nat. Commun.* **2018**, *9*, 163.
153. Varma, S.J.; Muchowska, K.B.; Chatelain, P.; Moran, J. Native iron reduces CO to intermediates and end-products of the acetyl-CoA pathway. *Nat Ecol Evol* **2018**, *2*, 1019–1024.
154. Garcia, A.K.; McShea, H.; Kolaczkowski, B.; Kaçar, B. Reconstructed ancient nitrogenases suggest Mo-specific ancestry. *Evol. Biol.* **2019**, 520.
155. Bray, M.S.; Lenz, T.K.; Haynes, J.W.; Bowman, J.C.; Petrov, A.S.; Reddi, A.R.; Hud, N.V.; Williams, L.D.; Glass, J.B. Multiple prebiotic metals mediate translation. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 12164–12169.
156. Bonfio, C.; Godino, E.; Corsini, M.; de Biani, F.F.; Guella, G.; Mansy, S.S. Prebiotic iron–sulfur peptide catalysts generate a pH gradient across model membranes of late protocells. *Nat. Catal.* **2018**, *1*, 616–623.
157. Kreysing, M.; Keil, L.; Lanzmich, S.; Braun, D. Heat flux across an open pore enables the continuous replication and selection of oligonucleotides towards increasing length. *Nat. Chem.* **2015**, *7*, 203–208.
158. Hazen, R.M.; Sverjensky, D.A. Mineral surfaces, geochemical complexities, and the origins of life. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, a002162.
159. Colón-Santos, S.; Cooper, G.J.T.; Cronin, L. Taming the Combinatorial Explosion of the Formose Reaction via Recursion within Mineral Environments. *ChemSystemsChem* **2019**, *54*, 104.
160. Sadownik, J.W.; Mattia, E.; Nowak, P.; Otto, S. Diversification of self-replicating molecules. *Nat. Chem.* **2016**, *8*, 264–269.
161. Semenov, S.N.; Kraft, L.J.; Ainla, A.; Zhao, M.; Baghbanzadeh, M.; Campbell, V.E.; Kang, K.; Fox, J.M.; Whitesides, G.M. Autocatalytic, bistable, oscillatory networks of biologically relevant organic reactions. *Nature* **2016**, *537*, 656–660.
162. Stelling, J. Mathematical models in microbial systems biology. *Curr. Opin. Microbiol.* **2004**, *7*, 513–518.
163. Goldman, A.D.; Bernhard, T.M.; Dolzhenko, E.; Landweber, L.F. LUCApedia: a database for the study of ancient life. *Nucleic Acids Res.* **2013**, *41*, D1079–82.
164. Nghe, P.; Hordijk, W.; Kauffman, S.A.; Walker, S.I.; Schmidt, F.J.; Kemble, H.; Yeates, J.A.M.; Lehman, N. Prebiotic network evolution: six key parameters. *Mol. Biosyst.* **2015**, *11*, 3206–3217.
165. Woese, C. The universal ancestor. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 6854–6859.
166. Tocheva, E.I.; Ortega, D.R.; Jensen, G.J. Sporulation, bacterial cell envelopes and the origin of life. *Nat. Rev. Microbiol.* **2016**, *14*, 535–542.
167. Hastings, J.; de Matos, P.; Dekker, A.; Ennis, M.; Harsha, B.; Kale, N.; Muthukrishnan, V.; Owen, G.; Turner, S.; Williams, M.; et al. The ChEBI reference database and ontology for biologically relevant chemistry:

enhancements for 2013. *Nucleic Acids Res.* **2013**, *41*, D456–63.

168. Kanehisa, M.; Furumichi, M.; Tanabe, M.; Sato, Y.; Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **2017**, *45*, D353–D361.
169. King, Z.A.; Lu, J.; Dräger, A.; Miller, P.; Federowicz, S.; Lerman, J.A.; Ebrahim, A.; Palsson, B.O.; Lewis, N.E. BiGG Models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Res.* **2016**, *44*, D515–22.
170. Kim, K.M.; Caetano-Anollés, G. The proteomic complexity and rise of the primordial ancestor of diversified life. *BMC Evol. Biol.* **2011**, *11*, 140.
171. Delaye, L.; Becerra, A. Cenancestor, the Last Universal Common Ancestor. *Evo Edu Outreach* **2012**, *5*, 382–388.
172. Theobald, D.L. A formal test of the theory of universal common ancestry. *Nature* **2010**, *465*, 219–222.
173. Surman, A.J.; Rodriguez-Garcia, M.; Abul-Haija, Y.M.; Cooper, G.J.T.; Gromski, P.S.; Turk-MacLeod, R.; Mullin, M.; Mathis, C.; Walker, S.I.; Cronin, L. Environmental control programs the emergence of distinct functional ensembles from unconstrained chemical reactions. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 5387–5392.
174. Sleep, N.H. Geological and Geochemical Constraints on the Origin and Evolution of Life. *Astrobiology* **2018**, *18*, 1199–1219.
175. Martin, W.; Russell, M.J. On the origin of biochemistry at an alkaline hydrothermal vent. *Philos. Trans. R. Soc. B.* **2007**, *362*, 1887–1926.
176. Martin, W.; Baross, J.; Kelley, D.; Russell, M.J. Hydrothermal vents and the origin of life. *Nat. Rev. Microbiol.* **2008**, *6*, 805–814.
177. Deamer, D. The Role of Lipid Membranes in Life's Origin. *Life* **2017**, *7*.
178. Westall, F.; Hickman-Lewis, K.; Hinman, N.; Gautret, P.; Campbell, K.A.; Bréhéret, J.G.; Foucher, F.; Hubert, A.; Sorieul, S.; Dass, A.V.; et al. A Hydrothermal-Sedimentary Context for the Origin of Life. *Astrobiology* **2018**, *18*, 259–293.
179. Pearce, B.K.D.; Pudritz, R.E.; Semenov, D.A.; Henning, T.K. Origin of the RNA world: The fate of nucleobases in warm little ponds. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 11327–11332.
180. Ross, D.S.; Deamer, D. Dry/Wet Cycling and the Thermodynamics and Kinetics of Prebiotic Polymer Synthesis. *Life* **2016**, *6*.
181. Morasch, M.; Liu, J.; Dirscherl, C.F.; Ianeselli, A.; Kühnlein, A.; Le Vay, K.; Schwintek, P.; Islam, S.; Corpinot, M.K.; Scheu, B.; et al. Heated gas bubbles enrich, crystallize, dry, phosphorylate and encapsulate prebiotic molecules. *Nat. Chem.* **2019**, *11*, 779–788.
182. Becker, S.; Feldmann, J.; Wiedemann, S.; Okamura, H.; Schneider, C.; Iwan, K.; Crisp, A.; Rossa, M.; Amatov, T.; Carell, T. Unified prebiotically plausible synthesis of pyrimidine and purine RNA ribonucleotides. *Science* **2019**, *366*, 76–82.
183. Ooka, H.; McGlynn, S.E.; Nakamura, R. Electrochemistry at Deep-Sea Hydrothermal Vents: Utilization of the Thermodynamic Driving Force towards the Autotrophic Origin of Life. *ChemElectroChem* **2019**, *6*, 1316–1323.
184. Mansy, S.S.; Schrum, J.P.; Krishnamurthy, M.; Tobé, S.; Treco, D.A.; Szostak, J.W. Template-directed synthesis of a genetic polymer in a model protocell. *Nature* **2008**, *454*, 122–125.
185. Lazcano, A. Alexandr I. Oparin and the Origin of Life: A Historical Reassessment of the Heterotrophic Theory. *J. Mol. Evol.* **2016**, *83*, 214–222.
186. Wood, A.P.; Aurikko, J.P.; Kelly, D.P. A challenge for 21st century molecular biology and biochemistry:

what are the causes of obligate autotrophy and methanotrophy? *FEMS Microbiol. Rev.* **2004**, *28*, 335–352.

187. Schönheit, P.; Buckel, W.; Martin, W.F. On the Origin of Heterotrophy. *Trends Microbiol.* **2016**, *24*, 12–25.
188. Ashkenasy, G.; Hermans, T.M.; Otto, S.; Taylor, A.F. Systems chemistry. *Chem. Soc. Rev.* **2017**, *46*, 2543–2554.
189. Pascal, R.; Pross, A.; Sutherland, J.D. Towards an evolutionary theory of the origin of life based on kinetics and thermodynamics. *Open Biol.* **2013**, *3*, 130156.
190. Semenov, S.N.; Wong, A.S.Y.; van der Made, R.M.; Postma, S.G.J.; Groen, J.; van Roekel, H.W.H.; de Greef, T.F.A.; Huck, W.T.S. Rational design of functional and tunable oscillating enzymatic networks. *Nat. Chem.* **2015**, *7*, 160–165.
191. Thauer, R.K.; Kaster, A.-K.; Seedorf, H.; Buckel, W.; Hedderich, R. Methanogenic archaea: ecologically relevant differences in energy conservation. *Nat. Rev. Microbiol.* **2008**, *6*, 579–591.
192. Duim, H.; Otto, S. Towards open-ended evolution in self-replicating molecular systems. *Beilstein J. Org. Chem.* **2017**, *13*, 1189–1203.
193. Deamer, D.; Weber, A.L. Bioenergetics and Life's Origins. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, a004929–a004929.
194. Boiteau, L.; Pascal, R. Energy sources, self-organization, and the origin of life. *Origins. Life Evol. B.* **2011**, *41*, 23–33.
195. Krishnamurthy, R. Life's Biological Chemistry: A Destiny or Destination Starting from Prebiotic Chemistry? *Chemistry* **2018**, *24*, 16708–16715.
196. Szathmáry, E.; Maynard Smith, J. From replicators to reproducers: the first major transitions leading to life. *J. Theor. Biol.* **1997**, *187*, 555–571.
197. Eigen, M. Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **1971**, *58*, 465–523.
198. Walker, S.I.; Davies, P.C.W. The algorithmic origins of life. *J. R. Soc. Interface* **2013**, *10*, 20120869.
199. Segré, D.; Lancet, D. Composing life. *EMBO Rep.* **2000**, *1*, 217–222.
200. Hordijk, W.; Steel, M.; Kauffman, S.A. Molecular Diversity Required for the Formation of Autocatalytic Sets. *Life* **2019**, *9*.
201. Kostyrka, G. What roles for viruses in origin of life scenarios? *Stud. Hist. Philos. Biol. Biomed. Sci.* **2016**, *59*, 135–144.
202. Stewart, J.E. The Origins of Life: The Managed-Metabolism Hypothesis. *Found. Sci.* **2019**, *24*, 171–195.
203. Vasas, V.; Szathmáry, E.; Santos, M. Lack of evolvability in self-sustaining autocatalytic networks constraints metabolism-first scenarios for the origin of life. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 1470–1475.
204. Sharma, V.; Annila, A. Natural process–Natural selection. *Biophys. Chem.* **2007**, *127*, 123–128.
205. Meléndez-Hevia, E.; Montero-Gómez, N.; Montero, F. From prebiotic chemistry to cellular metabolism—The chemical evolution of metabolism before Darwinian natural selection. *J. Theor. Biol.* **2008**, *252*, 505–519.
206. Frenkel-Pinter, M.; Haynes, J.W.; C, M.; Petrov, A.S.; Burcar, B.T.; Krishnamurthy, R.; Hud, N.V.; Leman, L.J.; Williams, L.D. Selective incorporation of proteinaceous over nonproteinaceous cationic amino acids in model prebiotic oligomerization reactions. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 16338–16346.
207. Hud, N.V.; Cafferty, B.J.; Krishnamurthy, R.; Williams, L.D. The Origin of RNA and "My Grandfather's Axe." *Chem. Biol.* **2013**, *20*, 466–474.
208. Joyce, G.F. Molecular evolution: booting up life. *Nature* **2002**, *420*, 278–279.
209. Calvin, M. Mineral Origins of Life Genetic Takeover and the Mineral Origins of Life A. G. Cairns-Smith. *BioScience* **1983**, *33*, 596–596.

210. Schmitt-Kopplin, P.; Hemmler, D.; Moritz, F.; Gougeon, R.D.; Lucio, M.; Meringer, M.; Müller, C.; Harir, M.; Hertkorn, N. Systems chemical analytics: introduction to the challenges of chemical complexity analysis. *Faraday Discuss.* **2019**, *218*, 9–28.
211. Geisberger, T.; Diederich, P.; Steiner, T.; Eisenreich, W.; Schmitt-Kopplin, P.; Huber, C. Evolutionary Steps in the Analytics of Primordial Metabolic Evolution. *Life* **2019**, *9*, 50.
212. Richert, C. Prebiotic chemistry and human intervention. *Nat. Commun.* **2018**, *9*, 5177.

II Biological CO_2 fixation

The autotrophic conversion of inorganic carbon to organic carbon is crucial for the development of life because it directly links the abiotic with the biotic world (Fuchs, 2011). In modern ecosystems, the most widespread biological CO_2 fixation pathways out of the six known is the reductive pentose phosphate (Calvin-Benson) cycle, converting CO_2 into sugar phosphates. It is found in most photosynthetic organisms, from prokaryotes to plants (Bassham, 1979). Among the remaining five autotrophic pathways that can turn inorganic into organic material (Fuchs, 2011), most central within this thesis is the acetyl-CoA (Wood-Ljungdahl) pathway. This is for several reasons, starting with the fact that the acetyl-CoA pathway traces back to the last universal common ancestor—LUCA (Weiss *et al.*, 2016). This pathway uses electrons and energy from hydrogen (H_2) and can simultaneously supply three key requirements for life: reduced carbon in the form of acetyl groups, electrons in the form of reduced ferredoxin and NADH, and ion gradients for energy conservation in the form of ATP (Fuchs, 2011; Müller, Chowdhury and Basen, 2018; Preiner, Igarashi, *et al.*, 2019). The pathway is linear, not cyclic as the other five pathways, a sort of one-way street into metabolism; it releases more energy than it consumes; and the enzymes involved contain metal cofactors that could be of primordial origin (Ragsdale and Pierce, 2008; Sousa and Martin, 2014).

The three publications presented in this chapter address the acetyl-CoA pathway and its connection to geochemical processes (Publication 3), its relevance for the physiology of LUCA (Publication 4), and its mechanistic peculiarities in comparison with other CO_2 fixation pathways (Publication 5) that provide clues about its antiquity and possibly about its origin.

Publication 3

Title:	Native metals, electron bifurcation, and CO_2 reduction in early biochemical evolution
Year:	2018
Authors:	Filipa L. Sousa, Martina Preiner, and William F. Martin
Published in:	*Current Opinion in Microbiology*, 43, pp. 77–83. doi: 10.1016/j.mib.2017.12.010.
Contribution:	Second author. Medium: provided references and edits. Inspired the content of substantial parts of the text based on insights from the laboratory work performed.
Summary:	This essay discusses the parallels of biochemical and geochemical CO_2 fixation with focus on the role of zero-valent metals and metal atoms. The hypothesis of such metals being the primary electron source of abiotic CO_2 reduction instead of hydrogen is presented.

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Microbiology

Native metals, electron bifurcation, and CO_2 reduction in early biochemical evolution

Filipa L Sousa[1], Martina Preiner[2] and William F Martin[2,3]

Molecular hydrogen is an ancient source of energy and electrons. Anaerobic autotrophs that harness the H_2/CO_2 redox couple harbour ancient biochemical traits that trace back to the universal common ancestor. Aspects of their physiology, including the abundance of transition metals, radical reaction mechanisms, and their main exergonic bioenergetic reactions, forge links between ancient microbes and geochemical reactions at hydrothermal vents. The midpoint potential of H_2 however requires anaerobes that reduce CO_2 with H_2 to use flavin based electron bifurcation — a mechanism to conserve energy as low potential reduced ferredoxins via soluble proteins — for CO_2 fixation. This presents a paradox. At the onset of biochemical evolution, before there were proteins, how was CO_2 reduced using H_2? FeS minerals alone are probably not the solution, because biological CO_2 reduction is a two electron reaction. Physiology can provide clues. Some acetogens and some methanogens can grow using native iron (Fe^0) instead of H_2 as the electron donor. In the laboratory, Fe^0 efficiently reduces CO_2 to acetate and methanol. Hydrothermal vents harbour awaruite, Ni_3Fe, a natural compound of native metals. Native metals might have been the precursors of electron bifurcation in biochemical evolution.

Addresses
[1] Division of Archaea Biology and Ecogenomics, Department of Ecogenomics and Systems Biology, University of Vienna, Althanstrasse 14 UZA I, 1090 Vienna, Austria
[2] Institute of Molecular Evolution, Heinrich-Heine-University, Universitätsstr. 1, 40225 Düsseldorf, Germany
[3] Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, 2780-157 Oeiras, Portugal

Corresponding author: Sousa, Filipa L (filipa.sousa@univie.ac.at)

Current Opinion in Microbiology 2018, **43**:77–83

This review comes from a themed issue on **Environmental microbiology**

Edited by **Alberto Scoma** and **Julia Vorholt**

https://doi.org/10.1016/j.mib.2017.12.010

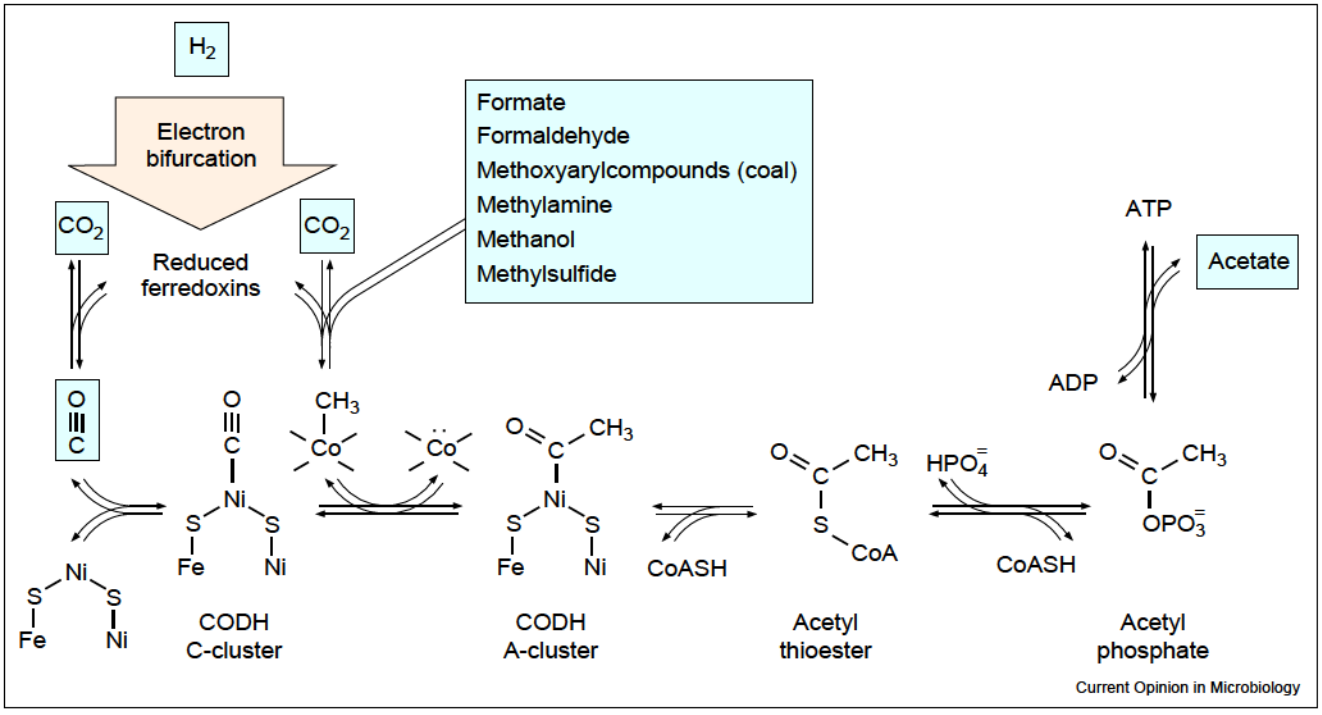
Introduction

The only thing we know for sure about life's origin some 3.95 billion years ago [1] is that energy was required. Without energy release, no chemical reactions can take place that could ultimately lead to complex chemicals, metabolism or primitive ecosystems [2•]. Two sources of energy at origins are mainly discussed: ultraviolet (UV) light emitted from the sun [3,4] and chemical energy at hydrothermal vents [3,4]. Although UV light can be conveniently integrated into elegant laboratory syntheses of organic molecules at low temperatures and pressures [5,6], it does not connect to the microbial world (life) because no known form of microbial physiology is powered by UV light. From the biological standpoint, chemical energy at hydrothermal vents, in particular the H_2/CO_2 redox couple, is interesting as a source of energy at origins. Why?

First, many forms of microbes use H_2 as a source of chemical energy for ATP synthesis in conjunction with a suitable electron acceptor such as CO_2 [7,8], and H_2-dependent autotrophs provided the initial primary production that supported the first heterotrophic metabolisms [9]. Second, the Earth's crust has been generating large amounts of H_2 since there was liquid water, through a process called serpentinization [10]. In addition, biologists have long held that anaerobic autotrophs that reduce CO_2 using electrons from H_2 harbour the most ancient forms of metabolism [11–13]. H_2 dependent anaerobic autotrophs are furthermore rich in transition metal catalysts such as Fe and Ni [13], traits long regarded as ancient, and hydrogenases that extract the electron pair from H_2 to provide reduction equivalents and energy that drive metabolism forward [14].

Among the kinds of carbon and energy metabolism known among modern microbes, the acetyl-CoA pathway (or Wood–Ljungdahl pathway) as it is used in acetogens (bacteria) and methanogens (archaea) appears to be the most ancient [12]. It is the only exergonic pathway of biological CO_2 fixation known [12], all others require energy input in the form of ATP. The intermediate product of the pathway is a thioester, reactive compounds that have long been thought to play an important role in early chemical evolution [15]. In bacteria, the acetyl-CoA pathway generates acetyl phosphate from H_2, CO_2 and phosphate (Figure 1). Acetyl phosphate is an excellent source of phosphorylation potential with a free energy of hydrolysis of −43 kJ/mol, 30% better than ATP. Even to

Figure 1



Scheme of energy conservation from ferredoxin to ATP in the acetyl CoA pathway. Redrawn after [12] incorporating intermediates in CO generation [13] and of substrate level phosphorylation [11]. Methanogenic growth on methoxy groups from coal was recently reported by [38•]. The beige arrow indicates the requirement for electron bifurcation in the synthesis of low potential reduced ferredoxins with electrons from H_2 [7]. Insights into the mechanisms of electron bifurcation were recently revealed by the structures of two bifurcating enzymes [25•,26•]. Substrates and endproducts of the reversible reaction sequence are boxed in blue. The synthesis of the methyl group from H_2 and CO_2 entails energy investment, such that net synthesis of acyl phosphate or ATP from H_2 requires chemiosmotic coupling [6–8].

the level of the energy rich thioester the reaction is exergonic:

$$2CO_2 + 4H_2 + CoASH \rightarrow CH_3COSCoA + 3H_2O$$

$$\Delta G_o' = -59\,kJ/mol[12].$$

Microbial genomes also point to the antiquity of anaerobic autotroph physiology. A recent study identified 355 ancient gene families that, based on their phylogenies, provided insights into the physiology and habitat of the last universal common ancestor, LUCA [16]. LUCA lived in a hot environment rich in gasses (H_2, CO_2, N_2, CO) and metals, a habitat very similar to hydrothermal vents, which existed on the early Earth [10]. It used the acetyl-CoA pathway [16], its metabolism was rich in FeS and thioester dependent reactions, which are enriched in the ancient ATP-independent core of metabolism [17••]. It harboured a diversified small molecule chemistry [18], many radical-based reactions that are dependent upon S-adenosyl methionine (SAM), which can form spontaneously *in vitro* [19•], and it harboured features found in microbes that today still inhabit ancient geochemical niches [20,21].

The problem with H_2 and FeS

Despite its chemical simplicity, its abundance in ancient environments, and the clear tendency of H_2-dependent autotrophs to branch deeply in phylogenetic trees [16], H_2 has a rather severe Achilles' heel as a source of electrons for CO_2 reduction at origins. Its midpoint potential is unfavourable for CO_2 reduction beyond the near-equilibrium reaction with formate [22]. That is why microbes that reduce CO_2 with H_2 employ flavin based electron bifurcation [23], a biochemical mechanism that generates reduced ferredoxins (Fd^-) with a midpoint potential on the order of –500 mV from H_2 with a more positive midpoint potential of only –414 mV [7]. Electrons from H_2 have to flow energetically uphill to low potential Fd. That might appear to violate the second law of thermodynamics, but electron bifurcation obeys the law in that one electron from H_2 is transferred energetically downhill to a high potential acceptor like NAD^+ or heterodisulfide CoB–S–S–CoM [24], while the other is transferred uphill to Fd so that the overall energetics of the reaction are favourable [8].

Why is that problematic? The problem with H_2 in an early evolution context is that flavin based electron bifurcation

is an elaborate physiological process that requires sophisticated proteins [25•,26•] working in concert with other proteins [7] as an energy metabolic pathway to reduce Fd (for reducing CO_2). This presents a familiar chicken-and-egg type paradox, namely how was CO_2 reduced with H_2 before there were proteins to catalyze electron bifurcation? One might counter that if early CO_2 fixation took place at hydrothermal vents, then electron bifurcation was not needed, because there was plenty of natural FeS around that could do the job of Fd^- when it comes to fixing CO_2.

But FeS minerals do not really solve the problem, because FeS has its own, different issues. FeS clusters in proteins are one electron donors, with iron undergoing Fe^{2+} to Fe^{3+} valence changes. The steps of biological CO_2 reduction in autotrophs are always two electron reactions [12]. In biology, the electrons from Fd^- are donated to C in CO_2 via metals that readily undergo two electron reactions, such as Ni, Mo, or W atoms coordinated in proteins or cofactors (Figure 2), or electron pairs are donated via hydride transfer from organic cofactors like NAD(P)H [12].

Huber and Wächtershäuser [27] obtained excellent yields (40 mol%) of the thioester methyl thioacetate from CH_3SH and CO using FeS, Ni^{2+} and Fe^{2+} salts, emulating the central anabolic reaction of the acetyl CoA pathway, but no one has reported genuine success involving FeS or other Fe^{2+} species as a reductant for CO_2 in an early evolution context, electrochemical experiments where external voltage is applied aside [28]. Could it be that in the beginning, CO_2 was not reduced directly by H_2 at all? What does nature say?

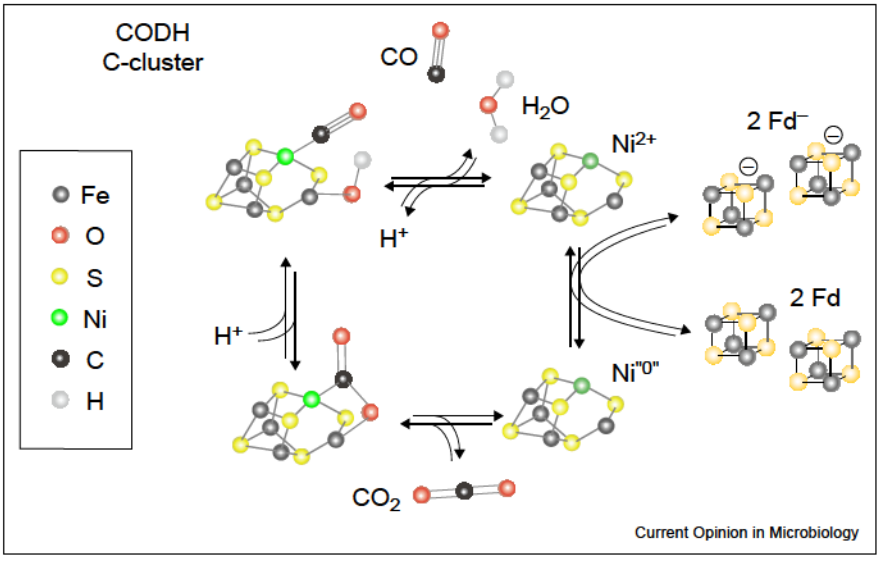
Geochemical CO_2 reduction

Modern hydrothermal systems provide a check for CO_2 reduction in that CH_4 (ca. 1 mM) [29] and other small organic compounds including formate are present in hydrothermal effluent [30–32]. Those small organic compounds are not synthesized where vent effluent discharges into the ocean on the sea floor, rather, the organics are apparently made deep in the crust, as organics in terrestrial vents attest [33]. CO_2 reduction in the modern crust is still not well understood in its details, although it is thought to stem from the same geochemical process that generates H_2 in hydrothermal effluent: serpentinization [30–32].

During serpentinization, H_2 synthesis stems from the reduction of water in hydrothermal systems via the oxidation of Fe^{2+} that is present in the Earth's crust in vast amounts as iron magnesium silicates [10]. Fresh iron silicates for serpentinization are continuously supplied anew at spreading zones such as the mid-Atlantic ridge, where crust emerges as magma that subsequently cools, ultimately being recycled back into the mantle at subduction zones [34]. The isotope signatures of CH_4 emergent from hydrothermal systems are distinct from that of marine CO_2 [29,32], the reasons for which are unclear. The rates of methane synthesis in laboratory scale serpentinization reactions so far are generally very slow [35•]. Thus, carbon is being reduced in serpentinizing systems, by a yet unidentified mechanism.

The process of serpentinization has been going on for the last 4.2 billion years, since there was liquid water on Earth [10]. Both its basic ingredients — water, Fe(II)-rich rocks,

Figure 2



Reaction mechanism of CO_2 reduction in bifunctional CO dehydrogenase acetyl-CoA synthase (CODH/ACS) synthase in the acetyl-CoA pathway, proposed by Ragsdale [13], modified from [13]. C-cluster refers to the NiFeS cluster at the active site of the CODH enzyme [13].

Box 1 The early Earth in a nutshell

There is broad agreement among geochemists and planetary researchers that the early Earth was molten at some point, by the moon forming impact roughly 4.4 billion years ago, at the latest [10,33]. On the molten Earth (>1500 K), carbon that had been brought to the freshly accreted planet was converted to CO_2, almost all of which was outgassed into the atmosphere, a small fraction being retained in magma oceans [10,33]. By about 4.4 Ga the magma oceans had cooled [33], by about 4.2 Ga there was liquid water on Earth [10], some outgassed from accretion and some delivered later by comets, and by around 4 Ga, perhaps as late as 3.7 Ga, the late heavy bombardment had come to an end [33]. By 3.95 Ga a carbon isotopic signature compatible with that produced by the acetyl-CoA pathway appeared [1].

As it relates to the source of energy for CO_2 reduction, the relevant sequence of events is this: magma oceans in the molten phase oxidized Earth's early carbon to atmospheric CO_2 and small amounts of mantle CO_2. Because it cooled from magma, the primordial crust was depleted in water, consisted mostly of iron magnesium silicates with very low water content [10,55]. As the crust cooled, water condensed to surface oceans. Gravity pulled water into cracks in the solid crust, creating convective currents — the process of serpentinization set in. Serpentinization drew very CO_2-rich water into the crust where serpentinization took place, such that H_2 was synthesized in an otherwise H_2-free environment and at sites where CO_2 existed in hydrothermal downcurrent water and as bound CO_2 (carbonates) in a generally dry crust. The primordial interaction between H_2 and CO_2 thus probably took place deep in the crust, not at sites where hydrothermal vent effluent reached the ocean floor in contrast to earlier views [6]; H_2 and CO_2 first interacted in the presence of vast amounts of dry rock, and at temperatures likely exceeding 100 °C.

and heat for convection — and the chemical reaction are simple, as such the process appears to be occurring on the Saturn moon Enceladus as well [36•]. Serpentinization is a spontaneous process that releases chemical energy, a notable property that it shares with microbial energy metabolism. If serpentinization releases chemical energy, where does the energy that serpentinization releases come from? A look at early Earth history is instructive: the energy comes from the molten state of the early Earth (magma oxidizes carbon to CO_2) and subsequent rock–water interactions, which then generate H_2 in the presence of CO_2 (Box 1).

Metagenomics tells us that approximately half of today's biomass lives in the crust in rocky, H_2 rich environments [37] and that substantial components of the modern subsurface biomass lives from the H_2/CO_2 redox couple as acetogens and methanogens [38•], which fuel subsurface primary production. Today, microbes in the crust can also grow from organic carbon deposits such as coal [39••], but that was not an option 4 billion years ago. Another main difference between today's crust and the primordial crust is that about half the water on Earth (i.e., roughly the volume of the ocean) is bound in the modern crust (and mantle), brought there by submarine hydrothermal activity [34]. The flipside of that coin is that the primordial oceans were twice as deep as today's [34], meaning that land for warm little ponds was probably in short supply.

Whence electrons, if not H_2 or FeS?

Serpentinization alters the rocks that host hydrothermal activity. A very notable component of hydrothermally altered rocks is the mineral awaruite. Awaruite is an intermetallic compound with the formula Ni_3Fe (or $Ni_{2-3}Fe$): native transition metals with oxidation state zero. It is a normal constituent of serpentinizing hydrothermal systems [35•,40], formed there naturally during serpentinization under conditions where high H_2 activities of ~200 mmol/kg [41] reduce the divalent metal ions.

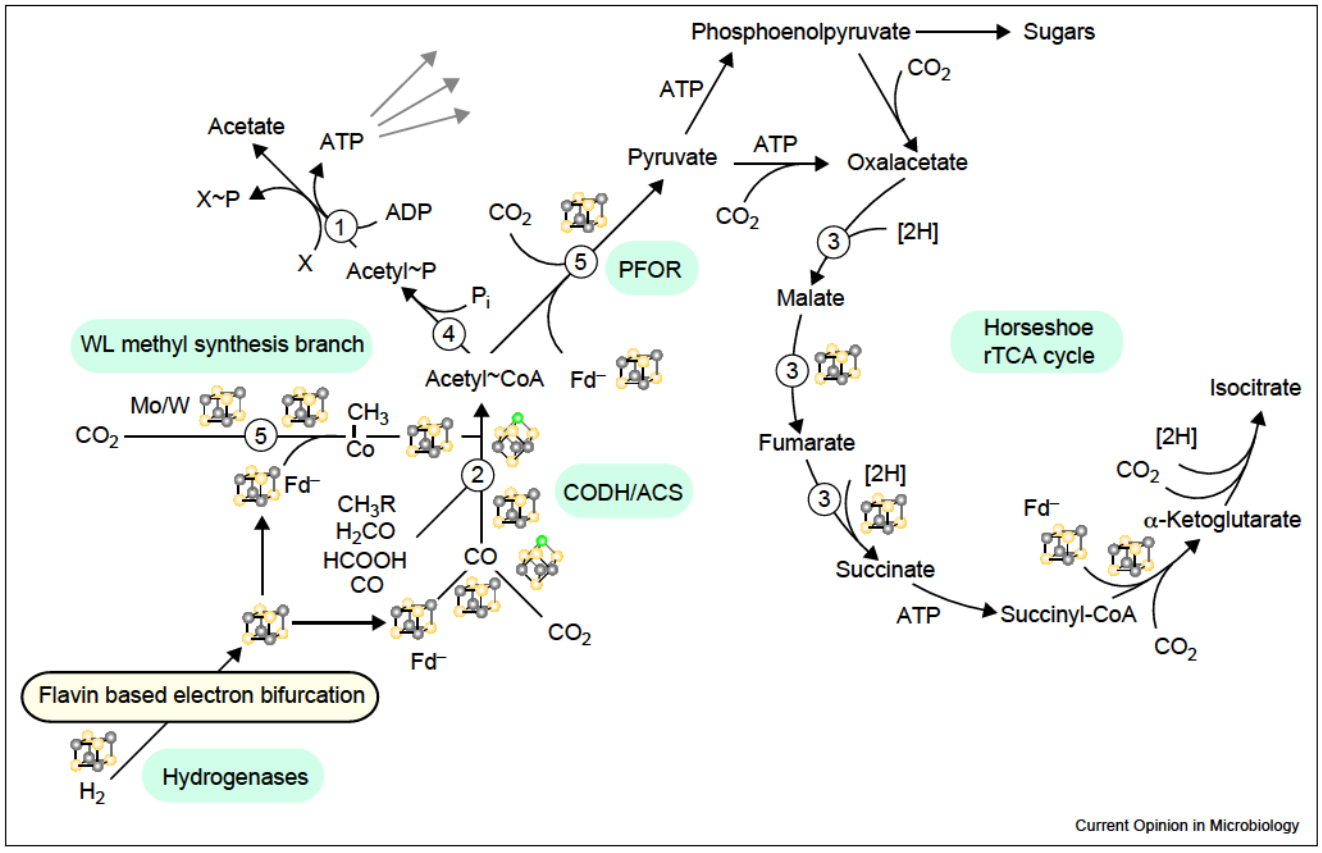
Almost 20 years ago Horita and Berndt [42] reported that Ni_3Fe could catalyze the synthesis of methane in mmol/kg amounts from H_2 and CO_2 [42] in simulated hydrothermal conditions (200–400 °C, ca. 50 MPa) although it cannot be excluded that awaruite was the reductant rather than the catalyst, at least in part. They also reported that after 1–2 weeks at lower temperature (200 °C) carbon compounds of oxidation state intermediate between CO_2 and CH_4 were obtained in amounts similar to or exceeding those observed for CH_4. Guan *et al.* [43] showed that Fe^0 in the presence of salts will reduce CO_2 to CH_4, C_3H_8, CH_3OH and C_2H_5OH in the 10–70 μM range at room temperature. He *et al.* [44] reported reduction of CO_2 to formate and acetate in the 1–10 mM range using nanoparticular Fe^0 at 80–200 °C. Moreover, Muchowska *et al.* [45••] recently showed native iron to accelerate and promote reactions of the reverse citric acid cycle (Figure 3). The message is that native metals can efficiently reduce CO_2. That is not the case for either H_2 alone or for FeS minerals. The native metals are interesting.

Awaruite is today synthesized where H_2 is produced and where organic compounds are being made. Are native metals involved in geochemical CO_2 reduction, and were they involved in primordial CO_2 reduction? Considering the mechanism proposed by Steve Ragsdale [13] for acetyl-CoA synthesis in acetogens (Figure 2), the active Ni species for CO_2 reduction in the (4 billion year old) biochemical reaction is formally Ni^0. The electrons are delivered to the enzyme one at a time via FeS clusters in Fd^-, but are delivered to carbon as a pair. The Fe^{2+} species in FeS clusters perform one electron chemistry, but the two electron carbon reduction reaction is performed by the transiently native metal. This might be a clue about ancient life.

Microbes always have the last word

What do anaerobic microbes say about native metals? Basically they say 'yes, please', and they are the source of much corrosion to things made of steel, as recently surveyed by Enning and Garrelfs [46]. Early reports showed that methanogens grow from Fe^0 and CO_2 [47], but questions remained whether the growth was really from iron or just from H_2 generated by interaction of iron with water. Dinh *et al.* [48] then showed that methanogens grew rapidly on native iron as an electron

Figure 3



The acetyl-CoA pathway and the incomplete reverse citric acid cycle as the core carbon end energy metabolism in some modern microbes and the first microbes, modified from Ref. [6] to include the energetic impediment imposed by flavin based electron bifurcation in ferredoxin reduction with electrons from H_2 [7] and updated to include new information. Participation of FeS or FeNiS clusters is indicated using database information for the corresponding enzymes indicated in Fuchs [12]. Note the paucity of ATP-dependent steps and the acyl phosphate/ATP-generating steps, indicated with ATP next to the reaction (see also Ref. [17••]). Circled numbers at reactions indicate that the non-enzymatic laboratory reaction has been reported as follows. (1) Kitani *et al.* reported that ATP is readily generated from acetyl phosphate and ADP using only Fe^{3+} [57] or Fe^{2+} [58] as catalysts. (2) Wächtershäuser and Huber reported the divalent metal ion-catalyzed synthesis of acetyl thioesters from CH_3SH and CO [26•]. (3) Muchowska et al. showed these steps to proceed spontaneously *in vitro* using native metal and metal ion catalysts [45••]. (4) Weber [56] reported the synthesis of pyrophosphate from an acetylthioester and P_i, whereby the presence of acetyl phosphate was inferred but not directly shown [6]. (5) Varma *et al.* [54••] recently demonstrated synthesis of metal bound methyl groups and synthesis of pyruvate from CO_2 with native metals as the source of electrons. The free energy of hydrolysis for acetyl phosphate (–43 kJ/mol) is greater than that for ATP (–31 kJ/mol) [6]. In the context of early biochemical evolution, reactions that are today coupled to ATP hydrolysis would be thermodynamically even more favourable if coupled to acetyl phosphate (or other acyl phosphate) hydrolysis, an argument in favour of acetyl phosphate as a primitive energy currency [6,59]. The involvement of reductants other than Fd^- in the enzymatic reaction is indicated with [2H]. *Abbreviations*: CODH/ACS, carbon monoxide dehydrogenase/acetyl-CoA synthase; WL, Wood–Ljungdahl; PFOR, pyruvate ferredoxin oxidoreductase.

source, suggesting that H_2 was not involved as an intermediate. More recently, Tan *et al.* [49] showed that *Methanosarcina barkeri* will grow on Fe^0 as an electron source, but only when deprived of standard electron donors such as CH_3OH or H_2.

The molecular mechanisms of microbial electron extraction from Fe^0 are so far elusive. Lohner *et al.* [50] reported Fe^0 oxidation by methanogens in the presence of externally applied voltage. The results suggested that methanogens can access electrons from Fe^0 via routes that do not involve H_2. More recent results by Deutzmann and colleagues using applied potentials suggest that methanogens might acquire electrons from Fe^0 via extracellular enzymes that oxidize the native metal to standard electron sources such as H_2 or formate [51,52]. There are many reports about anaerobic growth on iron, many involving methanogens or sulfate reducers [46]. But acetogens?

Kato *et al.* [53•] recently isolated acetogens from the genus *Sporomusa* (firmicutes) that grow on native iron without externally applied potentials. Most acetogens that Kato *et al.* [53•] tested do not grow on iron as the electron source, indicating the presence of genetically specified mechanisms to access electrons from Fe^0, in line

with the conclusions of Dinh *et al.* [48] for sulfate reducers and methanogens. For sulfate reducers, the terminal acceptor of electrons from Fe^0 can be a sulfur compound or CO_2, for acetogens and methanogens, the terminal acceptor is CO_2. The proteins, cofactors and mechanisms involved in Fe^0 oxidation are still unknown [46,52,53•]. It is possible that non-enzymatic reactions of Fe^0 with CO_2 such as those generating formate, methanol, and acetate in the laboratory [43,44] play a role in microbial growth on iron. Very recent work reports the non-enzymatic synthesis of pyruvate from aqueous CO_2 and Fe^0 under mild hydrothermal conditions [54••].

Conclusions

As usual, nature leaves us with observations and questions. Is the oxidation of native iron (and other metals) an ancient trait, preserved from the very earliest phases of biological CO_2 reduction and is it prevalent in hydrothermal vents, where Ni_3Fe is still made today? The two electron iron oxidation reaction $Fe^0 \rightarrow Fe^{2+} + 2e^-$ has a midpoint potential of $E_o' = -470$ mV [53•], more negative than that for hydrogen $H_2 \rightarrow 2H^+ + 2e^-$, with $E_o' = -410$ mV. Might anaerobic autotrophs that oxidize Fe^0 short circuit flavin based electron bifurcation to generate their low potential reduced ferredoxins? If so, they still would have to direct two electron (Fe^0) to one electron (FeS clusters in Fd^-) reactions. The physiological reactions by which microbes access Fe^0 and other native metals as electron sources might uncover hints about early life, possibly even probing a phase of physiological evolution before there was genetically encoded electron bifurcation.

Acknowledgements

We thank Joseph Moran, Carlo Kleinermanns, Harun Tüysüz, and Terry Bristol for helpful discussions, Olivia Judson for critical reading, the ERC (666053 to WFM), the VW Foundation (930 46 to WFM), and the WWTF (VRG15-007 to FLS) for funding.

References and recommended reading

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

- • of special interest
- •• of outstanding interest

1. Tashiro T, Ishida A, Hori M, Igisu M, Koike M, Méjean P, Takahata N, Sano Y, Komiya T: **Early trace of life from 3.95 Ga sedimentary rocks in Labrador, Canada**. *Nature* 2017, **549**:516-518.

2. • Judson OP: **The energy expansions of evolution**. *Nat Ecol Evol* 2017, **1** 0138.
A concise and interesting overview of how life accessed new sources of energy during evolution.

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

4. Bonfio C, Valer L, Scintilla S, Shah S, Evans DJ, Jin L, Szostak JW, Sasselov DD, Sutherland JD, Mansy SS: **UV-light-driven prebiotic synthesis of iron–sulfur clusters**. *Nat Chem* 2017, **60**:2817.

5. Baross JA, Hoffman SE: **Submarine hydrothermal vents and associated gradient environments as sites for the origin and evolution of life**. *Orig Life Evol Biosph* 1985, **15**:327-345.

6. Martin W, Russell MJ: **On the origin of biochemistry at an alkaline hydrothermal vent**. *Philos Trans R Soc Lond B: Biol Sci* 2007, **362**:1887-1925.

7. Basen M, Müller V: **"Hot" acetogenesis**. *Extremophiles* 2017, **21**:15-26.

8. Buckel W, Thauer RK: **Energy conservation via electron bifurcating ferredoxin reduction and proton/Na(+) translocating ferredoxin oxidation**. *Biochim Biophys Acta* 2013, **1827**:94-113.

9. Schönheit P, Buckel W, Martin W: **On the origin of heterotrophy**. *Trends Microbiol* 2016, **24**:12-25.

10. Sleep NH, Bird DK, Pope EC: **Serpentinite and the dawn of life**. *Philos Trans R Soc Lond B: Biol Sci* 2011, **366**:2857-2869.

11. Decker K, Jungermann K, Thauer RK: **Energy production in anaerobic organisms**. *Angew Chem Int Ed Engl* 1970, **9**:138-158.

12. Fuchs G: **Alternative pathways of carbon dioxide fixation: insights into the early evolution of life?** *Annu Rev Microbiol* 2011, **65**:631-658.

13. Ragsdale SW: **Nickel-based enzyme systems**. *J Biol Chem* 2009, **284**:18571-18575.

14. Søndergaard D, Pedersen CNS, Greening C: **HydDB: a web tool for hydrogenase classification and analysis**. *Sci Rep* 2016, **6**:34212.

15. Semenov SN, Kraft LJ, Ainla A, Zhao M, Baghbanzadeh M, Campbell VE, Kang K, Fox JM, Whitesides GM: **Autocatalytic, bistable, oscillatory networks of biologically relevant organic reactions**. *Nature* 2016, **537**:656-660.

16. Weiss MC, Sousa FL, Mrnjavac N, Neukirchen S, Roettger M, Nelson-Sathi S, Martin WF: **The physiology and habitat of the last universal common ancestor**. *Nat Microbiol* 2016, **1**:16116.

17. •• Goldford JE, Hartman H, Smith TF, Segrè D: **Remnants of an ancient metabolism without phosphate**. *Cell* 2017, **168**:1126-1134.
A comprehensive investigation of biochemical pathways that uncovers an ancient core of ATP-independent and phosphate-independent reactions in modern metabolism.

18. Michael AJ: **Evolution of biosynthetic diversity**. *Biochem J* 2017, **474**:2277-2299.

19. • Laurino P, Tawfik DS: **Spontaneous emergence of S-adenosylmethionine and the evolution of methylation**. *Angew Chem Int Ed Engl* 2017, **56**:343-345.
Shows that S-adenosylmethionine (SAM), an ancient and conserved cofactor that is central to the biosynthesis of other cofactors, can form from simpler components without enzymes.

20. Giovannelli D, Sievert SM, Hügler M, Markert S, Becher D, Schweder T, Vetriani C: **Insight into the evolution of microbial metabolism from the deep-branching bacterium, *Thermovibrio ammonificans***. *eLife* 2017, **6**:2050.

21. Nayak DD, Mahanta N, Mitchell DA, Metcalf WW: **Post-translational thioamidation of methyl-coenzyme M reductase, a key enzyme in methanogenic and methanotrophic Archaea**. *eLife* 2017, **6**:108.

22. Schuchmann K, Müller V: **Direct and reversible hydrogenation of CO_2 to formate by a bacterial carbon dioxide reductase**. *Science* 2013, **342**:1382-1385.

23. Herrmann G, Jayamani E, Mai G, Buckel W: **Energy conservation via electron-transferring flavoprotein in anaerobic bacteria**. *J Bacteriol* 2008, **190**:784-791.

24. Kaster A-K, Moll J, Parey K, Thauer RK: **Coupling of ferredoxin and heterodisulfide reduction via electron bifurcation in hydrogenotrophic methanogenic archaea**. *Proc Natl Acad Sci USA* 2011, **108**:2981-2986.

25. • Lubner CE, Jennings DP, Mulder DW, Schut GJ, Zadvornyy OA, Hoben JP, Tokmina-Lukaszewska M, Berry L, Nguyen DM,

Lipscomb GL *et al.*: **Mechanistic insights into energy conservation by flavin-based electron bifurcation**. *Nat Chem Biol* 2017, **13**:655-659.
Provides important insights into the mechanistic basis of flavin based electron bifurcation.

26. • Wagner T, Koch J, Ermler U, Shima S: **Methanogenic heterodisulfide reductase (HdrABC-MvhAGD) uses two noncubane [4Fe–4S] clusters for reduction**. *Science* 2017, **357**:699-703.
Provides important insights into the mechanistic basis of flavin based electron bifurcation.

27. Huber C, Wächtershäuser G: **Activated acetic acid by carbon fixation on (Fe, Ni)S under primordial conditions**. *Science* 1997, **276**:245-247.

28. Roldan A, Hollingsworth N, Roffey A, Islam H-U, Goodall JBM, Catlow CRA, Darr JA, Bras W, Sankar G, Holt KB *et al.*: **Bio-inspired CO_2 conversion by iron sulfide catalysts under sustainable conditions**. *Chem Commun (Camb)* 2015, **51**:7501-7504.

29. Proskurowski G, Lilley MD, Seewald JS, Früh-Green GL, Olson EJ, Lupton JE, Sylva SP, Kelley DS: **Abiogenic hydrocarbon production at lost city hydrothermal field**. *Science* 2008, **319**:604-607.

30. Schrenk MO, Brazelton WJ, Lang SQ: **Serpentinization, carbon, and deep life**. *Rev Miner Geochem* 2013, **75**:575-606.

31. McCollom TM, Seewald JS: **Serpentinites, hydrogen, and life**. *Elements* 2013, **9**:129-134.

32. McDermott JM, Seewald JS, German CR, Sylva SP: **Pathways for abiotic organic synthesis at submarine hydrothermal fields**. *Proc Natl Acad Sci USA* 2015, **112**:7668-7672.

33. Etiope G, Schoell M: **Abiotic gas: atypical, but not rare**. *Elements* 2014, **10**:291-296.

34. Zahnle K, Arndt N, Cockell C, Halliday A, Nisbet E, Selsis F, Sleep NH: **Emergence of a habitable planet**. *Geology and Habitability of Terrestrial Planets*. Springer; 2007:35-78.

35. • McCollom TM: **Abiotic methane formation during experimental serpentinization of olivine**. *Proc Natl Acad Sci USA* 2016, **113**:13965-13970.
An important investigation into the processes of serpentinization and its evolutionary significance.

36. • Waite JH, Glein CR, Perryman RS, Teolis BD, Magee BA, Miller G, Grimes J, Perry ME, Miller KE, Bouquet A *et al.*: **Cassini finds molecular hydrogen in the Enceladus plume: evidence for hydrothermal processes**. *Science* 2017, **356**:155-159.
Reports evidence for the existence of serpentinization on the Saturn moon Enceladus, which has a rocky core and liquid water.

37. Whitman WB, Coleman DC, Wiebe WJ: **Prokaryotes: the unseen majority**. *Proc Natl Acad Sci USA* 1998, **95**:6578-6583.

38. • Colman DR, Poudel S, Stamps BW, Boyd ES, Spear JR: **The deep, hot biosphere: twenty-five years of retrospection**. *Proc Natl Acad Sci USA* 2017, **114**:6895-6903.
A timely perspective on microbial communities in the deep biosphere and Thomas Gold's ideas on the origin of fossil fuels on Earth.

39. •• Mayumi D, Mochimaru H, Tamaki H, Yamamoto K, Yoshioka H, Suzuki Y, Kamagata Y, Sakata S: **Methane production from coal by a single methanogen**. *Science* 2016, **354**:222-225.
Reports growth of a methanogen on the arylmethoxygroups in coal.

40. Schwarzenbach EM, Gazel E, Caddick MJ: **Hydrothermal processes in partially serpentinized peridotites from Costa Rica: evidence from native copper and complex sulfide assemblages**. *Contrib Miner Petrol* 2014, **168**:1079.

41. Klein F, Bach W: **Fe–Ni–Co–O–S phase relations in peridotite–seawater interactions**. *J Petrol* 2009, **50**:37-59.

42. Horita J, Berndt ME: **Abiogenic methane formation and isotopic fractionation under hydrothermal conditions**. *Science* 1999, **285**:1055-1057.

43. Guan G, Kida T, Ma T, Kimura K, Abe E, Yoshida A: **Reduction of aqueous CO_2 at ambient temperature using zero-valent iron-based composites**. *Green Chem* 2003, **5**:630-634.

44. He C, Tian G, Liu Z, Feng S: **A mild hydrothermal route to fix carbon dioxide to simple carboxylic acids**. *Org Lett* 2010, **12**:649-651.

45. •• Muchowska KB, Varma SJ, Chevallot-Beroux E, Lethuillier-Karl L, Li G, Moran J: **Metals promote sequences of the reverse Krebs cycle**. *Nat Ecol Evol* 2017, **1**:1716-1721.
An important investigation showing that reactions of the citric acid cycle take place without enzymes, catalyzed by metals and metal ions.

46. Enning D, Garrelfs J: **Corrosion of iron by sulfate-reducing bacteria: new views of an old problem**. *Appl Env Microbiol* 2014, **80**:1226-1236.

47. Daniels L, Belay N, Rajagopal BS, Weimer PJ: **Bacterial methanogenesis and growth from CO_2 with elemental iron as the sole source of electrons**. *Science* 1987, **237**:509-511.

48. Dinh HT, Kuever J, Mußmann M, Hassel AW, Stratmann M, Widdel F: **Iron corrosion by novel anaerobic microorganisms**. *Nature* 2004, **427**:829-832.

49. Tan JL, Goh PC, Blackwood DJ: **Influence of H_2S-producing chemical species in culture medium and energy source starvation on carbon steel corrosion caused by methanogens**. *Corros Sci* 2017, **119**:102-111.

50. Lohner ST, Deutzmann JS, Logan BE, Leigh J, Spormann AM: **Hydrogenase-independent uptake and metabolism of electrons by the archaeon *Methanococcus maripaludis***. *ISME J* 2014, **8**:1673-1681.

51. Deutzmann JS, Spormann AM: **Enhanced microbial electrosynthesis by using defined co-cultures**. *ISME J* 2016, **11**:704-714.

52. Deutzmann JS, Sahinb M, Spormann AM: **Extracellular enzymes facilitate electron uptake in biocorrosion and bioelectrosynthesis**. *mBio* 2015, **6**:e00496-e515.

53. • Kato S, Yumoto I, Kamagata Y: **Isolation of acetogenic bacteria that induce biocorrosion by utilizing metallic iron as the sole electron donor**. *Appl Env Microbiol* 2015, **81**:67-73.
Reports growth of acetogens using native iron as the source of electrons for CO_2 reduction.

54. •• Varma SJ, Muchowska KB, Chatelain P, Moran J: **Metals enable a non-enzymatic acetyl CoA pathway**. *bioRxiv* 2017 http://dx.doi.org/10.1101/235523.
Reports the synthesis of formate, methanol, acetate, and pyruvate (10–100 micromolar range) from CO_2 and native metals under mild aqueous conditions.

55. Arndt N, Nisbet E: **Processes on the young Earth and the habitats of early life**. *Annu Rev Earth Planet Sci* 2012, **40**:521-549.

56. Weber AL: **Formation of pyrophosphate, tripolyphosphate, and phosphorylimidazole with the thioester, N,S diacetyl-cysteamine, as the condensing agent**. *J Mol Evol* 1981, **18**:24-29.

57. Kitani A, Tsunetsugu S, Suzuki A, Ito S, Sasaki K: **Fe(III)-ion-catalysed non-enzymatic transformation of ADP into ATP**. *Chem Soc Perkin Trans* 1991, **3**:323-329.

58. Kitani A, Tsunetsugu S, Suzuki A, Ito S, Sasaki K: **Fe(III)-ion-catalysed non-enzymatic transformation of adenosine diphosphate into adenosine triphosphate part II. Evidence of catalytic nature of Fe ions**. *Bioelectrochem Bioenerget* 1995, **36**:47-51.

59. Martin WF, Thauer RK: **Energy in ancient metabolism**. *Cell* 2017, **169**:953-955.

Publication 4

Title:	The last universal common ancestor between ancient earth chemistry and the onset of genetics
Year:	2018
Authors:	Madeline C. Weiss, Martina Preiner, Joana C. Xavier, Verena Zimorski, and William F. Martin
Published in:	*PloS Genetics,* 14(8), e1007518. doi: 10.1371/journal.pgen.1007518.
Contribution:	Second author. Minor: contributed detailed literature research, edited introductory paragraph on the geological setting on Early Earth and edited throughout the paper.
Summary:	This review provides a summary of research on the theoretical construct of the last universal common ancestor (LUCA) and its connection to alkaline hydrothermal vents as a possible location of OoL. Parallels between its reconstructed physiology and the physicochemical conditions on early Earth are drawn.

PLOS GENETICS

REVIEW

The last universal common ancestor between ancient Earth chemistry and the onset of genetics

Madeline C. Weiss[1], Martina Preiner[1], Joana C. Xavier[1], Verena Zimorski[1], William F. Martin[1,2]*

1 Institute of Molecular Evolution, Heinrich Heine University, Düsseldorf, Germany, **2** Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Oeiras, Portugal

* bill@hhu.de

Abstract

All known life forms trace back to a last universal common ancestor (LUCA) that witnessed the onset of Darwinian evolution. One can ask questions about LUCA in various ways, the most common way being to look for traits that are common to all cells, like ribosomes or the genetic code. With the availability of genomes, we can, however, also ask what genes are ancient by virtue of their phylogeny rather than by virtue of being universal. That approach, undertaken recently, leads to a different view of LUCA than we have had in the past, one that fits well with the harsh geochemical setting of early Earth and resembles the biology of prokaryotes that today inhabit the Earth's crust.

Citation: Weiss MC, Preiner M, Xavier JC, Zimorski V, Martin WF (2018) The last universal common ancestor between ancient Earth chemistry and the onset of genetics. PLoS Genet 14(8): e1007518. https://doi.org/10.1371/journal.pgen.1007518

Editor: Mark Achtman, Warwick Medical School, UNITED KINGDOM

Published: August 16, 2018

Funding: This work was supported by grants from the ERC (666053 https://erc.europa.eu/), Volkswagen Foundation (93 046 https://www.volkswagenstiftung.de/en/funding.html), and GIF ([I-1321-203.13/2015] http://www.gif.org.il/Pages/default.aspx) to WFM. The funders had no role in the preparation of the article.

Competing interests: The authors have declared that no competing interests exist.

Introduction

The very earliest phases of life on Earth witnessed the origin of life and genetics from the elements. There was a time when there was no life on Earth, and there was a time when there were DNA-inheriting cells. The transitions are hard to imagine. Some dates and constraints on the order of events helps us to better grasp the problem. The Earth is 4.5 billion years (Ga) old [1]. By about 4.4 Ga, the moon-forming impact turned the Earth into a ball of boiling lava [1]. Magma oceans with temperatures over 2,000˚K forced all water from early accretion into the gas phase and converted all early accreted carbon to atmospheric carbon dioxide (CO_2) [1,2]. By 4.2 to 4.3 Ga, the Earth had cooled sufficiently enough that there was liquid water [3] —those first oceans were about twice as deep as today's [1,2]. Only later, hydrothermal convection currents started sequestering water to the primordial crust and mantle, which today bind one extra ocean volume [4,5]. The first signs of life appear as carbon isotope signatures in rocks 3.95 billion years of age [6]. Thus, somewhere on the ocean-covered early Earth and in a narrow window of time of only about 200 million years, the first cells came into existence. Because the genetic code [7] and amino acid chirality [8] are universal, all modern life forms ultimately trace back to that phase of evolution. That was the time during which the last universal common ancestor (LUCA) of all cells lived.

LUCA, the tree of life, and its roots

LUCA is a theoretical construct—it might or might not have been something we today would call an organism. It helps to bridge the conceptual gap between rocks and water on the early Earth and ideas about the nature of the first cells. Thoughts about LUCA span decades. Various ideas exist in the literature about how LUCA was physically organized and what properties it possessed. These ideas are traditionally linked to our ideas about the overall tree of life and where its root might lie [9–18]. Phylogenetic trees are, however, ephemeral. It is their inescapable fate to undergo change as new data and new methods of phylogenetic inference emerge. Accordingly, the tree of life has been undergoing a great deal of change of late.

The familiar three-domain tree of life presented by ribosomal RNA [19] depicted LUCA as the last common ancestor of archaea, bacteria, and eukaryotes (Fig 1A). In that framework, efforts to infer the gene content, hence the properties of LUCA, boiled down to identifying genes that were present in eukaryotes, archaea, and bacteria. When the first genomes came out, there were a great many such investigations [20–22], all of which were confronted with the same two recurrent and fundamental problems: 1) How are the three domains related to one another so that gene presence patterns would really trace genes to LUCA as opposed to another evolutionarily more derived branch? 2) Does presence of a gene in two domains (or three) indicate that it was present in the common ancestor of those domains, or could it have reached its current distribution via late invention in one domain and lateral gene transfer (LGT) from one domain to another?

The first problem (the root of the domains) has been the subject of much recent work. Phylogenetic advances and new metagenomic data are changing the three-domain tree [19] into a two-domain tree [24,25]. This is partially a development around phylogenetic methods [24,26–28] but also entails new archaeal lineages that are now being assembled from metagenomic data and that appear to be more closely related to the host that acquired the mitochondrion than any other archaea known so far [29,30]. The two-domain tree showing an "archaeal origin of eukaryotes" [24,28] (Fig 1B) only tells part of the story, though, because eukaryote genomes harbor more bacterial genes than they do archaeal genes by a factor of about 3:1 [31–33], and those bacterial genes furthermore trace to the eukaryote common ancestor [23]. Eukaryotes are not just big, complex archaea; genomically and at the cellular level, they are true chimeras in that they possess archaeal ribosomes in the cytosol and bacterial ribosomes in mitochondria (Fig 1C) [34]. That polarizes cellular evolution in the right direction (there were once debates about eukaryotes being ancestral [10,13,14,22], as discussed elsewhere [35–37]) and identifies eukaryotes as latecomers in evolution, descendants of prokaryotes [38].

Current versions of the two-domain tree focus on the phylogeny of a handful of about 30 genes, mostly for ribosomal proteins (Box 1) but also on sequences from metagenomic samples. The metagenomic studies [29,30] have generated debate. Metagenomic data can bring forth alignments of genes that were sequenced accurately but have the wrong taxonomic label. For example, Da Cunha and colleagues [39] reported that published trees [29] hinge upon a strong signal stemming from one gene out of 30 and that the gene in question (an elongation factor [EF2]) might not be archaeal but eukaryotic instead. Spang and colleagues [40] defended their tree, eliciting more debate [41]. Errors can also occur in the assembly pipeline [42] en route to alignments [43], independent of contamination. Notwithstanding current debate about metagenomics-based trees of life [24,39,40,42,43], we should recall that rRNA itself produces the two-domain tree when various tree construction parameters are employed [24,26,27]. Both data and methods bear upon efforts to construct trees of life. It remains possible that some aspects of domain relationships might never be resolved to everyone's satisfaction—even the endosymbiotic origin of mitochondria is still debated [37]. But the bacterial

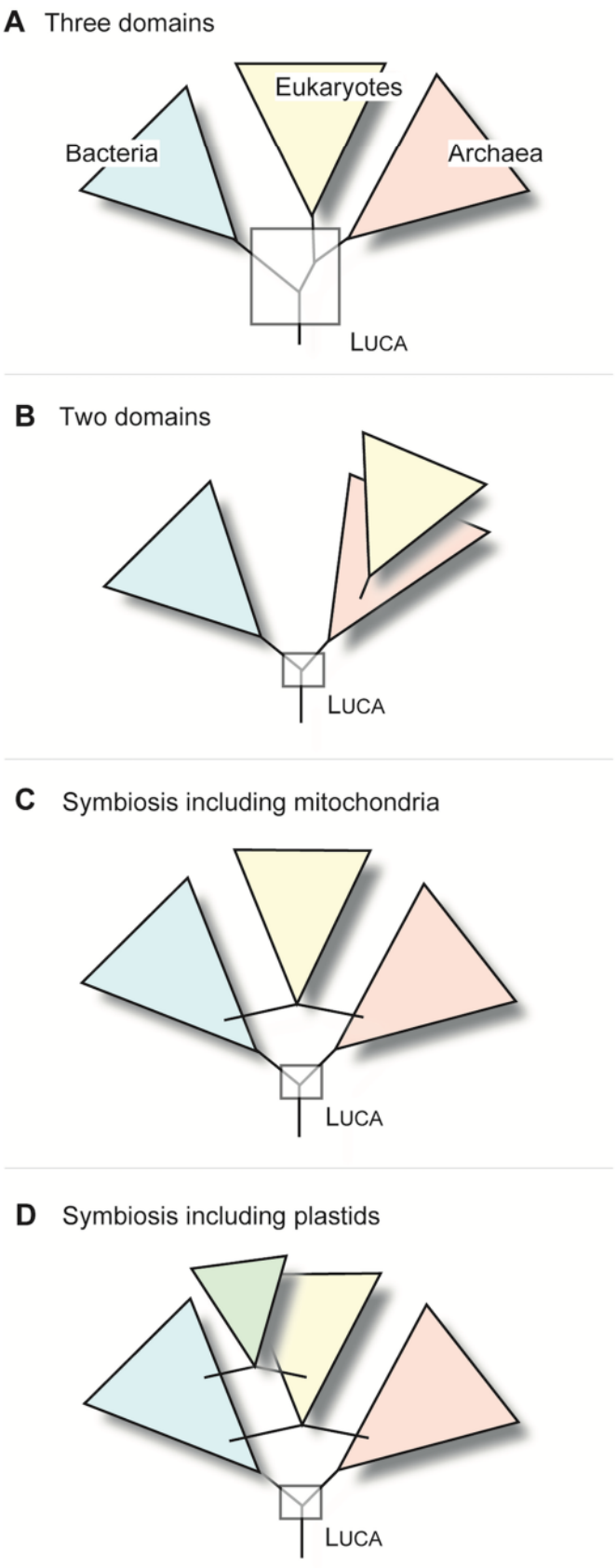

A Three domains
Eukaryotes
Bacteria
Archaea
LUCA
B Two domains
LUCA
C Symbiosis including mitochondria
LUCA
D Symbiosis including plastids
LUCA

Fig 1. Different views on domain relationships in the tree of life. (A) The three-domain tree: based on rRNA phylogeny, the three domains were of equal rank. (B) The two-domain tree: modern trees show eukaryote cytosolic ribosomes branching within the diversity of archaeal ribosomes. (C) As eukaryotes are not just grownup archaea, the eukaryote ancestor possessed mitochondria. If mitochondrial-derived genes are taken into account, the tree is no longer a bifurcating graph. (D) If plastids are included, the tree becomes even less tree-like because the photosynthetic lineages of eukaryotes also acquired many genes from the plastid ancestor [23].

https://doi.org/10.1371/journal.pgen.1007518.g001

origin of mitochondria and their presence in the eukaryote common ancestor [44–47], together with the tendency of eukaryotes to branch within archaeal lineages as archaeal lineage sampling [29,30,48] and phylogenetic methods [24,26,27,32] improve, indicates that eukaryotes arose from prokaryotes and that genes that trace to the common ancestor of archaea and bacteria trace to LUCA.

The second problem (how much LGT has there been between domains) that has impaired progress on LUCA has arguably been more difficult to resolve than the rooting issue. If a given gene is present in bacteria and archaea, was it present in LUCA, or could it have been transferred between domains via LGT? As one important example, early studies pondered the presence of bacterial type oxygen (O_2)-consuming respiratory chains in archaea [21]. Does that

Box 1. The tree of 1% and the tree of everything else

A traditional approach to LUCA has been to simply look for the genes that are present in all genomes. That is easy enough, but the results are sobering. What one finds is a collection of about 30 genes, mostly for ribosomal proteins, telling us that LUCA had a ribosome and had the genetic code, which we already knew [63–65]. That collection of about 30 genes has been in use for about 20 years as concatenated alignments to make trees of lineages based on larger amounts of data than rRNA sequences have to offer [66]. The genes that are present in all lineages (or nearly all) inform us about how LUCA translated mRNA into protein, but they do not tell us about how or where LUCA lived. That information concerns ecophysiology, and physiological traits are not universally conserved—they are what makes microbes different from one another. One can relax the criteria of universal presence a bit and allow for some gene loss in some lineages, in which case, one finds about 100 proteins that are nearly universal [67]. If one puts no size constraints on LUCA's genome and allows loss freely, then all genes present in at least one archaeon and one bacterium trace to LUCA, making it the most versatile organism that ever lived [51]. New insights about microbial phylogeny are emerging from concatenated alignments [24,29,30,42,48,68]. But one has to take care not to get genes from different lineages mixed up, which can be difficult when metagenomes are involved [39,43]. Furthermore, data concatenation has its own pitfalls [66,69,70]. Most modern concatenation studies [29,30,48] employ site-filtering methods in an attempt to remove "noise," but even sites that look "noise free" can still contain bias and conflicting data [63]. Another problem is that popular methods of phylogenetic inference produce inflated confidence intervals on phylogenies and branches [71]. Trees of ca. 30 concatenated proteins are no more immune to phylogenetic error than rRNA is and are prone to additional kinds of error [72]. As it relates to LUCA, regardless of the backbone tree, we still need to know what all proteins say individually about their own phylogenies.

mean that archaea are ancestrally O_2 consumers? As O_2 is the product of cyanobacterial photosynthesis [49] if we presume archaeal O_2 respiration to be an ancestral trait of archaea, it means that archaea arose after cyanobacteria, which are only about 2.5 billion years old and gave rise to plastids (Fig 1D) only about 1.5 billion years ago [50]. If ancestral archaea were oxygen respirers, and ancestral bacteria were too, suddenly neither the two-domain tree nor the three-domain tree (Fig 1) make sense because everything is upside down and rooted in cyanobacteria. Similar issues are encountered for many genes and traits [51]. Lateral gene transfer among prokaryotic domains helps to resolve such problems because it decouples physiology (ecological trait evolution) from phylogeny (ribosomal lineage evolution) [52], but it also makes genes more difficult to trace to LUCA.

Has lateral gene transfer obscured all records?

That takes us to the other extreme. If all genes have been subjected to LGT, as some early claims had it [53], then LUCA would be altogether unknowable from the standpoint of genomes. Early archaeal genomes did indeed uncover abundant transdomain LGT [54], and many bacteria to archaea transfers can be correlated to changes in physiology [55], including the transfer of O_2-consuming respiratory chains [55–58]. For reconstructing LUCA, the issue boils down to determining i) which genes are present in both archaea and bacteria, ii) which of those are present in both prokaryotic domains because of LGT between archaea and bacteria, and iii) which are present because of vertical inheritance from LUCA. For that, there are currently two methodological approaches. One involves making a backbone reference tree from universally conserved genes that are present in each genome—the tree of 1% [59] (see Box 1)—plotting all gene distributions on the tips of that tree, and then estimating which genes trace to LUCA on the basis of various assumed gain and loss parameters [60–62]. If we permit loss freely, many genes will trace back to LUCA; if we assume many gains, LUCA will have few genes [61]. Constraining ancestral genome sizes helps constrain estimates of which genes trace to LUCA [61] but only if we assume that the tree of each gene is compatible with the reference tree, which is a very severe assumption and unlikely to be true. Each gene has its own individual history (Box 1).

Each gene records its own evolutionary history

If any protein-coding genes have been vertically inherited from LUCA, their trees should reflect that. To find such trees, one has to make all trees for all proteins, meaning one has to make clusters for all protein-coding genes from large numbers (thousands) of sequenced genomes. Clusters correspond to "natural" protein families of shared amino acid sequence similarity. Given modern computers, making alignments for all such clusters and making maximum likelihood trees for all such alignments is a tractable undertaking. Because LGT among prokaryotes is a real and pervasive process shaping prokaryote genome evolution [55,58,73–77], one has to treat each gene as a marker of its own evolution, not as a proxy for other genes or as a function that is subordinate to ribosomal phylogeny.

Genes that are present in several bacterial lineages and one archaeal lineage (or vice versa) might have been present in LUCA, but they might also have been the result of LGT [55,56,58]. An example illustrates how each gene tree can discriminate between vertical inheritance from LUCA and interdomain LGT. A recent study investigated the 6.1 million proteins encoded in 1,981 prokaryotic genomes (1,847 bacteria and 134 archaea) [78]. The proteins were clustered using the standard Markov Cluster (MCL) method [79]. The first step in that procedure is a matrix containing 18.5 trillion elements ($(n^2-n)/2$), each element corresponding to a pairwise amino acid sequence comparison. The clustering of such a matrix requires substantial

computational power and is aided by the availability of several terabytes of memory in a single machine. The MCL algorithm samples the distribution of values in the matrix and then starts removing the weak edges, with the value of "weak" being specified by the user. Two kinds of thresholds are typically used in MCL clustering: BLAST e-values and amino acid identity in pairwise alignments.

When the goal of clustering is to make alignments and trees, our group has found that a clustering threshold of 25% amino acid identity is a good rule of thumb. At lower thresholds, amino acid identity starts to approach random values and generates random errors in alignments [80], carrying over as erroneous topologies in trees [81]. That is why Russell F. Doolittle coined the term "twilight zone" for amino acid identity at or below the 20% range [82,83]. Of course, many proteins or domains that clearly share a common ancestry by the measure of related crystal structures do not share more than a random amino acid sequence identity [84]. Such ancient folds will fall into separate clusters at the 25% identity threshold and might thus generate false negatives when it comes to presence in LUCA (but see next section).

From thousands of clusters and trees, a handful remain

Using the 25% identity threshold, the 6.1 million prokaryotic proteins sampled fall into 286,514 clusters of at least two sequences, and 11,093 of those clusters include sequences found in both archaea and bacteria [78]. Many of those clusters involve oxygen-dependent respiratory chains. Did LUCA have 11,000 genes in its genome and breathe oxygen? That is, was LUCA (and hence archaea) descended from cyanobacteria? Neither prospect seems likely enough to warrant further discussion [85]. Knowing that transdomain LGT is prevalent [54–56] and that thousands of typically bacterial genes are shared with only one archaeal group [58], Weiss and colleagues [78] reasoned that a simple way to exclude some LGTs would be to set the minimal phylogenetic criteria that 1) a gene needs to be present in bacteria and archaea, 2) it needs to be present in at least two phylum-level clades, and 3) the tree needs to preserve domain monophyly (Fig 2). Genes that do not fulfil criterion 1 are not candidates for LUCA anyway. The two-phylum-plus-monophyly criteria 2 and 3 make it less likely but not impossible that such a gene attained that distribution via LGT. How so? Criteria 2 and 3 would require one transdomain transfer followed by intradomain transfers to different phyla, while allowing no subsequent, independent transdomain transfers. The last condition is the restrictive one.

Of the 11,093 clusters that harbored sequences in bacteria and archaea, only 355 (3%) passed the simple LGT filter [78]. Put another way, 97% of the sequences present in bacteria and archaea apparently underwent some transdomain LGT, underscoring the degree to which transdomain LGT has influenced gene history since LUCA and underscoring the need to employ phylogenetic filters in search of genes that trace to LUCA [21,51]. The 97% LGT value is important with regard to the 25% clustering threshold and possible false negatives; 97% of all false negatives founded in low-sequence conservation would still not trace to LUCA because of transdomain LGTs. But transdomain LGT has apparently not erased all signals, as 355 genes passed the LGT test, and those genes tell us things about LUCA that we did not know before.

The physiology of LUCA

Most earlier depictions of LUCA focused on what it was like [16]; for example, whether it was like RNA [86], like a virus [87], whether it was like prokaryotes in terms of its genetic code [88], or like eukaryotes in terms of its cellular organization [22]. But traditional approaches lacked information about how and from what LUCA lived [16]. Our phylogenetic approach to LUCA [78] uncovered information about what LUCA was doing: its physiology, its ecology, and its environment. The genes for those physiological traits are not necessarily widespread

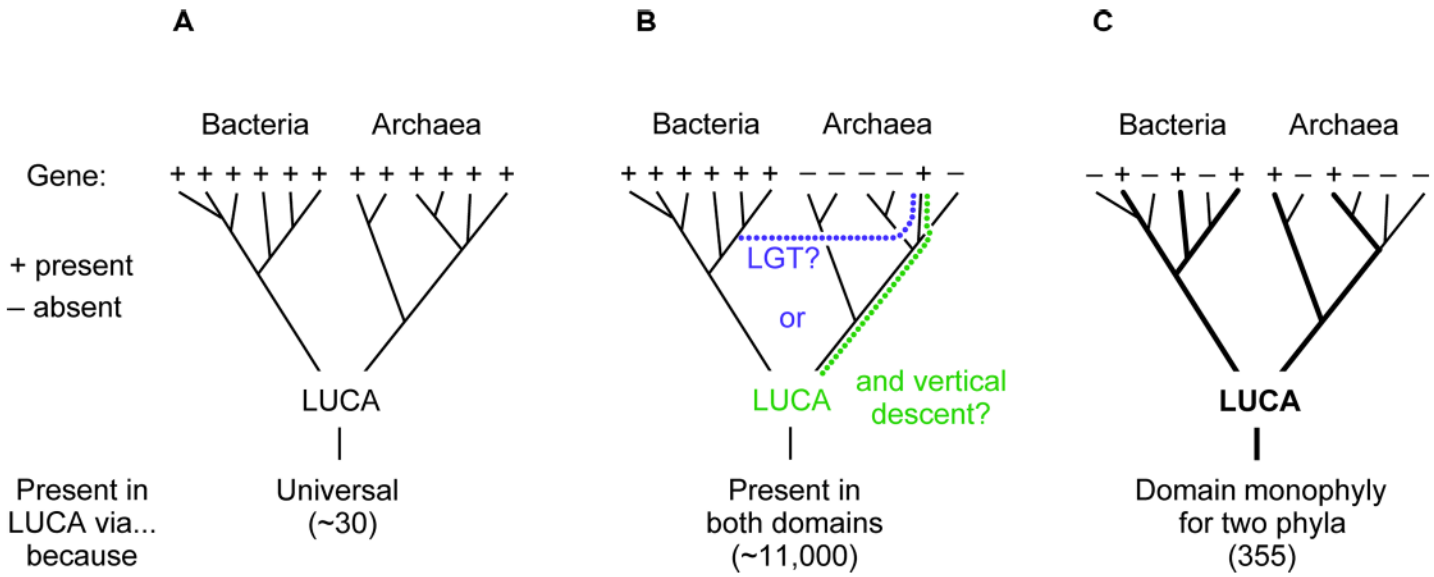


Fig 2. Three ways to infer genes present in LUCA. The gene presence is indicated with a plus sign, absence with a minus sign. a) Genes found universally in both domains, regardless of their tree, trace to LUCA. About 30 fulfil this criterion. b) Another way to trace genes to LUCA is to say that any gene found in both archaea and bacteria was present in LUCA. However, thousands of these genes will have been transferred between bacteria and archaea by LGT so were not necessarily present in LUCA. c) Genes present in only one bacterial or archaeal phylum could easily be the result of LGT and are removed. But pr[illegible] in two phyla per domain while preserving domain monophyly yields good candidates to have been present in LUCA. Such phylogenies would only result i[illegible] under very specific and restrictive conditions. They require exactly one transdomain transfer followed by either i) one additional transdomain LGT from the same donor lineage to a different recipient phylum or ii) retention during phylum divergence in the recipient domain, plus—in addition to either criteria i) or ii)—an additional, more subtle but highly restrictive criterion: No further transdomain LGTs occurred during all of evolution. Subsequent transdomain LGT would violate domain monophyly for the gene. Indeed, transdomain LGT is common, and 97% of the trees examined by Weiss and colleagues [78] did not exclude transdomain LGT (remaining 3%, 355 trees, provided in S1 Appendix). LGT, lateral gene transfer; LUCA, last universal common ancestor.

https://doi.org/10.1371/journal.pgen.1007518.g002

among modern genomes, but the filtering criteria by Weiss and colleagues [78] only require that these genes are ancient. What Weiss and colleagues [78] found is schematically summarized in Fig 3.

LUCA was an anaerobe, as long predicted by microbiologists [89]. Its metabolism was replete with O_2-sensitive enzymes. These include proteins rich in O_2-sensitive iron–sulfur (FeS) clusters and enzymes that entail the generation of radicals (unpaired electrons) via S-adenosyl methionine (SAM) in their reaction mechanisms. That fits well with the 50-year-old [90] but still modern view that FeS clusters represent very ancient cofactors in metabolism [91–93]. It also fits with newer insights about the ancient and spontaneous (nonenzymatic) chemistry underlying SAM synthesis [94].

LUCA lived from gasses. For carbon assimilation, LUCA used the simplest and most ancient of the six known pathways of CO_2 fixation, called the acetyl–CoA (or Wood–Ljungdahl) pathway [95–97], which is increasingly central for our concepts on early evolution because of its chemical simplicity [97,98] and exergonic nature [99–101]. In the acetyl–CoA pathway, CO_2 is reduced with hydrogen (H_2) to a methyl group and CO. The methyl group is synthesized by the methyl branch of the pathway, which employs different one-carbon (C1) carriers in bacteria (tetrahydrofolate) and archaea (tetrahydromethanopterin), cofactors that are synthesized by unrelated biosynthetic pathways [96]. Carbon monoxide (CO) is synthesized by carbon monoxide dehydrogenase (CODH), the archaeal and bacterial versions of which are distinct but related [96]. The methyl and carbonyl moieties are condensed to an enzyme-bound acetyl group that is removed from a metal cluster in acetyl–CoA synthase (ACS) as an energy rich thioester. Thioesters harbor chemically reactive bonds [102] that play

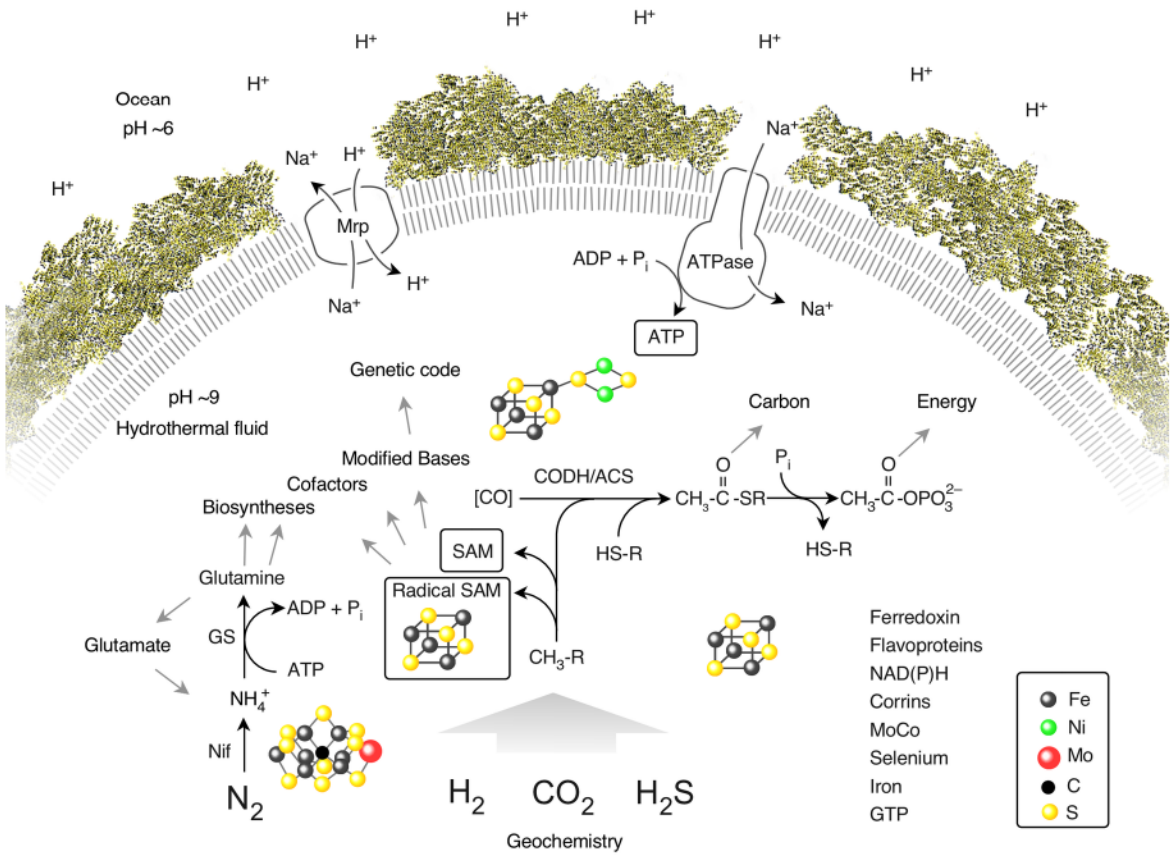


Fig 3. The physiology of LUCA. Summary of the main interactions of LUCA with its environment, reprinted with permission from [78] (supporting trees in S1 Appendix). Components listed at the lower right are present in LUCA. The figure does not make a statement regarding the source of CO in primordial metabolism, symbolized by [CO]. LUCA indisputably possessed genes because it had a genetic code. Transition metal clusters are symbolized. CH_3-R, methyl groups; CODH/ACS, carbon monoxide dehydrogenase/acetyl–CoA synthase; GS, glutamine synthetase; HS-R, organic thiols; LUCA, last universal common ancestor; Mrp, MrP type Na^+/H^+ antiporter; Nif, nitrogenase; SAM, S-adenosyl methionine.

https://doi.org/10.1371/journal.pgen.1007518.g003

a crucial role in energy metabolism [101] and in metabolism in general, both modern and ancient [101,103,104]. Although CODH/ACS clearly does trace to LUCA [78,96], this is not true for the methyl synthesis branch, which consists of unrelated enzymes in bacteria and archaea [78,96].

A recent report [105] argued that the presence of CODH in LUCA did not exclude a heterotrophic lifestyle for LUCA. This argument is problematic because no single enzyme defines a trophic lifestyle. Even Rubisco (D-ribulose-1, 5-bisphosphate carboxylase/oxygenase), the classical Calvin cycle enzyme, is not a marker for autotrophy because Rubisco also functions in a simpler heterotrophic pathway of RNA fermentation [106–108] that is common among archaea and bacteria in marine sediment environments [109]. Moreover, all heterotrophs are derived from autotrophs due to the former requiring the latter as a source of chemically defined growth substrates. The reason is that CO_2 constituted the main carbon source on Earth after the moon-forming impact [1,110], while carbon delivered from space was either too reduced to be fermented (polyaromatic hydrocarbons), too heterogeneous in structure to support microbial growth, or both [108]. Autotrophs with CODH can obtain ATP from CO_2 reduction with H_2 [98,101,110]. Autotrophs without CODH cannot. If we base inferences

about LUCA's lifestyle on broad criteria rather than single genes [105], LUCA was an autotroph [78,108].

Life is about harnessing energy [44]. Thioesters are chemically reactive—they forge direct links between carbon metabolism and energy metabolism (ATP synthesis) as they give rise to acetyl phosphate, the possible precursor of ATP in evolution as a currency of high-energy bonds [111]. Relics of ATP synthesis via acetyl phosphate were found in LUCA's genes [78], as were subunits of the rotor–stator ATP synthase itself. The ATP synthase might appear to present a paradox because no proteins of the proton-pumping machinery that cells use to generate the ion gradient that drives the ATP synthase traced to LUCA [78]. Yet some theories have it that the first cells arose at alkaline hydrothermal vents [91,96,111], meaning that the inside of the vent is more alkaline than the ocean outside. Such naturally existing pH gradients could have been harnessed by LUCA to synthesize ATP (Fig 3). Ancestral ATPases might have harnessed either proton gradients or sodium gradients generated by proton/sodium (H^+/Na^+) dependent antiporters [112], or they might have even been promiscuous for both kinds of ions, similar to the ATPase of modern microbes that live near the thermodynamic limits of life [113].

LUCA's environment was rich in sulfur; thioesters, SAM, proteins rich in FeS and iron–nickel–sulfur (FeNiS) clusters, sulfurtransferases, and thioredoxins were part of its repertoire, as were hydrogenases that could channel electrons from environmental H_2 to reduced ferredoxin, which is the main currency of reducing power (electrons) in anaerobes [114]. A recent report provided phylogenetic evidence that archaea are ancestrally H_2-dependent methanogens [62], compatible with an autotrophic, H_2-dependent lifestyle of LUCA.

LUCA had a reverse gyrase, an enzyme typical of thermophiles, suggesting that LUCA liked it hot. But independent of the reverse gyrase, simple chemical kinetics provide strong evidence in favor of a thermophilic origin for the first cells [115,116]. The reason is that only uncatalysed or inorganically catalysed reactions existed before there were enzymes. Their rates of reaction were lower than the enzymatically catalyzed reactions. Between 0˚C and 120˚C (the biologically relevant temperature range), organic chemical reaction rates generally increase with temperature [115,116]. Before there were enzymes, high-temperature environments were more conducive to organic chemical reactions than low-temperature environments [115,116]. Taken together, LUCA's requirement for gasses (CO_2, H_2, CO, nitrogen [N_2]), the prevalence of sulfide, its affinity to high temperature and metals, plus an ability to use but not generate ion gradients all point to the same environment: alkaline hydrothermal vents.

In addition to shedding light on physiology, the 355 trees that showed domain monophyly (S1 Appendix) [78] also have another interesting property: they are reciprocally rooted. That is, the bacteria are rooted in an archaeal outgroup and vice versa. Genes present in LUCA contain information about their lineages and about the groups of bacteria and archaea that branched most deeply in each domain. In both cases, the answer was clostridia (bacteria) and methanogens (archaea). Those are strictly anaerobic prokaryotes that use the acetyl–CoA pathway; live from CO_2, H_2, and CO; fix N_2, and today inhabit hydrothermal environments in the Earth's crust [117–119].

The onset of genetics

Though the organization of inanimate matter into living cells with genetics can be charted in mathematical terms [120,121], the biochemical details remain elusive. For example, it is controversial whether LUCA had DNA or not [87]. Several DNA-binding proteins trace to LUCA [78], so it would appear that LUCA possessed DNA, but it is unresolved whether LUCA could

actually replicate DNA. For LUCA, DNA might just have been a chemically stable repository for RNA-based replication [122].

A novel and interesting aspect of LUCA's biology concerns modified bases and the genetic code. Transfer RNA requires modified bases for proper interaction with mRNA (codon–anticodon wobble base pairing) and with rRNA in the ribosome during translation. That is, modified bases are part of the universal genetic code (Fig 4), which was present in LUCA. Many RNA-modifying enzymes trace to LUCA, particularly the enzymes that modify tRNA. Several of those enzymes are methyltransferases (many SAM dependent), and they remind us that, before the genetic code arose, the four main RNA bases could hardly have been in great supply in pure form because there were no genes or enzymes, only chemical reactions [123]. Spontaneous synthesis of bases in a real early Earth environment like a hydrothermal vent, an environment that lacks the control of a modern laboratory [124], is not likely to generate the four main bases in pure form. Many side products will accumulate, including chemically modified bases [111]. Chemically modified bases from living cells have been reported since the 1970s by pioneering RNA chemists such as Mathias Sprinzl [125] and Henri Grosjean [126]. There are 28 modified bases, mainly occurring in tRNA, that are shared by bacteria and archaea [127]. The modifications are chemically simple, such as the introduction of methyl groups or sulfur and occasionally of acetyl groups and the like (Fig 4).

Chemical modifications in the tRNA anticodon are essential for codon–anticodon interactions to work [128,129]. Modifications of the rRNA are concentrated around the peptidyl transferase site and are also essential for tRNA ribosome interactions [130]). It is possible that the genetic code itself arose in the same chemically reactive environment where LUCA arose and that modified bases in tRNA carry the chemical imprint of that environment [78]. That would forge a link between the early Earth and genetics as we know it. New laboratory syntheses of RNA molecules in the origin of life context now also include investigations of modified bases [131], as it is becoming increasingly clear that these are crucial components at the very earliest phases of molecular and biological evolution.

Moving forward

Investigations of LUCA based on phylogenies of all genes pose new opportunities and new challenges. As environmental sequencing and metagenomics progresses, the number of microbial sequences and new lineages is exploding [48,109]. How will that aspect of metagenomics affect investigations of LUCA? If the criteria for gene age are phylogenetic (prokaryote domain monophyly, presence in at least two bacterial and archaeal "phyla"), then the correct taxonomic assignment of each sequence is very important. A problematic aspect of metagenomic data is that some data handling steps can assign incorrect higher taxon labels to genes [39,41,43], which in turn can falsify phylogenetic relationships. Analyses of cultured microbes or complete genome sequences limit the available sample size but deliver reliable taxon labels, at least at the level of archaea versus bacteria. Clearly, there are trade-offs.

At first sight, LUCA's genome appears doomed to shrinkage. As the sample of complete genomes grows, the list of 355 genes that trace to LUCA by domain monophyly criteria [78] will shrink because each new genome offers new opportunities to uncover recent LGT events for the 355 genes. Recalling that only 3% of the 11,093 clusters investigated [78] appeared free of transdomain LGT, it is evident that the inclusion of new genomes will eventually cause the number 355 to asymptotically approach zero, unless some genes never undergo transdomain LGT, which seems unlikely. What to do? Filtering out recent LGT events would help save LUCA's genome from shrinking to zero. For example, the tree for gene X might violate domain monophyly by one LGT event. If the LGT was recent, affecting members of only one

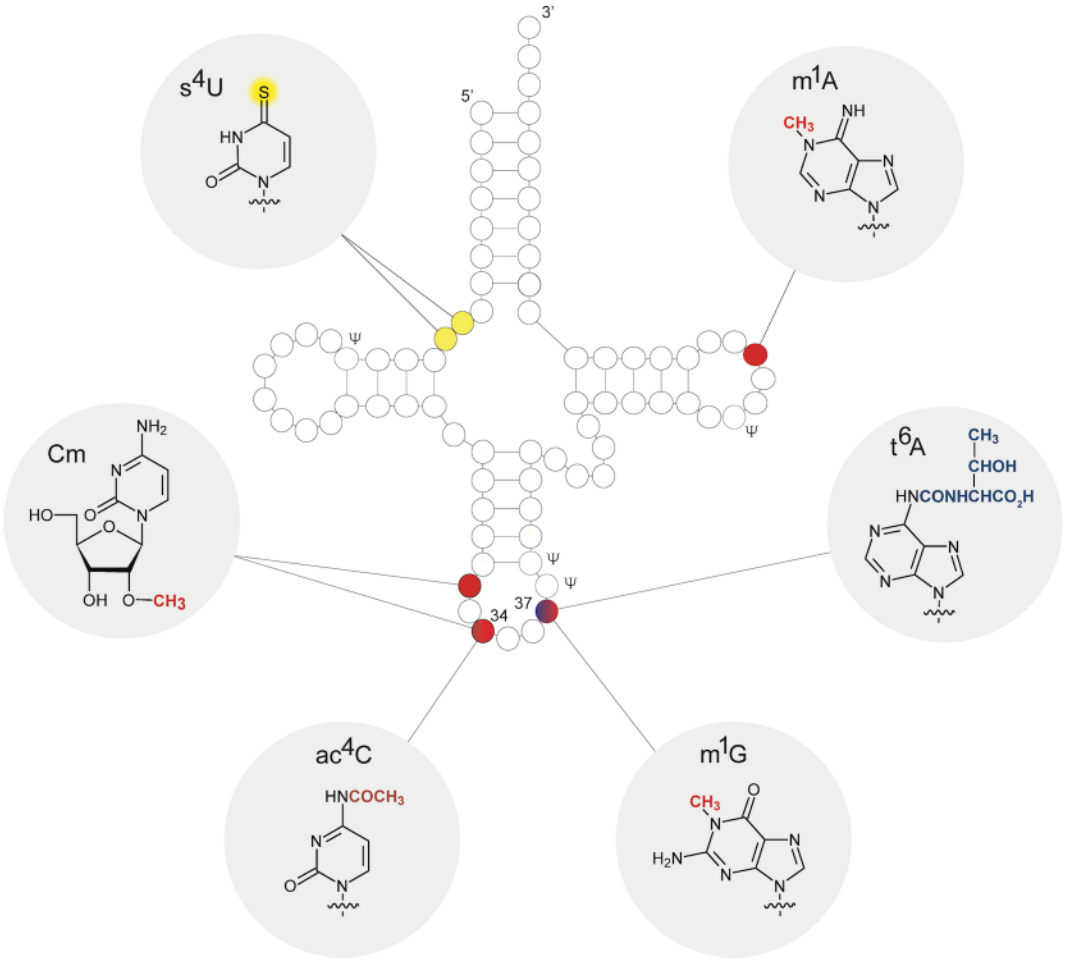


Fig 4. Modified tRNA and nucleoside structures (adapted from [78]). Cloverleaf secondary structure representation of tRNA showing post-transcriptional nucleoside modifications that are conserved among bacteria and archaea in both identity and position. The structures of respective conserved modified nucleosides are highlighted in grey. Methyl and acetyl groups are shown in red and dark red, respectively; sulfur in yellow; and the threonylcarbamoyl group in blue.

https://doi.org/10.1371/journal.pgen.1007518.g004

recipient genus or family, it would hardly affect inferences about LUCA, adding gene X to LUCA's list. To identify recent LGTs in prokaryote phylogeny, standard criteria like incomplete amelioration [132], anomalously high-sequence identity [133], or presence in the auxiliary genome [134] will be useful, as will new methods that root unrooted trees [135]. Identifying recent LGTs should allow us to trace more genes to LUCA.

There is also the issue of clustering thresholds to consider, as discussed above. Stringent thresholds produce many small clusters and more relaxed thresholds produce a smaller number of very large clusters [136]. One can argue that large clusters (low stringency) allow one to look further back into time, but they also can generate clusters whose origins trace to duplications in LUCA, in which domain monophyly is violated but not because of LGT. Another factor concerns gene fusions. Genes tend to undergo fusion and fission during evolution [137,138]. In clustering procedures, gene fusions tend to slightly reduce the number of clusters because when they occur, they can bring two fused genes into one alignment, and the weaker phylogenetic signal in the fusion is obscured [23]. Methods to detect fusions exist [139,140]. By detecting gene fusions and dissecting them into their component parts, it might be possible to increase the number of trees that trace to LUCA by phylogenetic criteria.

Investigations into early evolution always elicit protest. For example, there were criticisms [141] of the term "progenote," which Woese and Fox [142] introduced to designate a state of organization below that of a free-living cell [143,144], as shown in Fig 3. In addition, multiple LGTs can, in principle, generate false positives by mimicking vertical inheritance from LUCA [78], but very specific conditions have to be fulfilled (Fig 1C). The challenge is to distill a chronicle of microbial evolution that takes all genes and LGT [145] into account and that conveys information about physiology [146], the energy-releasing reactions that power microbial evolution.

Conclusions

More clues about LUCA's lifestyle are emerging. Investigations of modern biochemical pathways hone in on the same kinds of reactions as the phylogenetic approach [103]. Similarly, laboratory experiments also demonstrate the spontaneous synthesis of end products and intermediates of the acetyl–CoA pathway, the mainstay of LUCA's physiology; new findings show that formate, methanol, acetyl moieties, and even pyruvate arise spontaneously at high yields and at temperatures conducive to life (30˚C–100˚C) from CO_2, native metals, and water [98,147]. Those conditions are virtually impossible to underbid in terms of chemical simplicity [98], yet they bring forth the core of LUCA's carbon and energy metabolism [78,96,97,101, 103] overnight. Did the origin of genetics hinge upon hydrothermal chemical conditions that gave rise to the first biochemical pathways that in turn gave rise to the first cells? Genes that trace to LUCA [78], ancient biochemical pathways [103], and aqueous reactions of CO_2 with iron and water [98,110] all seem to converge on similar sets of simple, exergonic chemical reactions as those that occur spontaneously at hydrothermal vents [148]. From the standpoint of genes, physiology, laboratory chemistry, and geochemistry, it is beginning to look like LUCA was rooted in rocks.

Supporting information

S1 Appendix. ML trees for the 355 protein families that trace to LUCA by phylogenetic criteria. The trees are for the 355 clusters that, after alignment and tree construction, generated ML trees that preserve domain monophyly while also having homologues in $\geq$2 archaeal and $\geq$2 bacterial lineages. These 355 proteins trace to LUCA by those phylogenetic criteria [78]. LUCA, last universal common ancestor; ML, maximum likelihood.
(ZIP)

Acknowledgments

We thank Filipa Sousa (Universität Wien), Harun Tüysüz (Max-Planck-Institut für Kohlenforschung, Mülheim an der Ruhr), and Joseph Moran (University of Strasbourg) for discussions.

References

1. Zahnle K, Arndt N, Cockell C, Halliday A, Nisbet E, Selsis F, et al. Emergence of a habitable planet. Space Sci Rev. 2007; 129: 35±78.
2. Arndt N, Nisbet E. Processes on the young Earth and the habitats of early life. Annu Rev Earth Planet Sci. 2012; 40: 521±549.
3. Mojzsis SJ, Harrison TM, Pidgeon RT. Oxygen-isotope evidence from ancient zircons for liquid water at the Earth's surface 4,300 Myr ago. Nature. 2001; 409: 178±181. https://doi.org/10.1038/35051557 PMID: 11196638

4. Hirschmann MM. Water, melting, and the deep earth H_2O cycle. Annu Rev Earth Planet Sci. 2006; 34: 629±653.
5. Fei H, Yamazaki D, Sakurai M, Miyajima N, Ohfuji H, Katsura T, et al. A nearly water-saturated mantle transition zone inferred from mineral viscosity. Sci Adv. 2017; 3: e1603024. https://doi.org/10.1126/sciadv.1603024 PMID: 28630912
6. Tashiro T, Ishida A, Hori M, Igisu M, Koike M, Méjean P, et al. Early trace of life from 3.95 Ga sedimentary rocks in Labrador, Canada. Nature. 2017; 549: 516±518. https://doi.org/10.1038/nature24019 PMID: 28959955
7. Kubyshkin V, Budisa N. Synthetic alienation of microbial organisms by using genetic code engineering: Why and how? Biotechnol J. 2017; 12: 1600097±1600110.
8. Haldane JBS. Origin of Life. The Rationalist Annual. 1929; 148: 3±10.
9. Crick FHC. The recent excitement in the coding problem. Prog Nucleic Acid Res Mol Biol. 1963; 1: 163±217.
10. Doolittle WF, Brown JR. Tempo, mode, the progenote, and the universal root. Proc Natl Acad Sci USA. 1994; 91: 6721±6728. PMID: 8041689
11. Pace NR. Origin of life-facing up to the physical setting. Cell. 1991; 65: 531±533. PMID: 1709590
12. Woese C. The universal ancestor. Proc Natl Acad Sci USA. 1998; 95: 6854±6859. PMID: 9618502
13. Forterre P, Philippe H. Where is the root of the universal tree of life? Bioessays. 1999; 21: 871±879. https://doi.org/10.1002/(SICI)1521-1878(199910)21:10<871::AID-BIES10>3.0.CO;2-Q PMID: 10497338
14. Penny D, Poole A. The nature of the last universal common ancestor. Curr Opin Genet Dev. 1999; 9: 672±677. PMID: 10607605
15. Koonin EV. Comparative genomics, minimal gene-sets and the last universal common ancestor. Nat Rev Microbiol. 2003; 1: 127±136. https://doi.org/10.1038/nrmicro751 PMID: 15035042
16. Becerra A, Delaye L, Islas S, Lazcano A. The very early stages of biological evolution and the nature of the last common ancestor of the three major cell domains. Annu Rev Ecol Evol Syst. 2007; 38: 361±379.
17. Di Giulio M. The last universal common ancestor (LUCA) and the ancestors of archaea and bacteria were progenotes. J Mol Evol. 2011; 72: 119±126. https://doi.org/10.1007/s00239-010-9407-2 PMID: 21079939
18. Fox GE. Origin and evolution of the ribosome. Cold Spring Harb Perspect Biol. 2010; 2: a003483. https://doi.org/10.1101/cshperspect.a003483 PMID: 20534711
19. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci USA. 1990; 87: 4576±4579. PMID: 2112744
20. Kyprides N, Overbeek R, Ouzounis C. Universal protein families and the functional content of the Last Universal Common Ancestor. J Mol Evol. 1999; 49: 413±423. PMID: 10485999
21. Castresana J, Moreira D. Respiratory chain in the last common ancestor of living organisms. J Mol Evol. 1999; 49: 453±460. PMID: 10486003
22. Doolittle WF. The nature of the universal ancestor and the evolution of the proteome. Curr Opin Struct Biol. 2000; 10: 355±358. PMID: 10851188
23. Ku C, Nelson-Sathi S, Roettger M, Sousa FL, Lockhart PJ, Bryant D, et al. Endosymbiotic origin and differential loss of eukaryotic genes. Nature. 2015; 524: 427±432. https://doi.org/10.1038/nature14963 PMID: 26287458
24. Williams TA, Foster PG, Cox CJ, Embley TM. An archaeal origin of eukaryotes supports only two primary domains of life. Nature. 2013; 504: 231±236. https://doi.org/10.1038/nature12779 PMID: 24336283
25. McInerney J, Pisani D, O'Connell MJ. The ring of life hypothesis for eukaryote origin is supported by multiple kinds of data. Philos Trans R Soc Lond B Biol Sci. 2015; 370: 20140323. https://doi.org/10.1098/rstb.2014.0323 PMID: 26323755
26. Lake JA. Origin of the eukaryotic nucleus determined by rate-invariant analysis of rRNA sequences. Nature.1988; 331: 184±186. https://doi.org/10.1038/331184a0 PMID: 3340165
27. Galtier N, Tourasse N, Gouy M. A nonhyperthermophilic common ancestor to extant life forms. Science. 1999; 283: 220±221. PMID: 9880254
28. Cox CJ, Foster PG, Hirt RP, Harris SR, Embley TM. The archaebacterial origin of eukaryotes. Proc Natl Acad Sci USA. 2008; 105: 20356±20361. https://doi.org/10.1073/pnas.0810647105 PMID: 19073919

29. Spang A, Saw JH, Jørgensen SL, Zaremba-Niedzwiedzka K, Martijn J, Lind AE, et al. Complex archaea that bridge the gap between prokaryotes and eukaryotes. Nature. 2015; 521: 173±179. https://doi.org/10.1038/nature14447 PMID: 25945739

30. Zaremba-Niedzwiedzka K, Caceres EF, Saw JH, Bäckström D, Juzokaite L, Vancaester E, et al. Asgard archaea illuminate the origin of eukaryotic cellular complexity. Nature. 2017; 541: 353±358. https://doi.org/10.1038/nature21031 PMID: 28077874

31. Esser C, Ahmadinejad N, Wiegand C, Rotte C, Sebastiani F, Gelius-Dietrich G, et al. A genome phylogeny for mitochondria among alpha-proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes. Mol Biol Evol. 2004; 21: 1643±1660. https://doi.org/10.1093/molbev/msh160 PMID: 15155797

32. Pisani D, Cotton JA, McInerney JO. Supertrees disentangle the chimerical origin of eukaryotic genomes. Mol Biol Evol. 2007; 24: 1752±1760. https://doi.org/10.1093/molbev/msm095 PMID: 17504772

33. Cotton JA, McInerney JO. Eukaryotic genes of archaebacterial origin are more important than the more numerous eubacterial genes, irrespective of function. Proc Natl Acad Sci USA. 2010; 107: 17252±17255. https://doi.org/10.1073/pnas.1000265107 PMID: 20852068

34. Martin WF, Tielens AGM, Mentel M, Garg SG, Gould SB. The physiology of phagocytosis in the context of mitochondrial origin. Microbiol Mol Biol Rev. 2017; 81: e00008±17. https://doi.org/10.1128/MMBR.00008-17 PMID: 28615286

35. Forterre P.The common ancestor of Archaea and Eukarya was not an archaeon. Archaea 2013;372396. https://doi.org/10.1155/2013/372396 PMID: 24348094

36. Mariscal C, Doolittle WF. Eukaryotes first: How could that be? Philos Trans R Soc Lond B Biol Sci. 2015; 370: 20140322. https://doi.org/10.1098/rstb.2014.0322 PMID: 26323754

37. Harish A, Kurland CG. Mitochondria are not captive bacteria. J Theor Biol. 2017; 434: 88±98. https://doi.org/10.1016/j.jtbi.2017.07.011 PMID: 28754286

38. Dagan T, Roettger M, Bryant D, Martin W. Genome networks root the tree of life between prokaryotic domains. Genome Biol Evol. 2010; 2: 379±392. https://doi.org/10.1093/gbe/evq025 PMID: 20624742

39. Da Cunha V, Gaia M, Gadelle D, Nasir A, Forterre P. Lokiarchaea are close relatives of Euryarchaeota, not bridging the gap between prokaryotes and eukaryotes. PLoS Genet. 2017; 13: e1006810. https://doi.org/10.1371/journal.pgen.1006810 PMID: 28604769

40. Spang A, Eme L, Saw JH, Caceres EF, Zaremba-Niedzwiedzka K, Lombard J, Guy L, Ettema TJG. Asgard archaea are the closest prokaryotic relatives of eukaryotes. PLoS Genet. 2018; 14: e1007080. https://doi.org/10.1371/journal.pgen.1007080 PMID: 29596421

41. Da Cunha V, Gaia M, Nasir A, Forterre P. Asgard archaea do not close the debate about the universal tree of life topology. PLoS Genet. 2018; 14: e1007215. https://doi.org/10.1371/journal.pgen.1007215 PMID: 29596428

42. Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng J±F, et al. Insights into the phylogeny and coding potential of microbial dark matter. Nature. 2013; 499: 431±437. https://doi.org/10.1038/nature12352 PMID: 23851394

43. Williams TA, Embley TM. Archaeal "dark matter" and the origin of eukaryotes. Genome Biol Evol. 2014; 6: 474±481. https://doi.org/10.1093/gbe/evu031 PMID: 24532674

44. Judson OP. The energy expansions of evolution. Nat Ecol Evol. 2017; 1: 138. https://doi.org/10.1038/s41559-017-0138 PMID: 28812646

45. Zachar I, Szathmáry E. Breath-giving cooperation: critical review of origin of mitochondria hypotheses. Biology Direct. 2017; 12. https://doi.org/10.1186/s13062-017-0190-5 PMID: 28806979

46. Lane N. Serial endosymbiosis or singular event at the origin of eukaryotes? J Theor Biol. 2017; 434: 58±67. https://doi.org/10.1016/j.jtbi.2017.04.031 PMID: 28501637

47. Gould SB. Membranes and evolution. Curr Biol. 2018; 28: R381±R385. https://doi.org/10.1016/j.cub.2018.01.086 PMID: 29689219

48. Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, et al. A new view of the tree of life. Nat Microbiol. 2016; 1: 16048. https://doi.org/10.1038/nmicrobiol.2016.48 PMID: 27572647

49. Fischer WW, Hemp J, Johnson JE. Evolution of oxygenic photosynthesis. Annu Rev Earth Planet Sci. 2016; 44: 647±683.

50. Sánchez-Baracaldo P, Raven JA, Pisani D, Knoll AH. Early photosynthetic eukaryotes inhabited low-salinity habitats. Proc Natl Acad Sci USA. 2017; 114: E7737±E7745. https://doi.org/10.1073/pnas.1620089114 PMID: 28808007

51. Baymann F, Lebrun E, Brugna M, Schoepp-Cothenet B, Giudici-Orticoni MT, Nitschke W. The redox protein construction kit: Pre-last universal common ancestor evolution of energy-conserving enzymes.

Philos Trans R Soc Lond B Biol Sci. 2003; 358: 267±274. https://doi.org/10.1098/rstb.2002.1184 PMID: 12594934

52. Martin WF, Bryant DA, Beatty JT. A physiological perspective on the origin and evolution of photosynthesis. FEMS Microbiol Rev. 2017; 42: 205±231.

53. Doolittle WF. Phylogenetic classification and the universal tree. Science. 1999; 284: 2124±2129. PMID: 10381871

54. Deppenmeier U, Johann A, Hartsch T, Merkl R, Schmitz RA, Martinez-Arias R, et al. The genome of *Methanosarcina mazei*: Evidence for lateral gene transfer between bacteria and archaea. J Mol Microbiol Biotechnol. 2002; 4: 453±461. PMID: 12125824

55. Wagner A, Whitaker RJ, Krause DJ, Heilers JH, van Wolferen M, van der Does C, et al. Mechanisms of gene flow in archaea. Nat Rev Microbiol. 2017; 15: 492±501. https://doi.org/10.1038/nrmicro.2017.41 PMID: 28502981

56. Nelson-Sathi S, Dagan T, Landan G, Janssen A, Steel M, McInerney JO, et al. Acquisition of 1,000 eubacterial genes physiologically transformed a methanogen at the origin of Haloarchaea. Proc Natl Acad Sci USA. 2012; 109: 20537±20542. https://doi.org/10.1073/pnas.1209119109 PMID: 23184964

57. López-García P, Zivanovic Y, Deschamps P, Moreira D. Bacterial gene import and mesophilic adaptation in archaea. Nat Rev Microbiol. 2015; 13: 447±456. https://doi.org/10.1038/nrmicro3485 PMID: 26075362

58. Nelson-Sathi S, Sousa FL, Roettger M, Lozada-Chávez N, Thiergart T, Janssen A, et al. Origins of major archaeal clades correspond to gene acquisitions from bacteria. Nature. 2015; 517: 77±80. https://doi.org/10.1038/nature13805 PMID: 25317564

59. Dagan T, Martin W. The tree of one percent. Genome Biol. 2006; 7: 118. https://doi.org/10.1186/gb-2006-7-10-118 PMID: 17081279

60. Kunin V, Goldovsky L, Darzentas N, Ouzounis CA. The net of life: Reconstructing the microbial phylogenetic network. Genome Res. 2005; 15: 954±959. https://doi.org/10.1101/gr.3666505 PMID: 15965028

61. Dagan T, Martin W. Ancestral genome sizes specify the minimum rate of lateral gene transfer during prokaryote evolution. Proc Natl Acad Sci USA. 2007; 104: 870±875. https://doi.org/10.1073/pnas.0606318104 PMID: 17213324

62. Williams TA, Szöllősi GJ, Spang A, Foster PG, Heaps SE, Boussau B, et al. Integrative modeling of gene and genome evolution roots the archaeal tree of life. Proc Natl Acad Sci USA. 2017; 114: E4602±E4611. https://doi.org/10.1073/pnas.1618463114 PMID: 28533395

63. Hansmann S, Martin W. Phylogeny of 33 ribosomal and six other proteins encoded in an ancient gene cluster that is conserved across prokaryotic genomes: Influence of excluding poorly alignable sites from analysis. Int J Syst Evol Microbiol. 2000; 50: 1655±1663. https://doi.org/10.1099/00207713-50-4-1655 PMID: 10939673

64. Charlebois RL, Doolittle WF. Computing prokaryotic gene ubiquity: Rescuing the core from extinction. Genome Res. 2004; 14: 2469±2477. https://doi.org/10.1101/gr.3024704 PMID: 15574825

65. Ciccarelli FD, Doerks T, von Mering C, Creevey CJ, Snel B, Bork P. Toward automatic reconstruction of a highly resolved tree of life. Science. 2006; 311: 1283±1287. https://doi.org/10.1126/science.1123061 PMID: 16513982

66. Martin W, Stoebe B, Goremykin V, Hansmann S, Hasegawa M, Kowallik KV. Gene transfer to the nucleus and the evolution of chloroplasts. Nature. 1998; 393: 162±165. https://doi.org/10.1038/30234 PMID: 11560168

67. Puigbò P, Wolf YI, Koonin EV. Search for a 'Tree of Life' in the thicket of the phylogenetic forest. J Biol. 2009; 8: 59. https://doi.org/10.1186/jbiol159 PMID: 19594957

68. Brown CT, Hug LA, Thomas BC, Sharon I, Castelle CJ, Singh A, et al. Unusual biology across a group comprising more than 15% of domain Bacteria. Nature. 2015; 523: 208±211. https://doi.org/10.1038/nature14486 PMID: 26083755

69. Gadagkar SR, Rosenberg MS, Kumar S. Inferring species phylogenies from multiple genes: concatenated sequence tree versus consensus gene tree. J Exp Zool B Mol Dev Evol. 2005; 304: 64±74. https://doi.org/10.1002/jez.b.21026 PMID: 15593277

70. Thiergart T, Landan G, Martin WF. Concatenated alignments and the case of the disappearing tree. BMC Evol Biol. 2014; 14: 266. https://doi.org/10.1186/s12862-014-0266-0 PMID: 25547755

71. Yang Z, Zhu T. Bayesian selection of misspecified models is over confident and may cause spurious posterior probabilities for phylogenetic trees. Proc Natl Acad Sci USA. 2018; 115: 1854±1859. https://doi.org/10.1073/pnas.1712673115 PMID: 29432193

72. Lockhart PJ, Howe CJ, Barbrook AC, Larkum AWD, Penny D. Spectral analysis, systematic bias, and the evolution of chloroplasts. Mol Biol Evol. 1999; 16: 573±576.

73. Ochman H, Lawrence JG, Groisman EA. Lateral gene transfer and the nature of bacterial innovation. Nature. 2000; 405: 299±304. https://doi.org/10.1038/35012500 PMID: 10830951

74. Popa O, Dagan T. Trends and barriers to lateral gene transfer in prokaryotes. Curr Opin Microbiol. 2011; 14: 615±623. https://doi.org/10.1016/j.mib.2011.07.027 PMID: 21856213

75. Treangen TJ, Rocha EP. Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes. PLoS Genet. 2011; 7: e1001284. https://doi.org/10.1371/journal.pgen.1001284 PMID: 21298028

76. McInerney JO, McNally A, O'Connell MJ. Why Prokaryotes have pangenomes. Nat Microbiol. 2017; 2: 17040. https://doi.org/10.1038/nmicrobiol.2017.40 PMID: 28350002

77. Bapteste E, Boucher Y, Leigh J, Doolittle WF. Phylogenetic reconstruction and lateral gene transfer. Trends Microbiol. 2004; 12: 406±411. https://doi.org/10.1016/j.tim.2004.07.002 PMID: 15337161

78. Weiss MC, Sousa FL, Mrnjavac N, Neukirchen S, Roettger M, Nelson-Sathi S, et al. The physiology and habitat of the last universal common ancestor. Nat Microbiol. 2016; 1: 16116. https://doi.org/10.1038/nmicrobiol.2016.116 PMID: 27562259

79. Enright AJ, Van Dongen S, Ouzounis CA. An ancient algorithm for largescale detection of protein families. Nucleic Acids Res. 2002; 30: 1575±1584. PMID: 11917018

80. Landan G, Graur D. Heads or tails: A simple reliability check for multiple sequence alignments. Mol Biol Evol. 2007; 24: 1380±1383. https://doi.org/10.1093/molbev/msm060 PMID: 17387100

81. Lockhart PJ, Steel M, Hendy MD, Penny D. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol Biol Evol. 1994; 11: 605. https://doi.org/10.1093/oxfordjournals.molbev.a040136 PMID: 19391266

82. Doolittle RF. Of URFs and ORFs: A primer on how to analyze derived amino acid sequences. 1st ed. Mill Valley, CA: University Science Books; 1986.

83. Rost B. Twilight zone of protein sequence alignments. Protein Eng. 1999; 12: 85±94. PMID: 10195279

84. Rossmann MG, Moras D, Olsen KW. Chemical and biological evolution of nucleotide-binding protein. Nature. 1974; 250: 194±199. PMID: 4368490

85. Martin WF, Sousa FL. Early microbial evolution: The age of anaerobes. Cold Spring Harb Perspect Biol. 2015; 8: a018127. https://doi.org/10.1101/cshperspect.a018127 PMID: 26684184

86. Anantharaman V, Koonin EV, Aravind L. Comparative genomics and evolution of proteins involved in RNA metabolism. Nucleic Acids Res. 2002; 30: 1427±1464. PMID: 11917006

87. Forterre P. Three RNA cells for ribosomal lineages and three DNA viruses to replicate their genomes: A hypothesis for the origin of cellular domain. Proc Natl Acad Sci USA. 2006; 103: 3669±3674. https://doi.org/10.1073/pnas.0510333103 PMID: 16505372

88. Di Giulio M. An autotrophic origin for the coded amino acids is concordant with the coevolution theory of the genetic code. J Mol Evol. 2016; 83: 93±96. https://doi.org/10.1007/s00239-016-9760-x PMID: 27743002

89. Decker K, Jungerman K, Thauer RK. Energy production in anaerobic organisms. Angew. Chem. Int. Ed. 1970; 9: 138±158.

90. Eck RV, Dayhoff MO. Evolution of the structure of ferredoxin based on living relics of primitive amino acid sequences. Science. 1966; 152: 363±366. https://doi.org/10.1126/science.152.3720.363 PMID: 17775169

91. Russell MJ, Hall AJ. The emergence of life from iron monosulphide bubbles at a submarine hydrothermal redox and pH front. J Geol Soc Lond. 1997; 154: 377±402.

92. Camprubi E, Jordan SF, Vasiliadou R, Lane N. Iron catalysis at the origin of life. IUBMB Life. 2017; 69: 373±381. https://doi.org/10.1002/iub.1632 PMID: 28470848

93. Shock EL, McCollom T, Schulte MD. The emergence of metabolism from within hydrothermal systems. In: Wiegel J, Adams MWW, editors. Thermophiles: The Keys to Molecular Evolution and the Origin of Life. Taylor and Francis; 1998. pp. 59±76.

94. Laurino P, Tawfik DS. Spontaneous emergence of S-adenosylmethionine and the evolution of methylation. Angew Chem Int Ed Engl. 2017; 56: 343±345. https://doi.org/10.1002/anie.201609615 PMID: 27901309

95. Ragsdale SW. Enzymology of the Wood-Ljungdahl pathway of acetogenesis. Ann NY Acad Sci. 2008; 1125: 129±136. https://doi.org/10.1196/annals.1419.015 PMID: 18378591

96. Sousa FL, Martin WF. Biochemical fossils of the ancient transition from geoenergetics to bioenergetics in prokaryotic one carbon compound metabolism. Biochim Biophys Acta. 2014; 1837: 964±981. https://doi.org/10.1016/j.bbabio.2014.02.001 PMID: 24513196

97. Fuchs G. Alternative pathways of carbon dioxide fixation: insights into the early evolution of life? Annu Rev Microbiol. 2011; 65: 631±658. https://doi.org/10.1146/annurev-micro-090110-102801 PMID: 21740227

98. Varma SJ, Muchowska KB, Chatelain P, Moran J. Native iron reduces CO_2 to intermediates and end-products of the acetyl-CoA pathway. Nat Ecol Evol. 2018; 2: 1019±1024. https://doi.org/10.1038/s41559-018-0542-2 PMID: 29686234

99. Fuchs G. Variations of the acetyl-CoA pathway in diversely related microorganisms that are not acetogens. In Drake HL, editor. Acetogenesis. Chapman & Hall Microbiology Series (Physiology / Ecology / Molecular Biology / Biotechnology). Springer, Boston, MA; 1994. pp. 506±538.

100. Russell MJ, Martin W. The rocky roots of the acetyl-CoA pathway. Trends Biochem Sci. 2004; 29: 358±363. https://doi.org/10.1016/j.tibs.2004.05.007 PMID: 15236743

101. Martin WF, Thauer RK. Energy in ancient metabolism. Cell. 2017; 168: 953±955. https://doi.org/10.1016/j.cell.2017.02.032 PMID: 28283068

102. Semenov SN, Kraft LJ, Ainla A, Zhao M, Baghbanzadeh M, Campbell VE et al. Autocatalytic, bistable, oscillatory networks of biologically relevant organic reactions. Nature. 2016; 537: 656±660. https://doi.org/10.1038/nature19776 PMID: 27680939

103. Goldford JE, Hartman H, Smith TF, Segrè D. Remnants of an ancient metabolism without phosphate. Cell. 2017; 168: 1126±1134. https://doi.org/10.1016/j.cell.2017.02.001 PMID: 28262353

104. Goldford JE, Segrè D. Modern views of ancient metabolic networks. Curr Opin Syst Biol. 2018; in press. https://doi.org/10.1016/j.coisb.2018.01.004

105. Adam PS, Borrel G, Gribaldo S. Evolutionary history of carbon monoxide dehydrogenase/acetyl-CoA synthase, one of the oldest enzymatic complexes. Proc Natl Acad Sci USA. 2018; 115: E1166±E1173. https://doi.org/10.1073/pnas.1716667115 PMID: 29358391

106. Sato T, Atomi H, Imanaka T. Archaeal type III RuBisCOs function in a pathway for AMP metabolism. Science. 2007; 315: 1003±1006. https://doi.org/10.1126/science.1135999 PMID: 17303759

107. Aono R, Sato T, Imanaka T, Atomi H. A pentose bisphosphate pathway for nucleoside degradation in Archaea. Nat Chem Biol. 2015; 11: 355±360. https://doi.org/10.1038/nchembio.1786 PMID: 25822915

108. Schönheit P, Buckel W, Martin WF. On the origin of heterotrophy. Trends Microbiol. 2016; 24: 12±24. https://doi.org/10.1016/j.tim.2015.10.003 PMID: 26578093

109. Castelle CJ, Banfield JF. Major new microbial groups expand diversity and alter our understanding of the tree of life. Cell. 2018; 172: 1181±1197. https://doi.org/10.1016/j.cell.2018.02.016 PMID: 29522741

110. Sousa FL, Preiner M, Martin WF: Native metals, electron bifurcation, and CO_2 reduction in early biochemical evolution. Curr Opin Microbiol. 2018; 43: 77±83. https://doi.org/10.1016/j.mib.2017.12.010 PMID: 29316496

111. Martin W, Russell MJ. On the origin of biochemistry at an alkaline hydrothermal vent. Philos Trans R Soc Lond B. 2007; 362: 1887±1925.

112. Sojo V, Pomiankowski A, Lane N. A bioenergetic basis for membrane divergence in archaea and bacteria. PLoS Biol. 2014; 12: e1001926. https://doi.org/10.1371/journal.pbio.1001926 PMID: 25116890

113. Schlegel K, Leone V, Faraldo-Gómez JD, Müller V. Promiscuous archaeal ATP synthase concurrently coupled to Na^+ and H^+ translocation. Proc Natl Acad Sci USA. 2012; 109: 947±952. https://doi.org/10.1073/pnas.1115796109 PMID: 22219361

114. Buckel W, Thauer RK. Energy conservation via electron bifurcating ferredoxin reduction and proton/Na^+ translocating ferredoxin oxidation. Biochim Biophys Acta. 2013; 1827: 94±113. https://doi.org/10.1016/j.bbabio.2012.07.002 PMID: 22800682

115. Wolfenden R, Lewis CA Jr., Yuan Y, Carter CW Jr.. Temperature dependence of amino acid hydrophobicities. Proc Natl Acad Sci USA. 2015; 112: 7484±7488. https://doi.org/10.1073/pnas.1507565112 PMID: 26034278

116. Stockbridge RB, Lewis CA Jr., Yuan Y, Wolfenden R. Impact of temperature on the time required for the establishment of primordial biochemistry, and for the evolution of enzymes. Proc Natl Acad Sci USA. 2010; 107: 22102±22105. https://doi.org/10.1073/pnas.1013647107 PMID: 21123742

117. Chapelle FH, O'Neill K, Bradley PM, MethéBA, Ciufo SA, Knobel LL, et al. A hydrogen-based subsurface microbial community dominated by methanogens. Nature. 2002; 415: 312±315. https://doi.org/10.1038/415312a PMID: 11797006

118. Lever MA, Heuer VB, Morono Y, Masui N, Schmidt F, Alperin MJ, et al. Acetogenesis in deep subseafloor sediments of the Juan de Fuca Ridge Flank: A synthesis of geochemical, thermodynamic, and gene-based evidence. Geomicrobiol J. 2010; 27: 183±211.

119. Whitman WB, Coleman DC, Wiebe WJ. Prokaryotes: The unseen majority. Proc Natl Acad Sci USA. 1998; 95: 6578±6583. PMID: 9618454

120. Hordijk W, Steel M. Chasing the tail: The emergence of autocatalytic networks. Biosystems. 2017; 152: 1±10. https://doi.org/10.1016/j.biosystems.2016.12.002 PMID: 28027958

121. Steel M, Kauffmann S. A note on random catalytic branching processes. J Theoret Biol. 2018; 437: 222±224.

122. Koonin EV, Martin W. On the origin of genomes and cells within inorganic compartments. Trends Genet. 2005; 21: 647±654. https://doi.org/10.1016/j.tig.2005.09.006 PMID: 16223546

123. Baross JA, Martin WF. The ribofilm as a concept for life's origins. Cell 2015; 162: 13±15. https://doi.org/10.1016/j.cell.2015.06.038 PMID: 26140586

124. Sutherland JD. Opinion: Studies on the origin of life±the end of the beginning. Nat Rev Chem. 2017; 10: 0012. https://doi.org/10.1038/s41570-016-0012

125. Hartmann RK, Mörl M, Sprinzl M. The tRNA world. RNA. 2004; 10: 344±349. https://doi.org/10.1261/rna.5240904 PMID: 14970379

126. Grosjean H, Breton M, Sirand-Pugnet P, Tardy F, Thiaucourt F, Citti C, et al. Predicting the minimal translation apparatus: Lessons from the reductive evolution of mollicutes. PLoS Genet. 2014; 10: e1004363. https://doi.org/10.1371/journal.pgen.1004363 PMID: 24809820

127. Grosjean H, Gupta R, Maxwell ES. Modified nucleotides in archaeal RNAs. In: Blum P, editor. Archaea: New Models for Prokaryotic Biology. Norfolk, UK: Caister Academic Press; 2008. pp. 171± 196.

128. Yarian C, Townsend H, Czestkowski W, Sochacka E, Malkiewicz AJ, Guenther R, et al. Accurate translation of the genetic code depends on tRNA modified nucleosides. J Biol Chem. 2002; 277: 16391±16395. https://doi.org/10.1074/jbc.M200253200 PMID: 11861649

129. Gustilo EM, Vendeix FA, Agris PF. tRNA's modifications bring order to gene expression. Curr Opin Microbiol. 2008; 11: 134±140. https://doi.org/10.1016/j.mib.2008.02.003 PMID: 18378185

130. Decatur WA, Fournier MJ. rRNA modifications and ribosome function. Trends Biochem Sci. 2002; 27: 344±351. PMID: 12114023

131. Schneider C, Becker S, Okamura H, Crisp A, Amatov T, Stadlmeier M, et al. Noncanonical RNA nucleosides as molecular fossils of an early EarthÐG eneration by prebiotic methylations and carbamoylations. Angwandte Chemie Int Ed. 2018; 57: 1±5.

132. Lawrence JG, Ochman H. Molecular archaeology of the *Escherichia coli* genome. Proc Natl Acad Sci USA.1998; 95: 9413±9417. PMID: 9689094

133. Ku C, Martin WF. A natural barrier to lateral gene transfer from prokaryotes to eukaryotes revealed from genomes: The 70% rule. BMC Biol. 2016; 14: 89. https://doi.org/10.1186/s12915-016-0315-9 PMID: 27751184

134. Medini D, Donati C, Tettelin H, Masignani V, Rappuoli R. The microbial pan-genome. Curr Opin Genet Dev. 2005; 15: 589±594. https://doi.org/10.1016/j.gde.2005.09.006 PMID: 16185861

135. Tria FDK, Landan G, Dagan T. Phylogenetic rooting using minimal ancestor deviation. Nat Ecol Evol. 2017; 1: 193. https://doi.org/10.1038/s41559-017-0193 PMID: 29388565

136. Dagan T, Artzy-Randrup Y, Martin W. Modular networks and cumulative impact of lateral transfer in prokaryote genome evolution. Proc Natl Acad Sci USA. 2008; 105: 10039±10044. https://doi.org/10.1073/pnas.0800679105 PMID: 18632554

137. Kummerfeld SK, Teichmann SA. Relative rates of gene fusion and fission in multi-domain proteins. Trends Genet. 2005; 21: 25±30. https://doi.org/10.1016/j.tig.2004.11.007 PMID: 15680510

138. Méheust R, Watson AK, Lapointe FJ, Papke RT, Lopez P, Bapteste E. Hundreds of novel composite genes and chimeric genes with bacterial origins contributed to haloarchaeal evolution. Genome Biol. 2018. in press.

139. Henry CS, Lerma-Ortiz C, Gerdes SY, Mullen JD, Colasanti R, Zhukov A, et al. Systematic identification and analysis of frequent gene fusion events in metabolic pathways. BMC Genomics. 2016; 17: 473. https://doi.org/10.1186/s12864-016-2782-3 PMID: 27342196

140. Enright AJ, Iliopoulos I, Kyrpides NC, Ouzounis CA. Protein interaction maps for complete genomes based on gene fusion events. Nature. 1999; 402: 86±90. https://doi.org/10.1038/47056 PMID: 10573422

141. Gogarten JP, Deamer D. Is LUCA a thermophilic progenote? Nat Microbiol. 2016; 1: 16229. https://doi.org/10.1038/nmicrobiol.2016.229 PMID: 27886195

142. Woese CR, Fox GE. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc Natl Acad Sci USA. 1977; 74: 5088±5090. PMID: 270744

143. Weiss MC, Neukirchen S, Roettger M, Mrnjavac N, Nelson-Sathi S, Martin WF, et al. Reply to 'Is LUCA a thermophilic progenote?' Nat Microbiol. 2016; 1: 16230. https://doi.org/10.1038/nmicrobiol.2016.230 PMID: 27886196

144. Martin WF, Weiss MC, Neukirchen S, Nelson-Sathi S, Sousa FL. Physiology, phylogeny, and LUCA. Microb Cell. 2016; 3: 582±587. https://doi.org/10.15698/mic2016.12.545 PMID: 28357330

145. Martin W. Mosaic bacterial chromosomes: A challenge en route to a tree of genomes. BioEssays. 1999; 21: 99±104. https://doi.org/10.1002/(SICI)1521-1878(199902)21:2<99::AID-BIES3>3.0.CO;2-B PMID: 10193183

146. Liu Y, Beer LL, Whitman WB. Methanogens: A window into ancient sulfur metabolism. Trends Microbiol. 2012; 20: 251±258. https://doi.org/10.1016/j.tim.2012.02.002 PMID: 22406173

147. Muchowska KB, Varma SJ, Chevallot-Beroux E, Lethuillier-Karl L, Li G, Moran J. Metals promote sequences of the reverse Krebs cycle. Nat Ecol Evol. 2017; 1: 1716±1721. https://doi.org/10.1038/s41559-017-0311-7 PMID: 28970480

148. McCollom TM. Abiotic methane formation during experimental serpentinization of olivine. Proc Natl Acad Sci USA. 2016; 113: 13965±13970. https://doi.org/10.1073/pnas.1611843113 PMID: 27821742

Publication 5

Title: Something special about CO-dependent CO_2 fixation

Year: 2018

Authors: Joana C. Xavier, Martina Preiner, and William F. Martin

Published in: *FEBS Letters,* 285(22), pp. 4181–4195.
doi: 10.1111/febs.14664.

Contribution: Second author.
Medium: drafted and illustrated Figure 6. Analysed the chemical configurations of CO_2 and CO, drafted the discussion concerning this issue and edited throughout the manuscript.

Summary: This research article investigates the role of carbon monoxide (CO) in microbial CO_2 fixation with a focus on global metabolism. While CO_2 is omnipresent, CO is rare and interfaces in metabolism at only two enzymes. This and the differences between CO and CO_2 binding to the metal bearing enzyme centres leads to the conclusion that the role of CO in CO_2 fixation is ancient and is possibly connected to the abiotic beginnings of this kind of metabolism.

The FEBS Journal

FEBS PRESS science publishing by scientists

Something special about CO-dependent CO_2 fixation

Joana C. Xavier[1], Martina Preiner[1] and William F. Martin[1,2]

1 Institut für Molekulare Evolution, Heinrich Heine Universität Düsseldorf, Germany
2 Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Oeiras, Portugal

Keywords
carbon dioxide; carbon monoxide; CODH/ACS; enzymatic reactions; metabolic networks

Correspondence
J. C. Xavier, Institut für Molekulare Evolution, Heinrich Heine Universität Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany
Fax: +49 2118113554
Tel: +49 211 81 12736
E-mail: xavier@hhu.de
Website: https://www.molevol.hhu.de/

(Received 26 April 2018, revised 8 August 2018, accepted 19 September 2018)

doi:10.1111/febs.14664

Carbon dioxide enters metabolism via six known CO_2 fixation pathways, of which only one is linear, exergonic in the direction of CO_2-assimilation, and present in both bacterial and archaeal anaerobes – the Wood-Ljungdahl (WL) or reductive acetyl-CoA pathway. Carbon monoxide (CO) plays a central role in the WL pathway as an energy rich intermediate. Here, we scan the major biochemical reaction databases for reactions involving CO and CO_2. We identified 415 reactions corresponding to enzyme commission (EC) numbers involving CO_2, which are non-randomly distributed across different biochemical pathways. Their taxonomic distribution, reversibility under physiological conditions, cofactors and prosthetic groups are summarized. In contrast to CO_2, only 15 reaction classes involving CO were detected. Closer inspection reveals that CO interfaces with metabolism and the carbon cycle at only two enzymes: anaerobic carbon monoxide dehydrogenase (CODH), a Ni- and Fe-containing enzyme that generates CO for CO_2 fixation in the WL pathway, and aerobic CODH, a Mo- and Cu-containing enzyme that oxidizes environmental CO as an electron source. The CO-dependent reaction of the WL pathway involves carbonyl insertion into a methyl carbon-nickel at the Ni-Fe-S A-cluster of acetyl-CoA synthase (ACS). It appears that no alternative mechanisms to the CO-dependent reaction of ACS have evolved in nearly 4 billion years, indicating an ancient and mechanistically essential role for CO at the onset of metabolism.

Introduction

In autotrophs, carbon dioxide enters metabolism mainly via six known pathways of CO_2 fixation [1–5]. In discussions about novel synthetic CO_2 fixation pathways [6–10], it is often overlooked that heterotrophs also harbor a number of metabolic reactions that incorporate CO_2. For example, carbon atoms from CO_2 end up in the purine and pyrimidine rings during *de novo* nucleobase biosynthesis, from prokaryotes to humans [11], and CO_2 assimilation into membrane lipids has been measured as a proxy of metabolic activity in different heterotrophic bacteria [12]. Of the six natural pathways of autotrophic CO_2 fixation, only one involves CO as an intermediate – the Wood-Ljungdahl (WL) pathway, also called the reductive acetyl-CoA pathway.

Among CO_2 assimilation pathways, the WL pathway is unique in being the only linear pathway of carbon fixation that can occur exergonically [3,13,14]. Phylogenetic evidence traces the pathway to the genome of the Last Universal Common Ancestor (LUCA) [15]. It is the only known pathway of core CO_2 fixation present in both bacteria and archaea [2,3]. Gene distributions for both the enzymes of the pathway and the synthesis of its salient pterin cofactors – tetrahydrofolate (H_4F) in bacteria and tetrahydromethanopterin (H_4MPT) in

Abbreviations
ACS, acetyl-CoA synthase; CODH, carbon monoxide dehydrogenase; EC, enzyme commission; Gt, gigatonne; LUCA, last universal common ancestor; PAP, presence-absence pattern; rTCA, reverse citric acid; WL, Wood-Ljungdahl.

archaea – testify to the antiquity of the WL pathway [3,16], which is closely aligned with theories that posit a chemolithoautotrophic origin of life [17–19]. Its basic chemistry, the reduction in CO_2 to organic one-carbon (C1) moieties, occurs as spontaneous geochemical reactions in hydrothermal systems [20,21]. The WL pathway entails oxygen sensitive catalysts, as its enzymes are replete with iron and nickel sulfur centers essential for electron transfer and catalysis [14,22]. CO_2-reducing reactions of the WL pathway occur readily in the laboratory in the presence of native metals [23,24]. The WL pathway is the only pathway known that fixes CO_2 while conserving energy as ATP, the mechanisms of energy conservation entailing chemiosmotic coupling and flavin-based electron bifurcation [3,25]. The simplicity of the WL pathway [13,22], its antiquity [13–16,26,27], favorable energetics in the CO_2-reducing direction [3,25] and chemical similarity to exergonic geochemical reactions in hydrothermal vents [20,21] forge chemical links between early earth geochemistry and the biochemistry of the first cells.

The WL pathway works in a conceptually simple but chemically demanding manner – one carbon at a time [22]. The enzymology of the pathway has been reviewed [3,19,22,28,29]. In comparisons of the archaeal and bacterial pathway, the enzymes of the methyl synthesis branch show no sequence conservation across the prokaryotic domain divide [16], whereby the CO synthesis and thioesther synthesis are catalyzed by an enzyme well conserved between archaea and bacteria: bifunctional carbon monoxide dehydrogenase/acetyl CoA synthase (CODH/ACS). CODH catalyzes the reversible, ferredoxin-dependent interconversion of CO and CO_2 [30]. In the WL pathway, CO is generated as an intermediate of CO_2 fixation, but environmental CO can also enter the pathway as a carbon and electron source [3,31,32]. Both CODH and CO are central to carbon and energy metabolism in methanogens (archaea) [33], hydrogenogens, acetogens [22,34], some solventogenic bacteria, such as ethanol-producing *Clostridium ljungdahlii* [35] and other anaerobes including sulfate reducers [36], as reviewed in [32,37]. CODH contains FeS clusters, the active site contains an FeNiS cluster [38–40]. An anaerobic CODH preparation containing copper in the active site was reported [41], but the enzyme was inactive. The CODH enzyme of the WL pathway is oxygen sensitive. ACS catalyzes the cleavage and synthesis of acetyl-CoA, releasing or consuming CO, respectively. In *Moorella thermoacetica*, CO is carried inside the enzyme through a hydrophobic tunnel as proposed by scavenging experiments using hemoglobin [42] and subsequently supported by isotope exchange data [43] and structural data [44]. Some facultative aerobes, as *Rhodospirillum rubrum*, have the anaerobic CODH but no ACS, and use it to conserve energy in the reverse direction through CO oxidation [32].

In other aerobic and facultative aerobic bacteria, CO oxidation can also be catalyzed by an oxygen tolerant enzyme that shares no sequence similarity with CODH of the WL pathway. The oxygen tolerant CO oxidizing enzyme is encoded by the *cox* operon [45]. It is typically called aerobic CODH [45], but for clarity we will refer to it here by the name of its catalytic subunit, coxL. Importantly, coxL enzymes are not related to the CODH of the WL pathway, rather they are related to molybdenum hydroxylases [45,46]. The metals involved in coxL catalysis are molybdenum and copper [38,46–48]. *Cox* gene products only perform the oxidation of CO to CO_2, which in some species of Proteobacteria, Firmicutes and Actinobacteria can then be fixed via the Calvin cycle [45,47]. In aerobes that use coxL enzymes, CO is typically a source of electrons for respiratory processes coupled with exogenous electron acceptors such as oxygen [45,49], sulfate [50], anthraquinone disulfonate and fumarate [51].

Various lines of evidence point to the importance of CO in primordial metabolism [52–56]. Here, we queried large and well curated biochemical databases – KEGG and BRENDA – to investigate the number and nature of entry points of CO and CO_2 into metabolism.

Results and Discussion

CO_2 is everywhere in metabolism, CO is rare

The KEGG and BRENDA have different reaction nomenclatures, therefore to compare their content it is convenient to use Enzyme Commission (EC) numbers, which also link metabolic data with taxonomy and other catalysis metadata. Figure 1 shows the content of both databases regarding enzyme classes that use CO_2 and CO. Both databases reveal that CO is very rare in metabolism, whereas CO_2 is very common. KEGG contained 390 EC numbers involving CO_2, BRENDA Natural (a subset of BRENDA including only reactions tested *in vivo*) contained 323 EC numbers and both databases returned 13 EC numbers involving CO (Fig. 1).

The results obtained from both databases are not completely overlapping (Fig. 1). In KEGG, there are 91 EC numbers involving CO_2 that are not found in BRENDA. Conversely, 25 EC numbers involving CO_2 are found in BRENDA but not in KEGG. Regarding CO, each database has two unique EC numbers: in

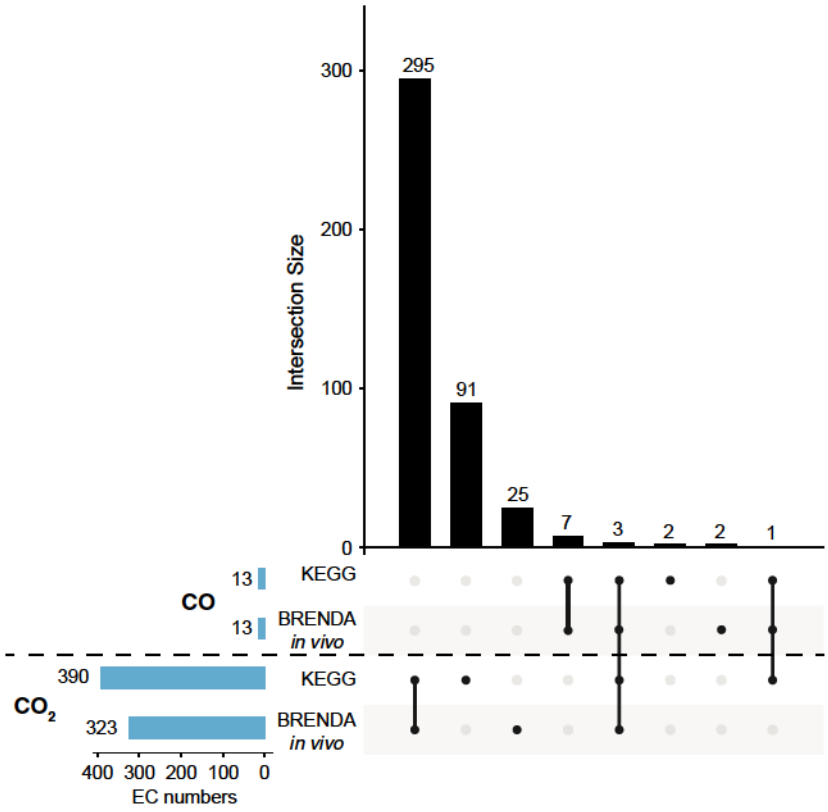


Fig. 1. Enzyme Commission (EC) numbers involving CO_2 and CO in KEGG and BRENDA (only *in vivo* reactions) and their overlaps. The horizontal bars display the total of EC numbers for each molecule in each database. The vertical bars display the size of the overlaps (intersections) between the databases.

BRENDA an additional dioxygenase, 1.13.11.54, and an additional heme oxygenase, 1.14.99.48, are listed as producing CO. In KEGG one of these non-overlapping EC numbers is a misannotation: 2.1.1.258, a 5-methyltetrahydrofolate: corrinoid/iron-sulfur protein Co-methyltransferase, represents a reaction of the Wood-Ljungdahl pathway that does not involve CO as substrate or product [29]. The second CO involving reaction unique to KEGG is EC 4.1.99.5, an O_2-dependent aldehyde oxygenase. O_2-dependent reactions cannot be primordial, because O_2 is the product of cyanobacterial metabolism (see Conclusion).

Eleven EC numbers that involve CO occur in both databases. Of those 11, seven entail CO only as a by-product of an O_2-dependent enzyme: two heme oxygenases, 1.14.14.18 and 1.14.15.20; four dioxygenases: 1.13.11.24, 1.13.11.47, 1.13.11.48 and 1.13.11.53 and one synthase 4.1.99.17. The remaining four EC numbers involving CO all trace directly to CODH. The first is 1.2.2.4, aerobic CODH with cytochrome b-561 as an electron acceptor. This reaction is disputed, however, as some authors argue that no cytochromes are involved in the aerobic CODH reaction [57], contrary to the original proposal [49]. The second is 1.2.5.3, aerobic CODH with quinones as an electron acceptor. The third is EC. 1.2.7.4, anaerobic CODH with ferredoxin. The fourth is 2.3.1.169, the CODH/ACS combined reaction, which in BRENDA is considered as including only the second step of acetyl-CoA synthesis and not the CO_2 fixation step.

CO_2 for all trades, CO only for CODH

CO_2 is involved throughout all major functional pathways in KEGG, while CO is assigned to only 7 (Fig. 2A). Each EC number had from 0 to a maximum of 11 KEGG pathways assigned. Multifunctionality is a known and important characteristic of enzymes, so the functional analysis done here preserved all classifications assigned to all enzymes, except for the large generalist categories ('Biosynthesis of antibiotics', 'Biosynthesis of secondary metabolites', 'Microbial metabolism in diverse environments' and 'Metabolic pathways'), which were discarded. A large number of enzymes do not have any pathways assigned (not shown in the plot) - 113 involving CO_2 and 5 involving CO. The functions of these 5 EC numbers involving CO were searched manually in the literature (see legend of Fig. 2; Table 1). All EC numbers involving CO as a substrate are assigned (or, if assigned Unknown, could be manually assigned) to 'carbon

Fig. 2. Functional analysis of EC numbers involving CO_2 and CO. (A) EC numbers involving CO_2 (dark and light blue for those in prokaryotes and in eukaryotes only, respectively) and CO (dark and light red, accordingly). (B) Enrichment analysis (Fisher's exact test with adjusted p-values by the Bonferroni correction) for high-level functional categories of EC numbers involving CO_2 (prokaryotes only).

fixation pathways in prokaryotes' and 'energy metabolism' through the CODH reaction. Other pathways involve CO always as a byproduct, with the exception of the additional assignments of anaerobic CODH 1.2.7.4 to 'methane metabolism' (in the methanogen pathway) and (through a reaction that does not involve CO) to 'nitrotoluene degradation'.

Each reaction in our set was annotated in a range from 1 to a maximum of 4320 taxa (species) in KEGG as per its occurrence, either through the corresponding gene in KEGG genomes or manually assigned upon examination of the literature (see Materials and methods). Out of the total 399 EC numbers gathered for CO_2 and CO, 99 reactions were found to be annotated only in eukaryotes, only one of which involves CO (a mammalian heme oxygenase that produces CO, 1.14.14.18, annotated in 110 KEGG genomes). In prokaryotes, 292 ECs involved with CO_2 were annotated, versus only 12 with CO.

To investigate the distribution of reactions across pathways where CO_2 was involved, a higher-level categorization was performed, using the KEGG pathway hierarchy, for prokaryotic EC numbers (Fig. 2B). A Fisher's exact test for enrichment of each pathway indicates amino-acid metabolism as highly enriched for CO_2–involving reactions, as well as a significant enrichment for xenobiotics biodegradation and metabolism, carbohydrate metabolism, metabolism of cofactors and vitamins and energy metabolism (adjusted p-values of 2.49×10^{-14}, 6.01×10^{-6}, 4.44×10^{-5}, 3.21×10^{-4} and 7.10×10^{-4}, respectively).

The ubiquity of CO_2 in metabolism is more clearly seen by highlighting CO_2-dependent reactions on the KEGG map 'Metabolic Pathways' (Fig. S1; portion

Table 1. Enzyme commission numbers associated with carbon monoxide.

EC number	Functional classifications in KEGG	Name and description	Proposed functional classification
1.2.2.4	Unknown	CO dehydrogenase (cytochrome b_{561}); although present in strict aerobes, O_2 is not required for the reaction. CO is oxidized to CO_2 with water as the oxidant [49]	Energy Metabolism; Carbon fixation pathways in prokaryotes
1.2.5.3	Unknown	Aerobic Carbon Monoxide dehydrogenase (quinone)	Energy Metabolism; Carbon fixation pathways in prokaryotes
1.2.7.4	Carbon fixation pathways in prokaryotes; Methane metabolism; Nitrotoluene degradation	Anaerobic carbon-monoxide dehydrogenase (ferredoxin)	
2.3.1.169	Carbon fixation pathways in prokaryotes	CO-methylating acetyl-CoA synthase	
2.1.1.258	Carbon fixation pathways in prokaryotes	5-methyltetrahydrofolate:corrinoid/iron-sulfur protein Co-methyltransferase; two step reaction: Tetrahydrofolate + acetyl-CoA ↔ 5-methyltetrahydrofolate + CoA + CO	
1.13.11.24	Unknown	Quercetin 2,3-dioxygenase. CO is a byproduct in this reaction.	Xenobiotics biodegradation and metabolism
1.13.11.47	Unknown	3-hydroxy-4-oxoquinoline 2,4-dioxygenase. CO is a byproduct in this reaction.	Xenobiotics biodegradation and metabolism
1.13.11.48	Unknown	3-hydroxy-2-methylquinolin-4-one 2,4-dioxygenase. CO is a byproduct in this reaction.	Xenobiotics biodegradation and metabolism
1.13.11.53	Cysteine and methionine metabolism	Acireductone dioxygenase (Ni^{2+}-requiring). CO is a byproduct in this reaction. Unknown function; the same enzyme, when binding iron, is the one leading to the salvage of methionine (EC 1.13.11.54) [97,98].	
1.14.14.18	Porphyrin and chlorophyll metabolism	Heme oxygenase (biliverdin-producing). CO is a byproduct in this reaction.	
1.14.15.20	Porphyrin and chlorophyll metabolism	Heme oxygenase (biliverdin-producing, ferredoxin). CO is a byproduct in this reaction.	
4.1.99.5	Cutin, suberine and wax biosynthesis	Aldehyde oxygenase (deformylating)	
4.1.99.17	Thiamine metabolism	Phosphomethylpyrimidine synthase. CO is a byproduct in this reaction.	

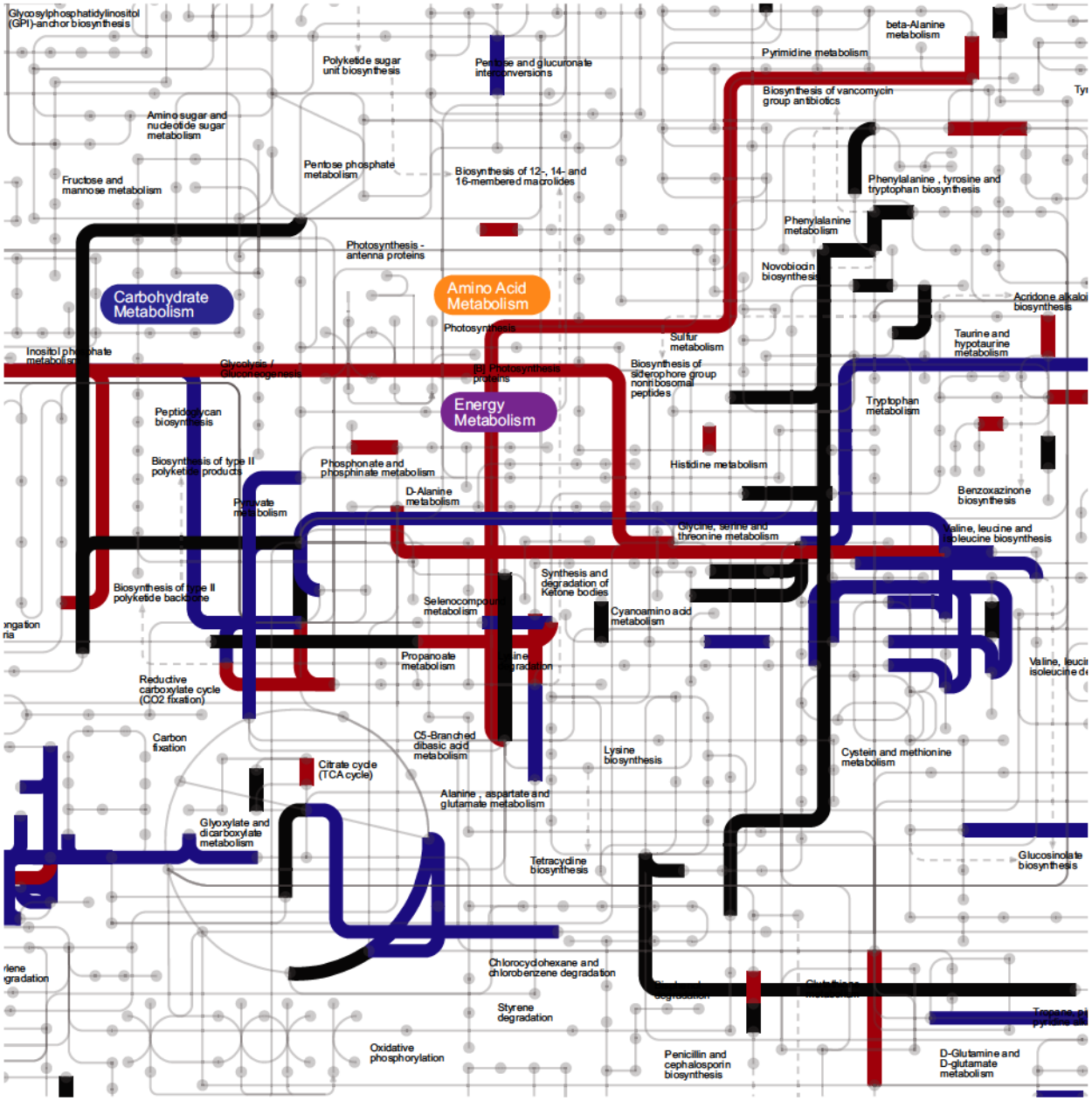


Fig. 3. CO_2 in a section of a global metabolic map (full map provided as Fig. S1). A portion of the KEGG map '01100 – metabolic pathways' with reactions involving CO_2 highlighted, portraying different directionality and reversibility assignments in BRENDA. In black, reactions not in BRENDA or where CO_2 is a product in BRENDA but reversibility is unknown; in blue, reactions where CO_2 is a substrate or it is a product and the reaction is classified as reversible in at least one study; in red, reactions classified as irreversible where CO_2 is a product.

shown in Fig. 3). The KEGG metabolic map used is the largest available for depiction, however it can still only plot 40% of the 292 prokaryotic EC numbers for CO_2. The chemical reactions in the map are all theoretically reversible, however not all can be reversibly catalyzed by the same enzyme under the same physiological conditions. We cross-checked all KEGG EC numbers involving CO_2 against the information regarding reversibility in BRENDA. KEGG reactions have no direct information regarding reversibility – all are assigned as reversible. Reversibility information in BRENDA is two-fold: (a) there is a direct assignment of compounds involved in the reaction as substrates or products and (b) each reaction assigned to an EC has an

independent reversibility classification, which is assigned manually by the database curators as 'reversible', 'irreversible' or 'unknown'. In Figs S1 and 3, EC numbers involving CO_2 are highlighted according to the reversibility of the reactions they encode in BRENDA. When looking at all reactions, including those not in the map, 65% were classified as producing CO_2 with reversibility unknown; 24% as utilizing CO_2 or reversibly producing it; 11% as irreversibly producing CO_2.

Metals and cofactors

The cofactors and metals involved in CO and CO_2 metabolism are different. We analyzed the number of studies reporting metal and cofactor utilization in BRENDA *in vivo*, only for EC numbers where CO and CO_2 are assigned as substrates, and only for anaerobic, prokaryotic reactions (Fig. 4).

For CO, 63 entries for metal utilization were retrieved linked with the only EC number where CO is an *in vivo* substrate in BRENDA (anaerobic CODH reaction, 1.2.7.4). Nickel and iron are by far the most commonly reported metals, occurring in 46 and 34.9% of all 63 entries, respectively. For CO_2, of the total 499 entries retrieved, magnesium and manganese are by far the preferred metals – 33.9 and 14.2%, respectively). Regarding organic cofactors, ATP, biotin and NADs are the most common for CO_2 utilization – 33.2, 26 and 20%, respectively from a total of 235 entries – whereas for CO nickel-iron-sulfur clusters are the only reported cofactors. Different types of Ni-Fe-S clusters have been synthesized in the laboratory [58,59], although none have yet been shown to catalyze the interconversion of CO_2 and CO.

CODH/ACS: Archaea and bacteria, but not aerobes

An earlier paper plotted the evolutionary distribution of the archaeal type and bacterial type CODH and ACS enzymes across genomes [16] demonstrating the antiquity of the enzyme. Single gene phylogenies also trace CODH and ACS to the universal common ancestor [15,26,27]. A fundamental limitation to gene phylogenies as a proxy of prokaryotic gene evolution is however that phylogenies only show in which lineages the gene is present, not the lineages in which it is missing. Plots of gene distributions reveal where genes are lacking. Figure 5A shows the current gene distribution at the prokaryotic phylum level for CODH and ACS as proxies for capacity to harness CO in metabolism and to fix it as acetyl-CoA. As the query sequences, homologues from eight prokaryotes were used to obtain insights into the distribution of the catalytic domains.

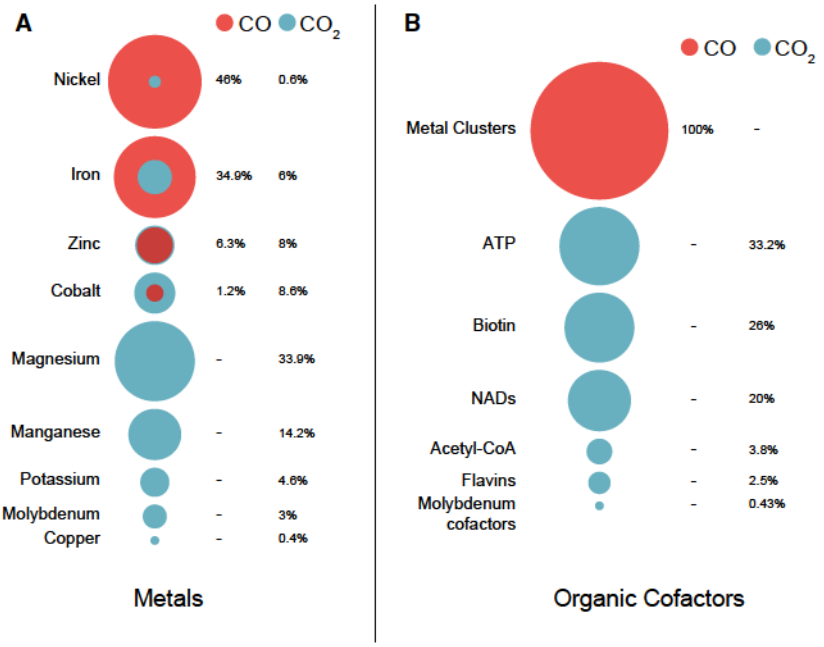


Fig. 4. Metals and organic cofactors in reactions that consume CO or CO_2. Percentage of entries of experimental evidence in BRENDA demonstrating the participation of different (A) metals and (B) cofactors in the catalytic activity of enzymes that use CO (red) or CO_2 (blue) as substrates.

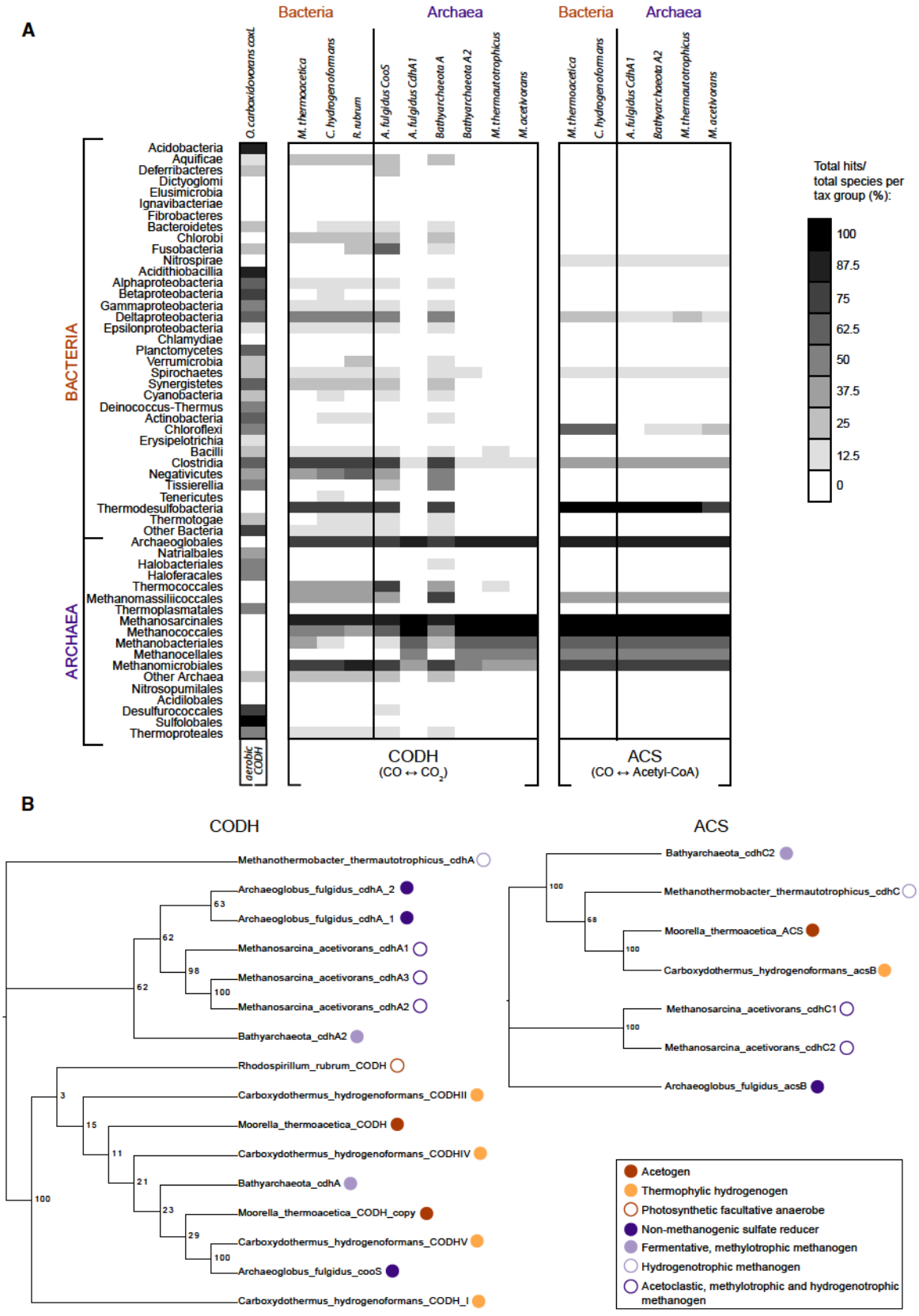


Fig. 5. Phylogenomic analysis of CO-interconverting enzymes. (A) Distribution of genes encoding the CODH and ACS reactions. The left part of the figure lists the taxonomic groups from 5655 completed sequenced genomes (212 archaeal and 5443 bacterial). The presence-absence patterns (PAPs) represent the proportion of genomes within a taxonomic group where each gene is present according to the discrete grey scale-bar of binned intervals (top right, value indicates upper value of each bin). Each column represents a different gene selected from a different query species capable of performing the aerobic CODH (oxidative) reaction, the anaerobic CODH or both the (anaerobic) CODH and the ACS reactions (*Oligotropha carboxidovorans*, *Moorella thermoacetica*, *Carboxydothermus hydrogenoformans*, *Rhodospirillum rubrum*, *Archaeoglobus fulgidus*, Candidatus *Bathyarchaeota archaeon* BA1, *Methanothermobacter thermautotrophicus* and *Methanosarcina acetivorans*). Homologous proteins were predicted by BLAST with an *E*-value threshold of 10^{-5} and filtering for global amino acid identities of at least 20% with Powerneedle (see Materials and methods). (B) Phylogenetic trees of the query sequences of CODH (left) and ACS (right) used to BLAST the RefSeq Database to build the PAPs in (A), numbers at branches are bootstrap values. Metabolic modes of the different species are marked in front of the respective sequences with colored circles according to the legend (bottom right).

The main observation from Fig. 5 is that CODH and ACS are typically distributed among anaerobic autotrophs. Some diversity is seen in the bacterial copy of the enzyme – the presence/absence patterns obtained with the different queries are not fully identical – indicating divergence after duplication. This contrasts with the archaeal forms, with one interesting exception. In both *A. fulgidus* and Bathyarchaeota, there is one copy of CODH with the same distribution as the bacterial CODH. This suggests interdomain lateral gene transfer for this CODH subunit (Fig. 5B). Gene transfers from bacteria to archaea are very common in evolution [60,61]. The distribution of ACS is clearer, and uniform within both domains, showing some homology in-between domains for the clostridial enzymes and Methanomicrobiales (Fig. 5).

CO forms stronger bonds to metals than CO_2

The large difference in the numbers of metabolic reactions that involve either the utilization or production of CO and CO_2 is striking. A closer look at the chemistry regarding the interaction of both compounds with metals provides further detail (Fig. 6). The orbitals in carbon atoms of both carbon monoxide and carbon dioxide are sp-hybridized, such that both molecules are linear. At the level of electron configurations, however, CO and CO_2 differ quite noticeably. In particular, the free electron pair of CO enables two complementing mechanisms (σ and π) that lead to very strong and short bonds with metals (Fig. 6A). The empty π-orbitals of CO support backbonding with metals, which results in a very high affinity to nickel and iron in particular [62]. The high affinity of CO for nickel leads to the facile formation of nickel carbonyl, $Ni(CO)_4$ (a volatile liquid), which formed the basis of the Mond process, an early method for industrial nickel preparation [63]. The strong affinity of CO to transition metals is the basis of its extreme toxicity to humans, it bonds with the iron in hemoglobin more strongly than does O_2.

By contrast, there are various bonding modes of CO_2 to transition metals (Fig. 6B) which depend mostly on whether the metal is rich or poor in electrons. In general, the bonds that CO_2 forms with metals are not as strong as those formed by CO. This can be a virtue in metabolism, as the rather weak bonds of CO_2 to metals permit faster and more versatile catalytic reactions than those of CO. Nevertheless, the special bond between CO and transition metals also enables carbonyl insertion, both in industrial chemistry (heterogenic catalysis) [64], and in one very ancient and important biological reaction – CODH/acetyl-CoA synthase [65], which requires the essential Ni-Fe-S cluster for achieving the slow reduction of CO_2 to CO. Recent studies showing that CO_2 is efficiently reduced by native metals to acetyl and pyruvoyl moieties entail metal bound carbonyl groups and carbonyl insertions in the proposed reaction mechanisms [23]. This parallels the Fischer-Tropsch type reaction mechanisms suggested for geochemical CO_2 reduction processes giving rise to abiotic organic molecules in hydrothermal vents [66,67].

Conclusion

For soil environments, it has been estimated that 0.2 gigatonnes (Gt) of CO is consumed each year globally [68] mainly through CO aerobic oxidation [69]. During methanogenesis in anoxic environments [70], about 0.6 Gt of CH_4 is produced annually from acetate [71,72], a process that generates one mol of CO as a pathway intermediate per mol of acetate cleaved [71,72], corresponding to roughly 1 Gt methanogenesis-dependent CO synthesis per year. Based on reviews of CO metabolism [22,37], and on our metabolic database search, it appears that CO interfaces with metabolism (the biotic segment of the carbon cycle) at only two enzymes: the anaerobic CODH, a Ni- and Fe-containing enzyme, and

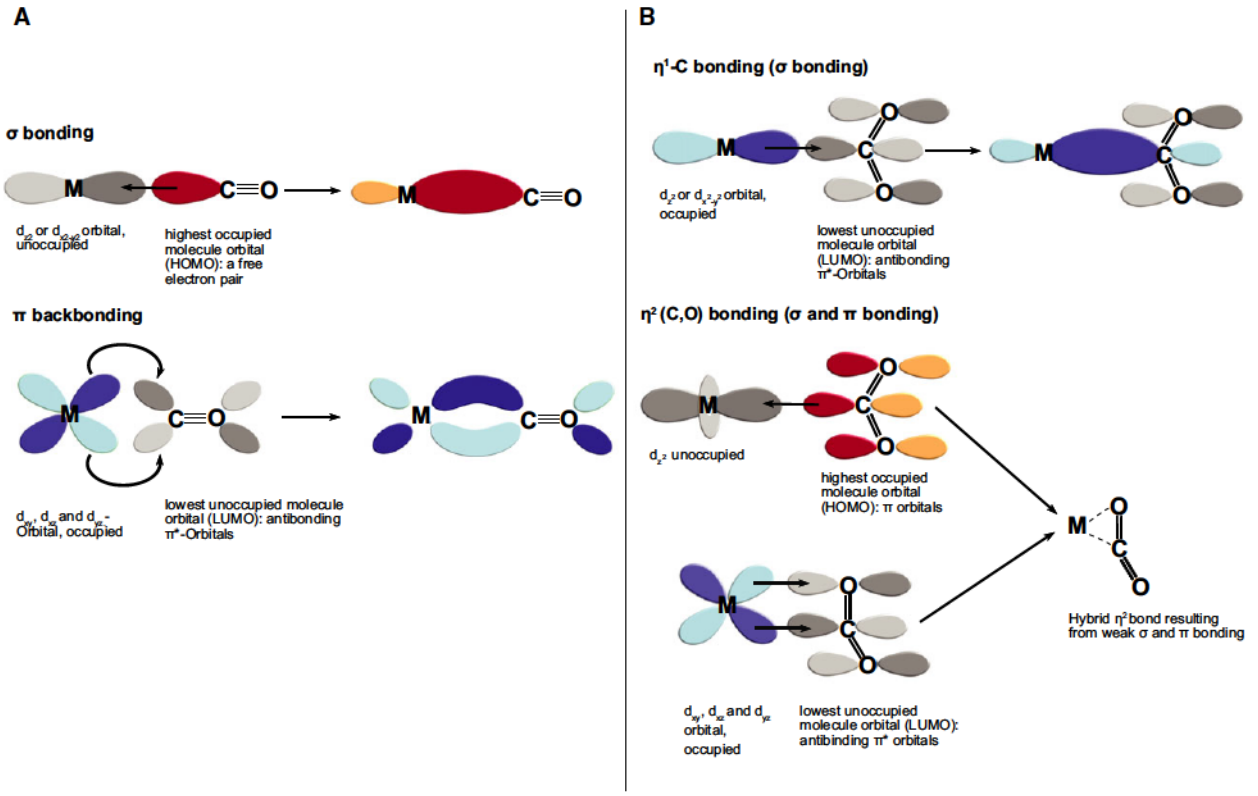


Fig. 6. CO (A) and CO_2 (B) bonding to transition metals. (A) CO binds to a transition metal (M) via the free electron pair of its carbon atom. The electron density in this orbital (red = positive phase, yellow = negative phase) can be placed into empty metal d orbitals forming a σ bond. Concurrently, a π bond is formed between an occupied d orbital and the antibonding empty π* orbital of CO (darker grey = positive phase, lighter grey = negative phase), so called 'π backbonding'. (B) Different bonding modes between CO_2 and transition metals include η^1-C coordination, which mostly happens with electron-rich metals (i.e. lower oxidation states), as they can transfer charge from the d_{z^2} orbitals to the antibonding π* orbitals of CO_2. A double bond-like interaction (dashed line) can also occur between a transition metal, carbon and oxygen, η^2-(C,O) bonding: an empty d_{z^2} orbital of a metal can take electron density from the π orbital of the CO_2 orbital (red/yellow), while electron density can also be transferred from occupied d orbitals (blue) into the antibonding π* orbitals of CO_2 (comparable to the backbonding of CO, but weaker).

the aerobic CODH, a Mo- and Cu-containing enzyme. Although they catalyze the same reversible reaction (Eq. 1),

$$CO + H_2O \rightarrow CO_2 + 2e^- + 2H^+ \qquad \Delta G'_0 = -20\,\mathrm{kJ \cdot mol^{-1}} \quad (1)$$

the aerobic and anaerobic CODH enzymes have different subunit structures, different cofactors and are not related at the amino acid sequence or structural level [37,39–41,45]. Prokaryotes that use aerobic CODH use CO as a source of electrons in energy metabolism, are typically aerobes or facultative aerobes and tend to transfer the electrons from CO to high potential acceptors such as O_2 or acceptors derived from it such as nitrate (NO_3^-) [45,73]. Because O_2 is a product of cyanobacterial metabolism [74], such high potential acceptors are latecomers in evolution, as current geochemical data have it that cyanobacterial O_2 first appeared about 2.5 billion years ago [74–76].

From the standpoint of thermodynamics, it is well-known that the WL pathway is the most favorable of the six known CO_2 fixation pathways [3,77]:

$$2CO_2 + 4H_2 + \mathrm{CoASH} \rightarrow \mathrm{CH_3COSCoA} + 3H_2O \qquad \Delta G'_0 = -59\,\mathrm{kJ \cdot mol^{-1}} \quad (2)$$

The reaction is exergonic when H_2 is the electron donor, which allows some acetogens and some methanogens to generate ion gradients and ATP at the expense of CO_2 fixation. All other pathways of CO_2 fixation require ATP hydrolysis to go forward. Recent findings show that the reverse citric acid (rTCA) cycle in some thermophiles requires hydrolysis of only one

ATP to go forward [78,79], but that ATP must still be generated by an independent energy metabolism. The reductive acetyl-CoA pathway is simultaneously a source of carbon and energy, a strong argument in favor of its ancestral status among carbon assimilation pathways [2,3]. CO_2-fixation via the rTCA cycle could have arisen via closure of the incomplete (horseshoe) version of the rTCA cycle [80] (starting from acetyl-CoA supplied by the WL-pathway) as it occurs in some acetogens and methanogens [18,81,82].

Its linear nature, chemical simplicity, favorable energetics, and occurrence among both Bacteria and Archaea set it apart from other pathways of CO_2 fixation and suggest that the WL is the most ancient of CO_2 fixation pathways [3,4]. Strong evidence supporting the antiquity of the WL pathway comes from new findings showing that its main reactions are facile, with its central intermediates including pyruvate arising spontaneous in laboratory reactions overnight from CO_2 and water at temperatures of 30–100 °C in the absence of enzymes, with native metals such as Fe^0 and Ni^0 functioning as catalysts and reductants [23]. From the standpoint of energetics, there is something very special about the reductive acetyl-CoA pathway among metabolic pathways. The involvement of CO as a reaction intermediate able to undergo carbonyl insertion might be the essential property that renders Ni-dependent C—C bond formation in the CODH/ACS reaction mechanism apparently immune to substitution by organic cofactors or alternative enzymes over the last 4 billion years. In physiological evolution, it appears that there is something very special about CO.

Materials and methods

Data retrieval and integration

Both KEGG and BRENDA databases were scanned for classes of reactions involving CO and/or CO_2 by parsing Enzyme Commission (EC) numbers. From BRENDA, we took only the subset of reactions tested *in vivo*. EC numbers involving bicarbonate (HCO_3^-) were also retrieved. Because of the chemical equilibrium between CO_2 and HCO_3^- and their rapid interconversion by carbonic anhydrases [83], which are widely distributed enzymes, throughout this work CO_2 and HCO_3^- were considered to be identical in database parsing procedures. EC numbers and the current list of KEGG organisms with the corresponding taxonomic classification were downloaded using the KEGG Rest API (http://www.kegg.jp/kegg/rest/keggapi.html), July 2017. The EC numbers from BRENDA were retrieved with the SOAP API Python interface. All integration was performed with Python scripts.

Taxonomy annotation

The taxonomic assignment of EC numbers was retrieved from the annotated genomes in the KEGG database. Among all 399 KEGG EC numbers used, 114 had no gene associated, and these were manually checked: for each EC number we checked the original literature linked in the KEGG entry to find the corresponding taxon where the EC number was identified. For 53 out of these 114 EC numbers, the taxon retrieved from the literature was not present in KEGG genomes. In these cases, a close phylogenetic cousin was assigned to the EC number so that it could be automatically assigned to the Prokaryotic or Eukaryotic domains.

Statistical analysis and metabolic maps

All the statistical analyses, including the Fisher's exact test for significance and Bonferroni correction, were performed with the package RPy2, that provides an interface between Python and the R statistical software. Overlapping sets of EC numbers were analyzed and plotted with UpSetR [84]. The metabolic map with highlighted reactions was produced with iPath v2.0 [85].

Analysis of distributions of CO enzymes

The query sequences for the catalytic domain of CODH and the catalytic domain of ACS were manually selected from nine different species (four archaea and five bacteria) that have been studied with respect to CO utilization. All annotated copies for both genes were taken for each genome. This exercise resulted in the collection of a total of 25 queries from: (bacterial) an acetogen, *Moorella thermoacetica*, with two copies of CODH and one copy of ACS [86]; a thermophilic hydrogenogen, *Carboxydothermus hydrogenoformans* with four CODH copies and one ACS [87]; a photosynthetic facultative anaerobe, *Rhodospirillum rubrum* with a single copy of CODH, capable of growth on carbon monoxide as sole energy source [88]; two aerobes with one CODH each, *Oligotropha carboxidovorans* (coxL I) and *Bradyrhizobium* sp. CPP (coxL II) [89] – the latter with a similar pattern to the former (data not shown); (archaeal) a non-methanogenic sulfate reducer, *Archaeoglobus fulgidus*, with 3 copies of CODH and one ACS [90]; a recently identified, fermentative and possibly methylotrophic methanogen, Candidatus Bathyarchaeota archaeon BA1 with two copies of CODH and one of ACS [91,92]; one hydrogenotrophic methanogen, *Methanothermobacter thermautotrophicus* with one copy of each enzyme [93] and finally an acetoclastic methylotrophic, hydrogenotrophic methanogen, *Methanosarcina acetivorans* with three copies of CODH and two of ACS [94]. Representative queries were taken from each genome when they were significantly similar. The queries were

aligned with ClustalW [95] and phylogenetic inferences were made with RAxML [96].

To characterize CODH and ACS gene distribution, a BLAST search was performed against all prokaryotic genomes in RefSeq (NCBI, version September 2016), of which the primary hits (*e*-value $\leq 1 \times 10^{-5}$) were selected. A pairwise global 'Needleman & Wunsch' – alignment was then performed with these sequences against the whole database of prokaryotes again to filter for hits with global identity >20%.

Acknowledgements

This work was supported by grants from the European Research Council (666053), the Volkswagen Foundation (93 046) to WFM and by a cooperation grant from the Deutsche Forschungsgemeinschaft to WFM (MA1426/21-1) and to Harun Tüysüz (TU 315/8-1), Max Planck Institute for Coal Research, Mülheim.

Author contributions

JCX collected and analyzed the data from KEGG and BRENDA. MP analyzed the chemical configurations of CO_2 and CO. WFM designed and supervised the study. The manuscript was written and proofread by all authors.

References

1 Tabita FR (1988) Molecular and cellular regulation of autotrophic carbon dioxide fixation in microorganisms. *Microbiol Rev* **52**, 155–189.
2 Berg IA (2011) Ecological aspects of the distribution of different autotrophic CO_2 fixation pathways. *Appl Environ Microbiol* **77**, 1925–1936.
3 Fuchs G (2011) Alternative pathways of carbon dioxide fixation: insights into the early evolution of life? *Annu Rev Microbiol* **65**, 631–658.
4 Hügler M & Sievert SM (2011) Beyond the Calvin cycle: autotrophic carbon fixation in the ocean. *Ann Rev Mar Sci* **3**, 261–289.
5 Appel AM, Bercaw JE, Bocarsly AB, Dobbek H, DuBois DL, Dupuis M, Ferry JG, Fujita E, Hille R, Kenis PJA *et al.* (2013) Frontiers, opportunities, and challenges in biochemical and chemical catalysis of CO_2 fixation. *Chem Rev* **113**, 6621–6658.
6 Jajesniak P, Eldin H, Omar M & Wong TS (2014) Carbon dioxide capture and utilization using biological systems : opportunities and challenges. *Bioprocess Biotech* **4**, 15.
7 Gong F, Liu G, Zhai X, Zhou J, Cai Z & Li Y (2015) Quantitative analysis of an engineered CO_2-fixing *Escherichia coli* reveals great potential of heterotrophic CO_2 fixation. *Biotechnol Biofuels* **8**, https://doi.org/10.1186/s13068-015-0268-1.
8 Antonovsky N, Gleizer S, Noor E, Zohar Y, Herz E, Barenholz U, Zelcbuch L, Amram S, Wides A, Tepper N *et al.* (2016) Sugar synthesis from CO_2 in *Escherichia coli*. *Cell* **166**, 115–125.
9 Schwander T, Schada von Borzyskowski L, Burgener S, Cortina NS & Erb TJ (2016) A synthetic pathway for the fixation of carbon dioxide in vitro. *Science* **354**, 900–904.
10 Claassens NJ (2017) A warm welcome for alternative CO_2 fixation pathways in microbial biotechnology. *Microb Biotechnol* **10**, 31–34.
11 Bolton E, Abelson P & Aldous E (1952) Utilization of carbon dioxide in the synthesis of nucleic acid by *Escherichia coli*. *J Biol Chem* **198**, 179–185.
12 Roslev P, Larsen MB, Jørgensen D & Hesselsoe M (2004) Use of heterotrophic CO_2 assimilation as a measure of metabolic activity in planktonic and sessile bacteria. *J Microbiol Methods* **59**, 381–393.
13 Fuchs G & Stupperich E (1985) Evolution of autotrophic CO_2 fixation. In *Evolution of Prokaryotes. FEMS Symposium No. 29* (Schleifer K & Stackebrandt E, eds), pp. 235–251. Academic Press, London.
14 Russell MJ & Martin W (2004) The rocky roots of the acetyl-CoA pathway. *Trends Biochem Sci* **29**, 358–363.
15 Weiss MC, Sousa FL, Mrnjavac N, Neukirchen S, Roettger M, Nelson-Sathi S & Martin WF (2016) The physiology and habitat of the last universal common ancestor. *Nat Microbiol* **1**, https://doi.org/10.1038/nmicrobiol.2016.116
16 Sousa FL & Martin WF (2014) Biochemical fossils of the ancient transition from geoenergetics to bioenergetics in prokaryotic one carbon compound metabolism. *Biochim Biophys Acta – Bioenerg* **1837**, 964–981.
17 Huber C & Wächtershäuser G (1997) Activated acetic acid by carbon fixation on (Fe, Ni)S under primordial conditions. *Science* **276**, 245–247.
18 Martin W & Russell MJ (2007) On the origin of biochemistry at an alkaline hydrothermal vent. *Philos Trans R Soc B Biol Sci* **362**, 1887–1926.
19 Berg IA, Kockelkorn D, Ramos-Vera WH, Say RF, Zarzycki J, Hügler M, Alber BE & Fuchs G (2010) Autotrophic carbon fixation in archaea. *Nat Rev Microbiol* **8**, 447–460.
20 Schrenk MO, Brazelton WJ & Lang SQ (2013) Serpentinization, carbon, and deep life. *Rev Mineral Geochemistry* **75**, 575–606.
21 McCollom TM (2016) Abiotic methane formation during experimental serpentinization of olivine. *Proc Natl Acad Sci USA* **113**, 13965–13970.
22 Ragsdale SW & Pierce E (2008) Acetogenesis and the Wood-Ljungdahl pathway of CO_2 fixation. *Biochim Biophys Acta – Proteins Proteomics* **1784**, 1873–1898.

23 Varma SJ, Muchowska KB, Chatelain P & Moran J (2018) Native iron reduces CO_2 to intermediates and end-products of the acetyl-CoA pathway. *Nat Ecol Evol* **2**, 1019–1024.

24 Sousa FL, Preiner M & Martin WF (2018) Native metals, electron bifurcation, and CO_2 reduction in early biochemical evolution. *Curr Opin Microbiol* **43**, 77–83.

25 Buckel W & Thauer RK (2018) Flavin-based electron bifurcation, ferredoxin, flavodoxin, and anaerobic respiration with protons (Ech) or NAD+(Rnf) as electron acceptors: a historical review. *Front Microbiol* **9**, https://doi.org/10.3389/fmicb.2018.00401

26 Adam PS, Borrel G & Gribaldo S (2018) Evolutionary history of carbon monoxide dehydrogenase/acetyl-CoA synthase, one of the oldest enzymatic complexes. *Proc Natl Acad Sci USA* **115**, E1166–E1173 (erratum appears in Proc Natl Acad Sci USA 115, E5836–E5837).

27 Techtmann SM, Lebedinsky AV, Colman AS, Sokolova TG, Woyke T, Goodwin L & Robb FT (2012) Evidence for horizontal gene transfer of anaerobic carbon monoxide dehydrogenases. *Front Microbio* **3**, https://doi.org/10.3389/fmicb.2012.00132

28 Vorholt J, Kunow J, Stetter KO & Thauer RK (1995) Enzymes and coenzymes of the carbon monoxide dehydrogenase pathway for autotrophic CO_2 fixation in *Archaeoglobus lithotrophicus* and the lack of carbon monoxide dehydrogenase in the heterotrophic *A. profundus*. *Arch Microbiol* **163**, 112–118.

29 Ragsdale SW (2008) Enzymology of the Wood-Ljungdahl pathway of acetogenesis. *Ann N Y Acad Sci* **1125**, 129–136.

30 Ragsdale SW (2004) Life with carbon monoxide. *Crit Rev Biochem Mol Biol* **39**, 165–195.

31 Sokolova TG, Henstra A-M, Sipma J, Parshina SN, Stams AJM & Lebedinsky AV (2009) Diversity and ecophysiological features of thermophilic carboxydotrophic anaerobes. *FEMS Microbiol Ecol* **68**, 131–141.

32 Diender M, Stams AJM & Sousa DZ (2015) Pathways and bioenergetics of anaerobic carbon monoxide fermentation. *Front Microbiol* **6**, https://doi.org/10.3389/fmicb.2015.01275

33 Borrel G, Adam PS & Gribaldo S (2016) Methanogenesis and the Wood-Ljungdahl pathway: an ancient, versatile, and fragile association. *Genome Biol Evol* **8**, 1706–1711.

34 Schuchmann K & Müller V (2014) Autotrophy at the thermodynamic limit of life: a model for energy conservation in acetogenic bacteria. *Nat Rev Microbiol* **12**, 809–821.

35 Kopke M, Held C, Hujer S, Liesegang H, Wiezer A, Wollherr A, Ehrenreich A, Liebl W, Gottschalk G & Durre P (2010) *Clostridium ljungdahlii* represents a microbial production platform based on syngas. *Proc Natl Acad Sci USA* **107**, 13087–13092.

36 Jansen K, Thauer RK, Widdel F & Fuchs G (1984) Carbon assimilation pathways in sulfate reducing bacteria. Formate, carbon dioxide, carbon monoxide, and acetate assimilation by *Desulfovibrio baarsii*. *Arch Microbiol* **138**, 257–262.

37 Oelgeschläger E & Rother M (2008) Carbon monoxide-dependent energy metabolism in anaerobic bacteria and archaea. *Arch Microbiol* **190**, 257–269.

38 Bender G, Pierce E, Hill JA, Darty JE & Ragsdale SW (2011) Metal centers in the anaerobic microbial metabolism of CO and CO_2. *Metallomics* **3**, 797–815.

39 Can M, Armstrong FA & Ragsdale SW (2014) Structure, function, and mechanism of the nickel metalloenzymes, CO dehydrogenase, and acetyl-CoA synthase. *Chem Rev* **114**, 4149–4174.

40 Gregg CM, Goetzl S, Jeoung JH & Dobbek H (2016) AcsF catalyzes the ATP-dependent insertion of Nickel into the Ni, Ni-[4Fe4S] cluster of Acetyl-CoA synthase. *J Biol Chem* **291**, 18129–18138.

41 Doukov TI, Iverson TM, Seravalli J, Ragsdale SW & Drennan CL (2002) A Ni-Fe-Cu center in a bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase. *Science* **298**, 567–572.

42 Maynard EL & Lindahl PA (1999) Evidence of a molecular tunnel connecting the active sites for CO_2 reduction and acetyl-CoA synthesis in acetyl-CoA synthase from *Clostridium thermoaceticum*. *J Am Chem Soc* **121**, 9221–9222.

43 Seravalli J & Ragsdale SW (2000) Channeling of carbon monoxide during anaerobic carbon dioxide fixation. *Biochemistry* **39**, 1274–1277.

44 Gong W, Hao B, Wei Z, Ferguson DJ, Tallant T, Krzycki JA & Chan MK (2008) Structure of the alpha2epsilon2 Ni-dependent CO dehydrogenase component of the *Methanosarcina barkeri* acetyl-CoA decarbonylase/synthase complex. *Proc Natl Acad Sci USA* **105**, 9558–9563.

45 King GM & Weber CF (2007) Distribution, diversity and ecology of aerobic CO-oxidizing bacteria. *Nat Rev Microbiol* **5**, 107–118.

46 Dobbek H, Gremer L, Meyer O & Huber R (1999) Crystal structure and mechanism of CO dehydrogenase, a molybdo iron-sulfur flavoprotein containing S-selanylcysteine. *Proc Natl Acad Sci USA* **96**, 8884–8889.

47 Wilcoxen J & Hille R (2013) The hydrogenase activity of the molybdenum/copper-containing carbon monoxide dehydrogenase of *Oligotropha carboxidovorans*. *J Biol Chem* **288**, 36052–36060.

48 Hille R, Dingwall S & Wilcoxen J (2015) The aerobic CO dehydrogenase from *Oligotropha carboxidovorans*. *J Biol Inorg Chem* **20**, 243–251.

49 Meyer O, Jacobitz S & Krüger B (1986) Biochemistry and physiology of aerobic carbon monoxide-oxidizing bacteria. *FEMS Microbiol Rev* **39**, 161–179.

50 Parshina SN, Sipma J, Henstra AM & Stams AJM (2010) Carbon monoxide as an electron donor for the biological reduction of sulphate. *Int J Microbiol* **2010**, https://doi.org/10.1155/2010/319527.
51 Henstra AM & Stams AJM (2004) Novel physiological features of *Carboxydothermus hydrogenoformans* and *Thermoterrabacterium ferrireducens*. *Appl Environ Microbiol* **70**, 7236–7240.
52 King GM (2015) Carbon monoxide as a metabolic energy source for extremely halophilic microbes: implications for microbial activity in Mars regolith. *Proc Natl Acad Sci USA* **112**, 4465–4470.
53 Nava-Sedeño JM, Ortiz-Cervantes A, Segura A & Domagal-Goldman SD (2016) Carbon monoxide and the potential for prebiotic chemistry on habitable planets around main sequence M stars. *Astrobiology* **16**, https://doi.org/10.1089/ast.2015.1435
54 Miyakawa S, Yamanashi H, Kobayashi K, Cleaves HJ & Miller SL (2002) Prebiotic synthesis from CO atmospheres: implications for the origins of life. *Proc Natl Acad Sci USA* **99**, 14628–14631.
55 Aylward N & Bofinger N (2001) The reactions of methanimine and cyanogen with carbon monoxide in prebiotic molecular evolution on earth. *Orig Life Evol Biosph* **31**, 481–500.
56 Cody GD, Boctor NZ, Filley TR, Hazen RM, Scott JH & Yoder HS Jr (2000) The primordial synthesis of carbonylated iron-sulfur clusters and the synthesis of pyruvate. *Science* **289**, 1337–1340.
57 Wilcoxen J, Zhang B & Hille R (2011) Reaction of the molybdenum- and copper-containing carbon monoxide dehydrogenase from *Oligotropha carboxydovorans* with quinones. *Biochemistry* **50**, 1910–1916.
58 Panda R, Zhang Y, McLauchlan CC, Rao PV, Tiago De Oliveira FA, Münck E & Holm RH (2004) Initial structure modification of tetrahedral to planar nickel(II) in a nickel-iron-sulfur cluster related to the C-cluster of carbon monoxide dehydrogenase. *J Am Chem Soc* **126**, 6448–6459.
59 Song LC, Li YL, Li L, Gu ZC & Hu QM (2010) Synthetic and structural investigations of linear and macrocyclic nickel/iron/sulfur cluster complexes. *Inorg Chem* **49**, 10174–10182.
60 Nelson-Sathi S, Sousa FL, Roettger M, Lozada-Chávez N, Thiergart T, Janssen A, Bryant D, Landan G, Schönheit P, Siebers B *et al.* (2015) Origins of major archaeal clades correspond to gene acquisitions from bacteria. *Nature* **517**, 77–80.
61 Wagner A, Whitaker RJ, Krause DJ, Heilers J-H, van Wolferen M, van der Does C & Albers S-V (2017) Mechanisms of gene flow in archaea. *Nat Rev Microbiol* **15**, 492–501.
62 Jitaru M (2007) Electrochemical carbon dioxide reduction - fundamental and applied topics. *J Univ Chem Technol Metall* **42**, 333–344.
63 Mond L, Langer C & Quincke F (1890) Action of carbon monoxide on nickel. *J Chem Soc Trans* **57**, 749–753.
64 Tchougreeff AL, Gulevich YV, Misurkin IA & Beletskaya IP (1993) A model for CO insertion in transition metal complexes. *J Organomet Chem* **455**, 261–270.
65 Evans DJ (2005) Chemistry relating to the nickel enzymes CODH and ACS. *Coord Chem Rev* **249**, 1582–1595.
66 McCollom TM & Seewald JS (2013) Serpentinites, hydrogen, and life. *Elements* **9**, 129–134.
67 McCollom TM (2013) Miller-Urey and beyond: What have we learned about prebiotic organic synthesis reactions in the past 60 years? *Annu Rev Earth Planet Sci* **41**, 207–229.
68 King GM (1999) Characteristics and significance of atmospheric carbon monoxide consumption by soils. *Chemosphere* **1**, 53–63.
69 Conrad R & Seiler W (1980) Role of microorganisms in the consumption and production of atmospheric carbon monoxide by soil. *Appl Environ Microbiol* **40**, 437–445.
70 Thauer RK, Kaster A-K, Seedorf H, Buckel W & Hedderich R (2008) Methanogenic archaea: ecologically relevant differences in energy conservation. *Nat Rev Microbiol* **6**, 579–591.
71 Thauer RK (1998) Biochemistry of methanogenesis: a tribute to Marjory Stephenson. 1998 Marjory Stephenson Prize Lecture. *Microbiology* **144**, 2377–2406.
72 Ferry JG (2010) How to make a living by exhaling methane. *Annu Rev Microbiol* **64**, 453–473.
73 Meyer O & Schlegel HG (1983) Biology of aerobic carbon monoxide-oxidizing bacteria. *Annu Rev Microbiol* **37**, 277–310.
74 Fischer WW, Hemp J & Johnson JE (2016) Evolution of oxygenic photosynthesis. *Annu Rev Earth Planet Sci* **44**, 647–683.
75 Lyons TW, Reinhard CT & Planavsky NJ (2014) The rise of oxygen in Earth's early ocean and atmosphere. *Nature* **506**, 307–315.
76 Allen JF (2016) A proposal for formation of Archaean stromatolites before the advent of oxygenic photosynthesis. *Front Microbiol* **7**, https://doi.org/10.3389/fmicb.2016.01784.
77 Fuchs G (1994) Variations of the acetyl-CoA pathway in diversely related microorganisms that are not acetogens. In *Acetogenesis* (Drake HL, ed.), pp. 507–520. Springer US, Boston, MA.
78 Mall A, Sobotta J, Huber C, Tschirner C, Kowarschik S, Bačnik K, Mergelsberg M, Boll M, Hügler M, Eisenreich W *et al.* (2018) Reversibility of citrate synthase allows autotrophic growth of a thermophilic bacterium. *Science* **359**, 563–567.
79 Nunoura T, Chikaraishi Y, Izaki R, Suwa T, Sato T, Harada T, Mori K, Kato Y, Miyazaki M, Shimamura

S *et al.* (2018) A primordial and reversible TCA cycle in a facultatively chemolithoautotrophic thermophile. *Science* **359**, 559–563.
80 Giovannelli S, Sievert SM, Hügler M, Markert S, Becher D, Schweder T & Vetriani C (2017) Insight into the evolution of microbial metabolism from the deep-branching bacterium, *Thermovibrio ammonificans*. *eLife* **6**, e18990. https://doi.org/10.7554/eLife. 18990
81 Simpson PG & Whitman WB (1993) Anabolic pathways in methanogens. In *Methanogenesis* (Ferry JG, ed.), pp. 445–472. Chapman & Hall Microbiology Series (Physiology/Ecology/Molecular Biology/ Biotechnology). Springer US, Boston, MA.
82 Furdui C & Ragsdale SW (2000) The role of pyruvate ferredoxin oxidoreductase in pyruvate synthesis during autotrophic growth by the Wood-Ljungdahl pathway. *J Biol Chem* **275**, 28494–28499.
83 Tresguerres M, Buck J & Levin LR (2010) Physiological carbon dioxide, bicarbonate, and pH sensing. *Pflügers Arch – Eur J Physiol* **460**, 953–964.
84 Lex A, Gehlenborg N, Strobelt H, Vuillemot R & Pfister H (2014) UpSet: visualization of intersecting sets. *IEEE Trans Vis Comput Graph* **20**, 1983–1992.
85 Yamada T, Letunic I, Okuda S, Kanehisa M & Bork P (2011) iPath2.0: interactive pathway explorer. *Nucleic Acids Res* **39**, W412–W415.
86 Pierce E, Xie G, Barabote RD, Saunders E, Han CS, Detter JC, Richardson P, Brettin TS, Das A, Ljungdahl LG *et al.* (2008) The complete genome sequence of *Moorella thermoacetica* (f. Clostridium thermoaceticum). *Environ Microbiol* **10**, 2550–2573.
87 Wu M, Ren Q, Durkin AS, Daugherty SC, Brinkac LM, Dodson RJ, Madupu R, Sullivan SA, Kolonay JF, Haft DH *et al.* (2005) Life in hot carbon monoxide: the complete genome sequence of *Carboxydothermus hydrogenoformans* Z-2901. *PLoS Genet* **1**, https://doi.org/10.1371/journal.pgen.0010065.
88 Munk AC, Copeland A, Lucas S, Lapidus A, Del Rio TG, Barry K, Detter JC, Hammon N, Israni S, Pitluck S *et al.* (2011) Complete genome sequence of *Rhodospirillum rubrum* type strain (S1T). *Stand Genomic Sci* **4**, 293–302.
89 King GM (2003) Molecular and culture-based analyses of aerobic carbon monoxide oxidizer diversity. *Appl Environ Microbiol* **69**, 7257–7265.
90 Klenk H-P, Clayton RA, Tomb J-F, White O, Nelson KE, Ketchum KA, Dodson RJ, Gwinn M, Hickey EK, Peterson JD *et al.* (1997) The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature* **390**, 364–370.
91 Evans PN, Parks DH, Chadwick GL, Robbins SJ, Orphan VJ, Golding SD & Tyson GW (2015) Methane metabolism in the archaeal phylum *Bathyarchaeota* revealed by genome-centric metagenomics. *Science* **350**, 434–438.
92 Vanwonterghem I, Evans PN, Parks DH, Jensen PD, Woodcroft BJ, Hugenholtz P & Tyson GW (2016) Methylotrophic methanogenesis discovered in the archaeal phylum *Verstraetearchaeota*. *Nat Microbiol* **1**, https://doi.org/10.1038/nmicrobiol.2016.170
93 Smith DR, Doucette-Stamm LA, Deloughery C, Lee H, Dubois J, Aldredge T, Bashirzadeh R, Blakely D, Cook R, Gilbert K *et al.* (1997) Complete genome sequence of *Methanobacterium thermoautotrophicum* deltaH: functional analysis and comparative genomics. *J Bacteriol* **179**, 7135–7155.
94 Galagan JE, Nusbaum C, Roy A, Endrizzi MG, Macdonald P, FitzHugh W, Calvo S, Engels R, Smirnov S, Atnoor D *et al.* (2002) The genome of *M. acetivorans* reveals extensive metabolic and physiological diversity. *Genome Res* **12**, 532–542.
95 Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948.
96 Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313.
97 Dai Y, Wensink PC & Abeles RH (1999) One protein, two enzymes. *J Biol Chem* **274**, 1193–1195.
98 Allpress CJ, Grubel K, Szajna-Fuller E, Arif AM & Berreau LM (2013) Regioselective aliphatic carbon-carbon bond cleavage by a model system of relevance to iron-containing Acireductone dioxygenase. *J Am Chem Soc* **135**, 659–668.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. CO_2 in a global metabolic map. KEGG map '01100 – metabolic pathways' with reactions involving CO_2 highlighted, portraying different directionality and reversibility assignments in BRENDA.

III CO_2 fixation in abiogenesis and beyond

It is established that CO_2 fixation can be considered a transition point between abiotic and biotic processes. Biology hurdles this transition daily, as evident by autotrophic organisms. The central question of this thesis is, however: how did that arise? The working hypothesis is that abiotic, geochemical organic syntheses in geochemical settings such as hydrothermal vents or serpentinizing systems in general (McDermott *et al.*, 2015; McCollom *et al.*, 2016) preceded biological processes. As such, the possible product spectrum within such settings is key.

The three publications within this chapter present the core of this dissertation. They discuss in detail the importance of the physicochemical setting for prebiotic processes with special focus on the importance of heterogeneous catalysis. Publication 6 draws parallels between known industrial processes and possible synthetic pathways within hydrothermal geochemical settings. Publication 7 presents experimental results showing actual observable parallels between the bio- and geochemistry of H_2-dependent CO_2 fixation. Here, segments resembling intermediates and products of biotic pathways are found as the main products of the abiotic reaction. Publication 8 finally provides an outlook on the possibilities of such systems for more complex experimental approaches and also theoretical transition points to more complex protometabolic systems.

Publication 6

Title: Serpentinization: connecting geochemistry, ancient metabolism and industrial hydrogenation

Year: 2018

Authors: Martina Preiner, Joana C. Xavier, Filipa L. Sousa, Verena Zimorski, Anna Neubeck, Susan Q. Lang, H. Chris Greenwell, Karl Kleinermanns, Harun Tüysüz, Tom M. McCollom, Nils G. Holm, and William F. Martin

Published in: *Life,* 8(4), 41.
doi: 10.3390/life8040041.

Contribution: First and shared corresponding author.
Major: wrote substantial sections of the paper, edited heavily throughout all review processes and coordinated the distribution and responses with all authors.

Summary: This article points out the striking parallels between biochemical, geochemical and industrial CO_2 fixation. Minerals such as magnetite are presented as a common denominator between geologically abundant and industrially applied catalysts, expanding the possibilities of mechanistic and product oriented studies within prebiotic chemical research.

life

MDPI

Review

Serpentinization: Connecting Geochemistry, Ancient Metabolism and Industrial Hydrogenation

Martina Preiner [1,*], Joana C. Xavier [1], Filipa L. Sousa [2], Verena Zimorski [1], Anna Neubeck [3], Susan Q. Lang [4], H. Chris Greenwell [5], Karl Kleinermanns [6], Harun Tüysüz [7], Tom M. McCollom [8], Nils G. Holm [9] and William F. Martin [1,*]

1 Institute of Molecular Evolution, University of Düsseldorf, 40225 Düsseldorf, Germany; xavier@hhu.de (J.C.X.); zimorski@hhu.de (V.Z.)
2 Division of Archaea Biology and Ecogenomics, Department of Ecogenomics and Systems Biology, University of Vienna, Althanstrasse 14 UZA I, 1090 Vienna, Austria; filipa.sousa@univie.ac.at
3 Department of Earth Sciences, Palaeobiology, Uppsala University, Geocentrum, Villavägen 16, SE-752 36 Uppsala, Sweden; anna.neubeck@geo.su.se
4 School of the Earth, Ocean, and Environment, University of South Carolina, 701 Sumter St. EWS 401, Columbia, SC 29208, USA; slang@geol.sc.edu
5 Department of Earth Sciences, Durham University, South Road, DH1 3LE Durham, UK; chris.greenwell@durham.ac.uk
6 Institute for Physical Chemistry, University of Düsseldorf, 40225 Düsseldorf, Germany; shangrila1@icloud.com
7 Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, 45470 Mülheim an der Ruhr, Germany; tueysuez@mpi-muelheim.mpg.de
8 Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder, CO 80309, USA; tom.mccollom@lasp.colorado.edu
9 Department of Geological Sciences, Stockholm University, SE-106 91 Stockholm, Sweden; Nils.Holm@geo.su.se
* Correspondence: preiner@hhu.de (M.P.); bill@hhu.de (W.F.M.); Tel.: +49-(0)211-81-12736 (M.P.)

Received: 27 August 2018; Accepted: 20 September 2018; Published: 22 September 2018

Abstract: Rock–water–carbon interactions germane to serpentinization in hydrothermal vents have occurred for over 4 billion years, ever since there was liquid water on Earth. Serpentinization converts iron(II) containing minerals and water to magnetite (Fe_3O_4) plus H_2. The hydrogen can generate native metals such as awaruite (Ni_3Fe), a common serpentinization product. Awaruite catalyzes the synthesis of methane from H_2 and CO_2 under hydrothermal conditions. Native iron and nickel catalyze the synthesis of formate, methanol, acetate, and pyruvate—intermediates of the acetyl-CoA pathway, the most ancient pathway of CO_2 fixation. Carbon monoxide dehydrogenase (CODH) is central to the pathway and employs Ni^0 in its catalytic mechanism. CODH has been conserved during 4 billion years of evolution as a relic of the natural CO_2-reducing catalyst at the onset of biochemistry. The carbide-containing active site of nitrogenase—the only enzyme on Earth that reduces N_2—is probably also a relic, a biological reconstruction of the naturally occurring inorganic catalyst that generated primordial organic nitrogen. Serpentinization generates Fe_3O_4 and H_2, the catalyst and reductant for industrial CO_2 hydrogenation and for N_2 reduction via the Haber–Bosch process. In both industrial processes, an Fe_3O_4 catalyst is matured via H_2-dependent reduction to generate Fe_5C_2 and Fe_2N respectively. Whether serpentinization entails similar catalyst maturation is not known. We suggest that at the onset of life, essential reactions leading to reduced carbon and reduced nitrogen occurred with catalysts that were synthesized during the serpentinization process, connecting the chemistry of life and Earth to industrial chemistry in unexpected ways.

Keywords: rock–water–carbon interactions; origin of life; carbides; iron sulfur; early metabolism

1. Abiotic Chemical Synthesis at Hydrothermal Vents

Since their discovery, hydrothermal vents have been of interest in thoughts about the origin of life [1,2]. They are relevant to origins for a number of reasons. From the standpoint of thermodynamics, hydrothermal systems harbor chemical reactions that are continuously far from equilibrium, a property they share with life [3–5], and they harbor gradients: Temperature gradients, pH gradients, and redox gradients [2,6]. Today, those gradients are most pronounced at the vent ocean interface, where vent effluent emerges into sea water and forms hydrothermal mounds [6,7]. But life is—to our knowledge—not arising anew today. Rather life emerged once. We can say that because all life forms we know (the only ones that demand an explanation) share the same genetic code, this can only be reasonably explained by common ancestry [8]. Carbon isotopes with signatures typical for life that appear in rock 3.95–3.8 Ga of age [9,10] provide the currently known date for the emergence of life.

Though views on the origin of life are traditionally marked by debate, everyone agrees that energy was important, because without energy uptake and release, no chemical reactions can take place, and if no chemical reactions take place, life forms can neither arise nor multiply. Chemical reactions require the flow of energy in order to be initiated and proceed. The main sources of energy at origins that are currently discussed in the literature are *UV* light [11,12], lightning that generates nitric oxides as oxidants [13], kinetic energy of meteorite impacts [14,15], hydrothermally generated ion gradients [16–18], and chemical energy in the form of the H_2–CO_2 redox couple [19–25].

This paper will focus on H_2-dependent CO_2 reduction in hydrothermal systems and its possible significance at origins. We will discuss H_2 synthesis in hydrothermal systems via serpentinization, H_2-dependent organic synthesis at modern vents, and H_2-dependent reduction of inorganic compounds in the crust to generate catalysts (native metals and carbides) that are used in the laboratory and in industry for H_2-dependent reduction of CO_2 and N_2. Catalysts are important because they increase reaction rates by lowering the activation energy. They also have a drastic effect on the type of products formed because in thermodynamically controlled reactions the most stable products accumulate, whereas in kinetically controlled reactions the most rapidly formed products accumulate. The nature of the most rapidly formed products is usually governed by the chemical and physical properties of the catalyst.

We will outline clear links between vents and well-studied existing life forms. Modern anaerobic autotrophs that live from the H_2–CO_2 couple, acetogens and methanogens, have a conserved native metal (Ni^0) in their most central CO_2 reducing enzyme—carbon monoxide dehydrogenase (CODH) [26,27]—and a carbide in the only enzyme that introduces N_2 into biology, nitrogenase [28,29]. At the same time, H_2-dependent reduction of inorganic compounds uncovers interesting links between vents and important industrial processes, because serpentinization yields not only H_2, but also Fe_3O_4, which is the starting catalyst for CO_2 hydrogenation to synthetic gasoline [30] and for N_2 reduction to NH_3 through the Haber–Bosch process [31,32].

That the chemical conditions, minerals and metallic compounds of hydrothermal vents could serve as sites of chemical synthesis of organic compounds on the early Earth is hardly a new idea [3,33,34]). Shock and Schulte [35] investigated the thermodynamics of hydrothermal systems, predicting nearly complete conversion of inorganic CO_2 to organic compounds under some conditions, emphasizing that organic compounds should even be more stable than mixtures of their precursors H_2 and CO_2 under a variety of conditions. Amend and Shock [36] showed that the synthesis of amino acids was thermodynamically favorable under hydrothermal vent conditions. Amend and McCollom [37] explored the thermodynamics of low temperature vents similar to Lost City (low temperature, slight pH gradients between effluent and sea water), and found that synthesis of organic compounds in the composition and stoichiometry of biomass from inorganic precursors is thermodynamically favorable (−1016 to −628 Joules per cell) at 50 to 100 °C. Clearly, there is substantial potential for organic synthesis at today's hydrothermal vents and likely more so in vents on the early Earth. If life

really started 3.8–3.95 billion years ago [9], we need to consider the state of the very early Earth, the setting within which the first hydrothermal systems formed.

2. The Early Earth: Magma, then Crust, then Oceans

The early Earth was molten following the moon-forming impact roughly 4.4 billion years ago, at the latest [38,39]. With temperatures on the molten Earth around > 1200 °C, carbon that had been brought to Earth by accretion was converted to CO_2, which was outgassed into the atmosphere [38,40], with a portion of CO_2 remaining in magma oceans [41]. Magma oceans also retained little water, which was predominantly converted to atmospheric steam. By about 4.4 Ga the magma ocean had cooled [42] and by about 4.2 Ga there was liquid water on Earth [38,39], some condensed from the atmosphere and some delivered later by comets. By around 4 Ga the late heavy bombardment had ended [42,43]. By 3.95 Ga a carbon isotopic signature compatible with that produced by the acetyl-CoA pathway had appeared [9].

As it relates to hydrothermal vents and organic synthesis, the relevant sequence of events starts from CO_2 as a result of magma oceans and minor amounts of mantle CO_2. In the molten state, the densest material of the early Earth—metals in the elemental state—was drawn to the planet's center where it remains to this day as the core (mainly 85% Fe and 5% Ni). This process of differentiation (gravitational metal migration to the core) did not completely separate heavy material from light, because some light elements also exist in the core, such as Si, S, and C, estimated at 6%, 2%, and 0.2% respectively [44]. Lighter material consisting mainly of silicates (iron, magnesium, and aluminum silicates), possibly with residual metals, was displaced to the surface accordingly, where it formed the primordial mantle and crust. The crust started as magma and, thus, had a very low water content [38]. As the crust cooled, water eventually condensed over the Earth's surface to form oceans of liquid water [38,40,45,46].

As liquid oceans formed, gravity pulled water into cracks of the steadily cooling crust. Water in the crust became heated and resurfaced, creating convective currents. The primordial ocean was about twice as deep as today's because the modern crust and mantle bind about one ocean volume of water [40,47], which was originally in the ocean before rock–water interactions in the primordial crust commenced. The CO_2 content of the primordial atmosphere was perhaps 100–1000 times higher [40] than today's, and very large amounts of CO_2 were, thus, dissolved in the ocean. Convective water currents through the very dry iron magnesium silicate crust led to rock–water interactions, initiating a process called serpentinization, as sketched in Figure 1.

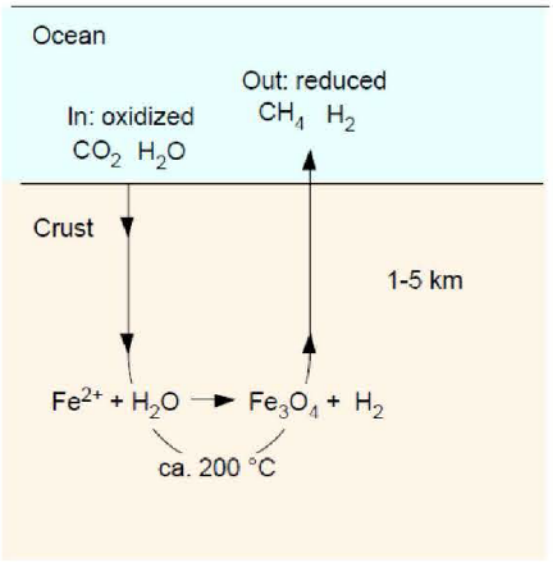


Figure 1. Schematic representation of serpentinization in a hydrothermal vent. See text and [5,48–52].

3. Serpentinization: Rock–Water Interactions

Serpentinization occurs when ultramafic rocks, enriched in the minerals olivine and orthopyroxene, react with water and are converted to rocks containing a suite of minerals dominated by serpentine. A significant by-product of the reaction is H_2. In general terms, the process can be summarized by the general reaction:

$$\text{Olivine} \pm \text{orthopyroxene} + H_2O \rightarrow \text{serpentine} \pm \text{brucite} + \text{magnetite} + H_2. \quad (1)$$

In somewhat more precise terms, the reaction can be expressed as:

$$\begin{gathered} Mg_{1.8}Fe_{0.2}SiO_4[\text{olivine}] + Mg_{1.8}Fe_{0.2}Si_2O_6[\text{orthopyroxene}] + wH_2O \\ 0.5(Mg, Fe^{2+}, Fe^{3+})_3(Si, Fe^{3+})_2O_5(OH)_4[\text{serpentine}] + x(Mg, Fe^{2+})(OH)_2[\text{brucite}] \\ + yFe_3O_4[\text{magnetite}] + zH_2 \end{gathered} \quad (2)$$

where the stoichiometric coefficients w, x, y, and z are variable and depend on a number of factors including temperature and the relative proportions of olivine and orthopyroxene in the reacting rock [48,49]. The production of H_2 during serpentinization results from the oxidation of ferrous iron (Fe^{2+}) from the reactant minerals to ferric iron (Fe^{3+}) in the products through reaction with water, liberating H_2. This process can be expressed as:

$$2(Fe^{2+}O) + H_2O \rightarrow (Fe^{3+}{}_2O_3) + H_2 \quad (3)$$

where ($Fe^{2+}O$) and ($Fe^{3+}{}_2O_3$) represent components of reacting rocks and product minerals, respectively. Magnetite is often the major product of Fe^{2+} oxidation during serpentinization. In some serpentinizing systems, sufficient amounts of H_2 accumulate to reduce Fe^{2+} and Ni^{2+} (a trace component of olivine), converting them to native metal alloys such as awaruite (Ni_3Fe) [53,54].

$$(Fe^{2+}O) + 3(Ni^{2+}O) + 4H_2 \rightarrow Ni_3Fe\ [\text{awaruite}] + 4H_2O \quad (4)$$

where ($Fe^{2+}O$) and ($Ni^{2+}O$) again represent components of the reacting rock. The essence of serpentinization is, ultimately, that Fe^{2+} is oxidized by water, water is consumed by the reaction, the oxidized Fe^{2+} and the newly formed oxygen remain as magnetite (Fe_3O_4), the water is reduced to H_2, and the vent fluid becomes alkaline (pH 9–11) [6] due to hydroxides being generated. Notably, magnetite, Fe_3O_4, and H_2 are the starting materials for industrial catalysts used for N_2 reduction via the Haber–Bosch process, to which we will return in a later section. Liquid gasoline synthesis from H_2 and CO_2, according to new findings [30] is also efficiently and specifically (up to 22% CO_2 conversion and 78% of products) catalyzed with magnetite. In industrial CO_2 reduction, iron catalysts are often the choice for high-temperature processes.

However, if we consider serpentinization as a source of energy and electrons for organic synthesis at the origin of life [5,50–52] or as a source of energy and electrons for the first ecosystems on Earth [38,55,56], we have to extrapolate back from modern systems. Importantly, only ultramafic (silica-poor) rocks can undergo serpentinization. The ultramafic rocks that are serpentinized in today's ocean lithosphere represent chunks of the mantle that are tectonically uplifted to shallow environments where serpentinization happens owing to lower temperatures [57,58]. Since the composition of the mantle does not appear to have changed significantly over time, primordial serpentinization should not have been fundamentally different from the modern process. Furthermore, a possibly thinner crust and the eruption of extrusive ultramafic volcanic rocks called komatiites because of higher mantle temperatures would have made this process more widespread on the early Earth than it currently is [35]. The primordial crust contained minor amounts of residual Fe^0 and Ni^0 from differentiation, and the ocean water circulating through the crust contained much more CO_2 than present day.

On the modern seafloor, fluids discharged from serpentinizing systems contain up to 16 mmol H_2/kg [59,60]. Laboratory simulations of serpentinization generate comparable and higher H_2 concentrations [61–63]. Hydrothermal fluids from serpentine-hosted seafloor systems are also commonly enriched in CH_4, up to several mmol/kg [59,60]. The exact temperature and pressure in natural serpentinizing systems are not known, serpentinization experiments in simulated hydrothermal systems show good results from 400 °C down to nearly ambient temperatures [64,65].

4. Serpentinization: Awaruite and Carbon

Some serpentinization loci are naturally richer in nickel and iron minerals than others, leading to alloys containing compounds like taenite (NiFe, 25–40% Ni), kamacite (NiFe, ca. 7% Ni), or awaruite [65,66]. The Ni–Fe alloys produced in serpentinizing systems are primarily awaruite, which can have a somewhat narrow range of composition from Ni_2Fe to Ni_3Fe ($Ni_{2-3}Fe$). Awaruite is a minor but fairly widespread component of serpentinized rocks [53,66,67]. Its mechanism of synthesis and/or deposition are not known in detail, but it is known to be generated by serpentinization because it occurs in hydrothermally altered serpentinite rocks [52,68]. It is thought to arise from the H_2-dependent reduction of Fe^{2+} and Ni^{2+} containing minerals during the serpentinization process [53,66]. In a laboratory simulation of serpentinization, McCollom [68] reported the laboratory synthesis of awaruite from Ni-containing olivine under conditions that generated ca. 60 mmol/kg H_2. In the temperature range 200–400 °C, high H_2 activities and low H_2S activities favor awaruite formation [53]. Foustoukos et al. [54] report that awaruite can be formed at H_2S activities of ca. 1 mmol/kg and at H_2 activities of <100 mmol/kg, in line with findings from laboratory awaruite synthesis [68].

Serpentinization in submarine hydrothermal systems has been around since there was water on Earth [55]. Contingent upon a heat source and cooling rates at the ridge axis, the reaction runs continuously until the host rocks are completely serpentinized, or the supply of water is exhausted. Lost City is at least 30,000 years old [69] and possibly 100,000 years old [70], indicating that these systems can be fairly long-lived. Lost City effluent contains 1–2 mmol/kg methane [71], with formate being the second most prevalent carbon species [72]. Like methane, the isotopic composition of Lost City formate indicates an abiotic origin [56,72]. Smaller amounts of ethane, propane, and butane are also observed in a serpentinizing environment and attributed to abiological synthesis [71,73–75].

During the lifespan of a serpentinizing vent, the rock reacts and its composition changes, altering the redox state and chemical composition of its surfaces, an aspect that might be relevant in terms of catalyst formation during serpentinization.

A study by Horita and Berndt [76] illustrates the principle. In laboratory experiments to better understand the presence of methane in hydrothermal effluent, they examined the reduction of CO_2 with H_2 to CH_4 using awaruite as the catalyst. At temperatures from 200–400 °C, pressure at 50 bar (5 MPa), and H_2 activities around 200 mmol/kg, they obtained 1–10 mmol/kg CH_4, often with >50% conversion rates of CO_2 added. The point is that awaruite (the catalyst) is not a pre-existing component of the rocks that host hydrothermal systems. Awaruite is first synthesized on site by the serpentinization process, and once formed it can then catalyze the H_2-dependent reduction of CO_2.

The study of Horita and Berndt [76] points to the significance of native metals in rock–water–carbon interactions more generally. Heinen and Lauwers [77] investigated the ability of Fe^0 to catalyze CO_2 reduction in the presence of H_2S to obtain thiols in an early evolution context. They obtained modest thiol yields (nmol), but they only analyzed volatile S containing compounds.

In the laboratory, good rates of CO_2 conversion have been reported under hydrothermal conditions using native metals as catalysts and reductants. Guan et al. [78] showed that Fe^0 in the presence of potassium, copper and aluminum will reduce CO_2 to CH_4, C_3H_8, CH_3OH, and C_2H_5OH in the 10–70 µM range in 20 h at room temperature. He et al. [79] reported a reduction of CO_2 to formate and acetate in the 1–10 mM range using nanoparticular Fe^0 at 80–200 °C for 5 to 200 h. Varma et al. [80] used Fe, Ni, Mo, Co, and W at temperatures between 30–100 °C, all with some success, but the

best yields (10–200 µM concentrations of reduced carbon compounds) were observed with Fe^0 in presence of potassium salts. Importantly, the reduced carbon products observed by Guan et al. [78], He et al. [79], and Varma et al. [78] are compounds that also occur as intermediates and end products in the metabolism of organisms, such as acetogenic bacteria and methanogenic archaea [81,82], that live on H_2 and CO_2 as the substrates for their carbon and energy metabolism: Formate, methyl moieties, and acetate. Varma et al. [80] even reported the synthesis of pyruvate at temperatures of 100 °C and below.

In terms of microbial physiology, Varma et al.'s [80] findings are very significant because the reduction of CO_2 to pyruvate—presumably via formate, methyl groups, and acetyl groups—exactly mirrors the reaction sequence in the acetyl-CoA pathway, the pathway of carbon and energy metabolism in organisms that live from the reduction of CO_2 with H_2 [83,84]. Stated another way, when CO_2 and native metals react in water overnight under strictly anaerobic conditions, the core reaction sequence of the most ancient pathway of microbial carbon and energy metabolism, the acetyl-CoA pathway [24,83,85,86], which operates in the most ancient microbial lineages [87,88] unfolds in a series of spontaneous non-enzymatic reactions in the presence of water. The pressures employed by Varma et al. [80] do not preclude, however, the existence of a gas phase. At depths of several km, very high gas activities can be attained, but the pressure (hundreds of bars) would seem to be too high for a gas phase. However, Früh-Green et al. [89] reported gas bubbles arising during drilling around vents at the Atlantis Massif at a depth of 1,140 m.

Noteworthy in the study of Varma et al. [80] is that the products of CO_2 reduction appear to be synthesized and bound on the surface of the metals, such that they had to be cleaved by alkaline hydrolysis to be assayed. It is not clear whether the products were bound to the Fe particles via C–Fe bonds or C–O–Fe bonds, because there was no surface analysis so far, but alkaline hydrolysis was required to obtain the soluble products. In the studies of He et al. [79] and Guan et al. [78] some fraction of the reduced carbon products was probably discarded, bound to metal surfaces. Reaction mechanisms for CO_2 reduction have been proposed by He et al. [79] and Varma et al. [80] but the exact role of H_2 and Fe^0 in CO_2 reduction is not yet clear. Heinen and Lauwers [77] showed that under anaerobic conditions, Fe^0 and H_2O readily generate H_2 which, in the presence of the metal, readily reduces CO_2. Guan et al. [78] reported that roughly 0.3 mol H_2 was generated per mol Fe^0.

Awaruite was also reported to catalyze the reduction of N_2 with H_2, but at very low rates just above background [90]. Under simulated deep crust conditions comparable to the Haber–Bosch process (>300 °C), but without exogenous H_2, the reduction of N_2 to NH_3 on native iron was reported [91], as were low conversion rates (~0.1%) of N_2 to NH_3 with H_2S in the presence of FeS at ambient pressure and 90 °C [92]. On the scales of submarine crust volume and geological time, a constant supply of small amounts of reduced nitrogen, or activated nitrogen species on metal catalyst surfaces, could easily be sufficient to underpin prebiotic synthesis of nitrogenous carbon compounds [93].

5. Serpentinization: Methane

Distinctive C and H isotope signatures and other evidence indicate that the methane discharged from many serpentinizing systems has an abiotic origin [71,74,75]. Presumably, this methane is produced as the H_2 produced by serpentinization reacts with CO_2:

$$CO_2 + 4\,H_2 \rightarrow CH_4 + 2H_2O \quad (5)$$

Although exactly where in the system and how it is produced remains a matter of ongoing debate [74,75,94,95]. In the absence of catalysis, reduction of dissolved CO_2 is extremely slow even at temperatures up to 350 °C [33,68,96]. However, when suitable catalysts such as awaruite are available, or when CO_2 and H_2 are present in a gas phase, the reaction can proceed much more readily [22,33,75,76]. It is this kinetic inhibition that allows CO_2 and H_2 to remain in disequilibrium and be exploited as an energy source by methanogenic organisms in hydrothermal environments.

Reduction of dissolved CO_2 to formate proceeds readily under simulated hydrothermal conditions and, as mentioned above, formate is observed in modern hydrothermal vent effluent, although typically at lower concentrations than methane [72,74,97].

Today, the formation of abiotic methane is important as a substrate for methanotrophic organisms in an oxidizing environment. In an origin of life context, methane itself is probably not of central interest, because of its very strong C–H bond. At great depths in the oceanic crust and in the upper mantle dominated by the fayalite-magnetite-quartz (FMQ) redox buffer, CO_2 is the dominant carbon species at stable equilibrium [20]. At temperatures above about 350 °C, equilibrium between CO_2 and CH_4 favors CO_2 even at elevated H_2 concentrations [5,20]. However, with decreasing temperature, the equilibrium shifts to favor CH_4, so that at temperatures below ~350 °C reaction of CO_2 with H_2 to form CH_4 is thermodynamically favored. If the activation energy is too high and kinetic hindrances arise that prevent equilibrium conditions from being established, the organic chemistry will be locked up in different organic metastable compounds such as carboxylic acids. If nitrogen is present, even amino acids and nitrogen bases may be formed. Therefore, it is the interrupted transition from CO_2 to CH_4 that has the potential to create prebiotic constituents of life processes in the ocean floor and not CH_4 itself. Organic chemists have often rejected igneous environments as a likely site for the origin of life because CO_2 is a common component of fluids and gases (because of kinetic hindrances). Oparin [98], for instance, claims that 'carbon dioxide is not the beginning but the end of life'—a statement that is, of course, only true in an oxidizing environment.

6. Serpentinization: Magnetite (Fe_3O_4)

Magnetite is a common (although not ubiquitous) product of serpentinization. It is of interest in the context of CO_2 reduction because it is the starting point to reach industrial catalysts for CO_2 hydrogenation and for N_2 reduction to ammonia via the Haber–Bosch process. Wei et al. [30] reported the H_2-dependent reduction of CO_2 to hydrocarbons (C_2–C_{11}), methane, and aromatics, with up to 22% CO_2 conversion using Fe_3O_4 as the catalyst. During the reaction process, at about 320 °C and 30 bar (3 MPa), with H_2/CO_2 ratios of 1:1 to 6:1, Fe_3O_4 is converted by H_2 in situ to iron carbide, Fe_5C_2, as analysis of the spent catalyst reveals. Fe_5C_2 is, in turn, thought to be the decisive catalyst for Fischer–Tropsch (FT) synthesis of longer hydrocarbons from CO, which is generated on Fe_3O_4 sites via a reverse water gas shift (WGS) reaction [30,99,100]. Note the involvement of WGS means that H_2O is present during the reaction. FT and WGS reactions have long been discussed in the context of hydrothermal organic synthesis [101,102]. That magnetite itself is not an effective catalyst, as experimental studies reveal [68,103], is not the main point here. The point, as we see it, is that the effective catalyst, an iron carbide, is synthesized from CO_2 and Fe_3O_4 in the presence of H_2 during the reaction that mirrors conditions and chemical components found in serpentinizing hydrothermal vents, as outlined in Figure 2.

The participation of carbides in CO_2 hydrogenation is well known from processes developed for industrial application. Metal–carbide interfaces catalyze CO_2 conversion into CO [104] or methanol [105]. The literature on Fischer–Tropsch synthesis [106,107] shows that carbides are a product of reactions with iron catalysts. This happens especially with CO, because it decomposes to CO_2 and chemisorbed carbon. The latter reacts further to produce iron carbides. As one can commonly observe magnetite, Fe_3O_4, in both fresh and spent iron catalysts, this compound is thought to be an active part of such reactions. Small amounts of native iron also seem to help the catalysis from CO to carbide [108]. Carbides have not been widely considered in an early evolution context, although they are also found in natural systems: Iron carbides are formed in the lower parts of the Earth's crust (called ophiolite), but are uplifted over time—some of them even containing nickel [109–111]. In all of biology, there is only one carbide carbon known. Discovered in 2011 [28,29], it resides at one of the most crucial reactions for fueling ecosystems and in one of the most ancient enzymes known, nitrogenase.

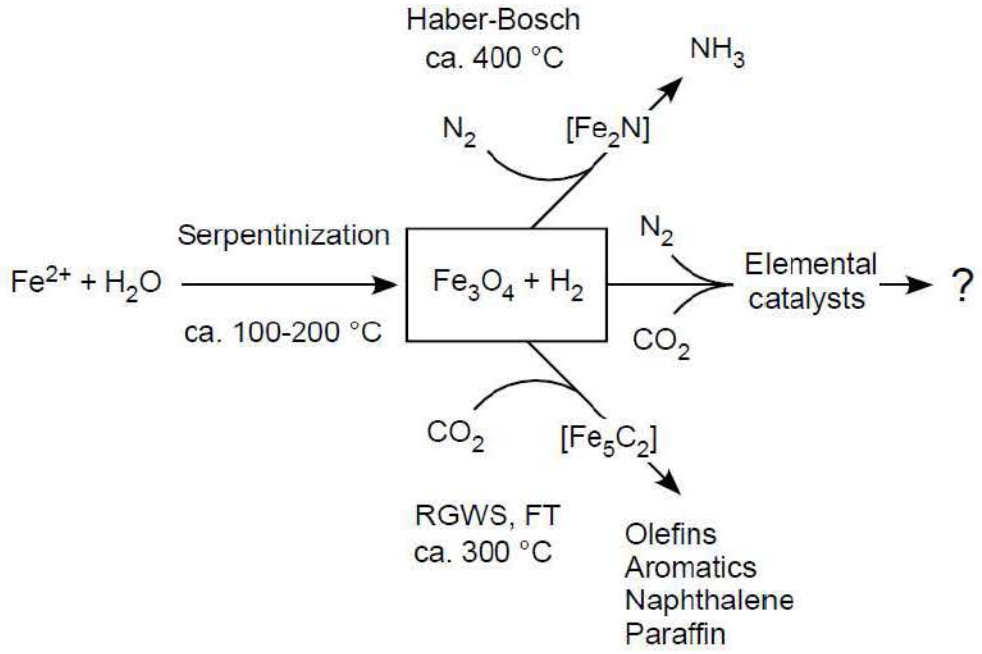


Figure 2. Possible connections between industrial processes and chemical evolution. See text. RWGS, reverse gas water shift reaction. FT, Fischer–Tropsch. The Haber–Bosch process starts with magnetite and generates nitrides [31,32]. The synthesis of gasoline from CO_2 starts with magnetite and generates iron carbide, both catalysts appear to fulfill important but distinct roles [28]. Reaction parameters aimed at simultaneous reduction of N_2 and CO_2 are not well explored.

Magnetite itself, without carbides and without exogenous H_2, can reduce CO_2 in H_2O to acetate with roughly 3% conversion at 250 °C [112]. The greater high-temperature catalytic ability of reduced magnetite-containing carbides [106] or nitrides [32] and the lower-temperature (100–200 °C) catalytic ability of native metals [80] and awaruite [76] raise a question that needs to be spelled out clearly: Do serpentinizing hydrothermal systems improve their organic synthetic capacity during their lifespan by synthesizing better (more reduced) inorganic catalysts of organic reactions? That is, does serpentinization generate carbides and nitrides from magnetite and H_2 in a similar manner to what chemical industry does? And if so, what kinds of products might one expect under conditions where both CO_2 and N_2 could be reduced simultaneously?

In a typical heterogeneously driven gas phase catalytic reaction, the first step of a catalytic reaction is adsorption of reactants on the surface of catalysts, followed by dissociation of reactants, product formation, and, finally, desorption of products from the surface of the catalyst. Heterogeneous catalysts naturally contain different types of catalytically active surface sites depending on their composition, crystallinity, particle size, and facet, and also impurities. If the catalyst contains several components (as the starting materials for the serpentinization process do) the surface composition might be different from a bulk structure. Moreover, the surface structure and composition of the catalyst can change dramatically during the catalytic reaction as it is the case for CO_2 reduction over magnetite [30] and N_2 hydrogenation to NH_3 in the Haber–Bosch process [32].

This is potentially an interesting avenue of thought. The sequence of events underlying awaruite synthesis can, generally, be understood as large amounts of reduced iron and water giving rise to lesser amounts of H_2, which then reduces divalent Ni and Fe in minerals to generate reduced magnetite and elemental metals Ni^0 and Fe^0 (Ni_3Fe) which appear to have catalytic properties when it comes to H_2-dependent CO_2 reduction [76,80]. If so, the geochemical process is synthesizing catalysts for organic synthesis—constantly. As outlined above, processes of catalyst synthesis from Fe_3O_4 and H_2 are demonstrably occurring in industrial applications for reduction of CO_2 and N_2.

7. Conserved Relicts in Metabolism: The Ni^0 in CODH

The reaction mechanism of an ancient enzyme called carbon monoxide dehydrogenase (CODH) entails the generation of zero valent Ni (Ni^0) as an intermediate in biological CO synthesis as shown in Figure 3a. CODH is involved in a pathway that is thought to be just as ancient (probably even more so) than the enzyme itself: The acetyl-CoA pathway. It is the only exergonic pathway of autotrophic carbon metabolism known [83,85,113]. Its exergonic nature allows acetogens and methanogens to generate transmembrane ion gradients in a process involving flavin-based electron bifurcation [114] during the process of CO_2 fixation [81,115] and thereby conserve energy in the form of ATP via electron transfer phosphorylation (chemiosmosis).

Flavin-based electron bifurcation is a newly discovered mechanism of soluble (as opposed to membrane-associated) energy conservation [114,116,117]. Its principle is significant and we will encounter it again in a later section so it should be briefly explained here. The midpoint potential, E_0', of H_2 (–414 mV) is not sufficiently negative to generate the low-potential reduced ferredoxin (E_0' = ca. –500 mV) that acetogens and methanogens require and use to reduce CO_2 under physiological conditions. How, then, do acetogens and methanogens send electrons from H_2 energetically uphill by roughly –100 mV? The electron pair from H_2 is transferred to a flavoprotein, the flavin of which splits the pair: One electron goes energetically uphill to ferredoxin while the other goes energetically downhill to a more positive electron acceptor such as NAD^+ (E_0' = ca. –320 mV) [115] or a heterodisulfide (E_0' = ca. –140 mV) [118] or similar [114]. The reduction of the downhill (more positive) acceptor energetically finances the reduction of the uphill acceptor so that ferredoxin is reduced and CO_2 in the acetyl-CoA pathway can be fixed.

Although five of the six known pathways of autotrophic carbon metabolism also generate acetyl-CoA as the net end product of CO_2 fixation starting with electrons from H_2 [83] (the Calvin cycle generates glyceraldehyde 3-phosphate), only the acetyl-CoA pathway generates ATP from CO_2 fixation. The other five require ATP input to fix CO_2 [84]. That ATP input comes from an independent energy metabolism, typically aerobic or anaerobic respiration (sulfate reduction for example), that is independent of CO_2 reduction. The reverse TCA cycle only requires the input of one ATP per CO_2 [119,120], but it still requires ATP input. The acetyl-CoA pathway permits ATP synthesis from H_2-dependent CO_2 reduction. Why? It involves CO as an intermediate, a carbonyl moiety that becomes a carboxylate only after going through a sequence of metal carbonyl (CO–Ni), thioester (CO–S) and acyl phosphate ($CO–OPO_3^{2-}$) bonds, the latter of which phosphorylates ADP to generate the carboxylate. In all other pathways of CO_2 fixation, CO_2 is fixed as a carboxylate that subsequently has to be reduced, ultimately at the expense of ATP hydrolysis [83,84]. It is the zero valent Ni of CODH that generates CO for the acetyl-CoA pathway; that is the chemistry that makes the pathway exergonic [27,121,122].

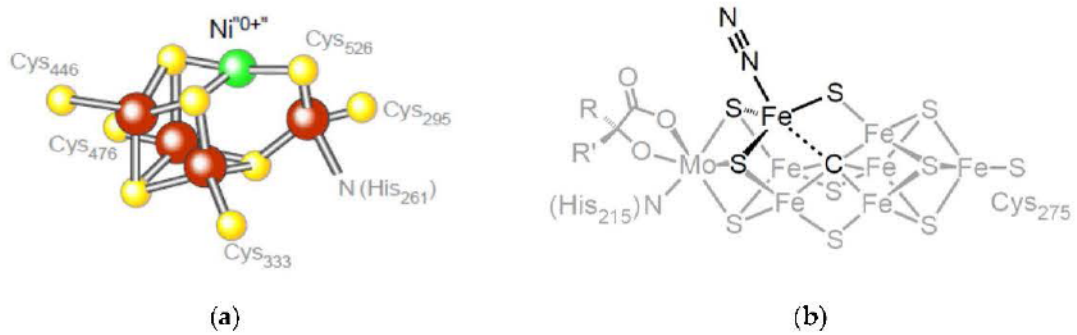


Figure 3. Relicts in metabolism. (**a**) The active site of CODH that interconverts CO_2 and CO, redrawn from supplemental figure S5 in Ragsdale [26] underscoring the reduced Ni atom that binds CO_2 in the proposed mechanism for the CO-generating reaction. (**b**) The carbide carbon in the active site of nitrogenase [28,29] and its proposed role in the catalytic mechanism [123].

8. Conserved Relicts in Metabolism: The Carbide in Nitrogenase

In terms of dry weight, life is 50% carbon and 10% nitrogen. Though CO_2 can enter metabolism via six known carbon fixation pathways [83], there is only one known entry point for N_2: Nitrogenase. It has long been known for its complex 7Fe-9S cluster harboring a light atom at its center and for existing in three related forms that differ with respect to the metal cofactor that is peripherally associated with the active site: Mo, V, or Fe [123–125]. It requires 16 ATP and 8 electrons per conversion of N_2 into 2 NH_4^+ and it produces H_2 as an unavoidable reaction by-product. The ATP is consumed by "archerases" in nitrogenase which hydrolyze ATP to induce conformational changes that alter the midpoint potential of 4Fe4S clusters to more negative values for N_2 reduction [126].

In 2011, the light atom at the nitrogenase active site was identified as a carbon atom in the elemental state that is coordinated by four iron atoms, forming an essential Fe_4C carbide at the catalytic heart of the enzyme [28,29] (Figure 3b). The carbide is generated during nitrogenase maturation from a methyl group donated to an active site Fe by S-adenosyl methionine (SAM) [127]. Synthetic analogs of the active site of nitrogenase even catalyze the synthesis of hydrocarbons from CO [128].

Nitrogenase is a very ancient enzyme that traces to the last universal common ancestor (LUCA) in ancestral genome reconstructions [87]. It has never been replaced during 4 billion years of evolution, nor has an alternative enzymatic mechanism been invented among microbes that would serve the same purpose of making N from N_2 available for biosynthesis (amino acids, cofactors, nucleotides) and growth [129]. The incorporation of nitrogen into metabolism requires an iron carbide, as models for the nitrogenase catalytic mechanism suggest [130]. There is apparently no other way for biological systems to reduce N_2 with the tools of enzymes and cofactors. That suggests to us that the carbide-containing active site of nitrogenase is a biological reconstruction of the naturally occurring inorganic catalyst that gave rise to organic N at the onset of biochemistry.

This inference parallels the situation with a native metal (Ni^0) and CO synthesis in the exergonic acetyl-CoA pathway: There is apparently no mechanistic alternative to exergonic CO_2 reduction with H_2 that can be readily realized during 4 billion years of evolution. We suggest that elemental carbon (carbide) in nitrogenase and Ni^0 in CODH represent relicts from the chemical environment that supported organic synthesis at life's origin.

9. Weighing in on Caveats

One of the strengths of hydrothermal vent theories is that over the years since their first formulations, the gaps between spontaneous chemistry at vents and the chemistry of life have narrowed, not widened. Methane made abiogenically in substantial amounts at vents [22,52,71,72] stems from the same overall exergonic chemical reaction at the heart of methanogenesis in archaea, namely Equation (5). A similar case can be made for the acetogenic reaction [81,131].

$$4H_2 + 2CO_2 \rightarrow CH_3COOH + 2H_2O \quad (6)$$

Methanogens and acetogens are identified as the most ancient archaea and bacteria respectively, in reciprocally rooted trees for genes that trace to the last universal common ancestor (LUCA) [87,88]. Furthermore, H_2-dependent methanogens appear as the most ancient archaea in some phylogenies [132] and clostridia (acetogens) appear as the most ancient bacteria, too [133], although lineage phylogenies among prokaryotes are always in flux and inferences from them are generally problematic because lateral gene transfer decouples physiology from phylogeny [134]. More direct observations are that methanogens abundantly inhabit vents [23,56] and the deep, subsurface oceanic crust [135,136] today.

But what about acetate or chemically reactive methyl groups like methanol or methyl sulfide? For example, the absence of more than nanomolar amounts of methyl sulfide was taken as evidence against the hydrothermal origins theory [137], but methanethiol is avidly assimilated by methanogens and acetogens in the acetyl-CoA pathway [83]. Similar reasoning applies to acetate, which is

very scarce in hydrothermal effluents sampled so far [72], even though it readily accumulates in experiments intended to simulate vent conditions [80]. Acetate is the carbon and energy source for aceticlastic methanogens [81] belonging to the Methanosarcinales, which are abundant at hydrothermal vents [138,139], because the substrates they require for growth abound. The deep biosphere contains very large amounts of biomass [140,141] consisting of microorganisms that are generally starved for reducible carbon substrate, Heberling et al. [142] estimate 200×10^9 tons of biomass in the marine igneous crust (that is, excluding sediment). That abiotic acetate or methylsulfide is scarce in vent effluents reported so far [56,137] might not directly reflect synthesis processes at depth, but microbial scavenging instead. Methane accumulates within vents because strong oxidants, which are typically lacking within vents, are required for methane oxidation [143].

Another criticism concerns synthesis. Miller and Bada [144] argued that FT type reactions readily catalyzed by native metals cannot generate organic compounds in hydrothermal systems because of catalyst inhibition by H_2O and H_2S. Holm et al. [52] explain, however, that off-axis hydrothermal systems of the Lost City type have low concentrations of H_2S [145]. Additionally, awaruite can be formed at higher H_2S activities (ca. 1 mmol/kg) than previously thought [54], and metals other than Fe, including Ni, are not inhibited by H_2S.

Water activities are another issue typically raised by critics of hydrothermal vent theories [146]. Water hydrolyzes RNA and interferes with many reactions that generate specific prebiotic-type syntheses of particular organic molecules. This has spawned arguments that life must have arisen on land in dry or even desert-like conditions [147] or that essential reaction sequences giving rise to RNA monomers took place in the absence of water because that is how they work best in the laboratory with energy from *UV* light [12]. Nobody has yet proposed that RNA replication took place without water. Life is 80% water by weight, about 60% protein and 25% RNA by dry weight [148], with protein synthesis consuming about 75% of a cell's energy budget [149]; life counters hydrolysis problems by synthesizing polymers faster than they are hydrolyzed, at the expense of chemical energy.

An underappreciated aspect concerning water activity is that serpentinization takes place in rock that is initially very dry and that roughly 12 to 37 mol water is consumed per mol H_2 produced (reactions **1**–**4**). Placed in the context of methane or acetate synthesis (reactions **5**–**6**), roughly 100 mol of H_2O is consumed per CO_2 or carbonate that is reduced to the level of a methyl group. Furthermore, a great deal of additional water is simply converted into hydroxylated minerals during serpentinization. At the same time that water is being consumed, reactive gasses (H_2, CO) are being generated. Gas reactions on solid phase catalysts in serpentinization studies should not be neglected, nor should the circumstances that hydrophobic organics synthesized at hydrothermal vents will undergo phase separation and generate molecular environments of low water activity, which is the main function that enzymes provide in catalysis: Exclusion of water from the active site [150]. The crust hosting hydrothermal systems is necessarily replete with localities of low water activity and water disappears where reducing power for carbon reduction is generated.

10. Do Genomes Help?

Recent work has examined the origins problem from a novel top-down perspective. Standard approaches analyzed genomes to see which genes are universal to all cells, recovering about 30 genes involved in information processing [88]. Weiss et al. [87] constructed all trees for all proteins from 2000 genomes and identified proteins that are not universal to all cells but that are ancient by phylogenetic criteria. They looked for genes that trace to the LUCA because they are present in bacteria and archaea, but present by virtue of vertical inheritance from LUCA as opposed to late lateral gene transfer [87,88]. Based upon the kinds of genes that trace to LUCA, the results indicated that LUCA lived from dissolved gasses: H_2, CO_2, CO, N_2, and H_2S. The results indicated that for carbon assimilation, LUCA used the simplest and most ancient of the six known pathways of CO_2 fixation, the acetyl-CoA pathway germane to methanogenesis and acetogenesis, but it lacked the methyl synthesis branch suggesting that it relied on a geochemical supply of methyl groups. LUCA had nitrogenase, it had no traces of light utilization,

its environment was hot, contained H_2, CO_2, CO, N_2, H_2S, and metals—an environment that looks very much like a hydrothermal vent. Genomes trace LUCA to a site of rock–water–carbon interactions.

LUCA was able to harness ion gradients via the rotor-stator ATPase, but proteins of the ion gradient generation (pumping) were missing [87], which would have been possible, if LUCA lived at the vent of a serpentinizing system. There, geochemically generated ion gradients (alkaline inside versus the more neutral ocean) could be harnessed before the machinery was invented that allows cells to pump with the help of an ion gradient generated by a chemistry that is specified by genes [24,131]. In short, LUCA was half-alive, dependent upon geochemical organic synthesis and the chemical disequilibrium of its environment to harness carbon and energy. That might sound radical, but no theory for origins can operate without some kind of sustained chemical synthesis from the environment. A strong argument for origins at vents is the congruence between the basic CO_2 reduction reactions of serpentinization and the chemistry of acetogen and methanogen metabolism [24,86]. Metabolic pathway reconstructions of phosphate independent reactions uncovered the essential role of ferredoxin-dependent redox reaction and thioesters in ancient anabolic biochemistry [151,152]. LUCA's metabolism and genetic code were heavily dependent upon methyl groups [87].

The kind of LUCA that emerges from genomic reconstruction would have starved (low H_2 and CO_2 activities) and perished (*UV* light) at the surface—what it needed to survive was provided by rock–water interactions in the crust: Reactive gasses. Perhaps the crust was the first environment on Earth to have been inhabited [4,55]. Biologists and chemists have long held that inorganic catalysts/electron donors like FeS centers preceded organic catalysts/electron donors like NADH in evolution [153,154]. Similarly conserved base modifications (methylations, sulfur additions) in tRNA, in particular in the anticodon loop that allows the modern genetic code to operate, reflect the chemical environment within which the genetic code arose [87,88].

LUCA also harbored hydrogenases [87]. In Figure 3, the active sites of CODH and nitrogenase underscore the role of Ni^0 and a carbide in the entry of carbon (from CO_2) and nitrogen (from N_2) into the biosphere. These active sites represent bottlenecks in the origins and early evolution of metabolism. For the reduction of carbon and nitrogen, electrons stemming from H_2 as the proximal carrier were required. But in metabolism, H_2 never interacts directly with C or N, rather it enters metabolism via metal atoms in hydrogenases. Electrons from H_2 always enter metabolism via metals in the active site of hydrogenase or hydrogenase domains. There are three different kinds of hydrogenases known, they are phylogenetically unrelated and have very different active sites: The [NiFe] hydrogenase, the [FeFe] hydrogenase, and the [Fe] hydrogenase, the latter lacking FeS clusters [155]. Common to all three active sites, however, is that H_2 relinquishes its electrons via interaction with an Fe atom that is coordinated by one or more CO molecules [156]: Iron(II) carbonyls participate in H_2 activation in three independently arisen hydrogenase active sites. In contrast to CODH and nitrogenase, hydrogenases have appeared three times independently in evolution, but the iron carbonyl catalyst is conserved, indicating that the catalyst is older than the proteins that harbor it.

11. What Next?

What has been missing from hydrothermal vent-based research on origins is a decisive experiment where a wide variety of amino acids and bases arise under genuinely realistic conditions, although an early attempt was made by Hennet et al. [34]. FeS generally does not efficiently convert CO_2 into reduced carbon, either in the laboratory or in metabolism. Native metals do [80], and they furthermore generate the biologically relevant carbon species that occur at the core of metabolism: Formate, acetate, pyruvate, and methyl groups. The reason that FeS does not work well, but Fe^0 does, stated most simply, is that CO_2 reduction in biology (and perhaps more generally) is always a two-electron reaction, whereas FeS only undergoes one-electron reactions. Wächtershäuser has stressed the role of FeS in primordial biochemistry [157] and reported extensive experiments using FeS as a catalyst in reactions involving CO as a reductant, in which amino acids were observed [158]. Bases (heterocycles) have not been observed so far in FeS catalyzed reactions, nor has CO_2 reduction

using FeS catalysts ever been convincingly demonstrated in terms of reaction rates or product yields so far. CO_2 reduction using H_2 and Ni_3Fe catalysis at higher temperatures (200–400 °C) produce good yields of methane [76], but methane is rather an end product of metabolism than a building block of life.

That is a complicated way of saying that work investigating the role of hydrothermal chemistry in an origins context should possibly be looking at higher temperatures, higher pressures, low water activities, and catalysts containing native metals, carbides, and minerals like magnetite. It is possible that efficient synthesis of amino acids and bases from CO_2 and N_2 requires conditions where both gases are reduced to an appreciable extent simultaneously. This does not mean that Haber–Bosch-like conditions (500 °C) need to be employed for prebiotic type synthesis, because there is no imperative to generate high yields or specificity from such experiments. It does, however, suggest that FeS, despite its unassailably essential role in the physiology of life, despite its unquestionable antiquity [153], and despite its unquestioned ability to generate thioesters from CH_3SH and CO [159] might not be the right electron donor to get CO_2 reduced in a hydrothermal context. Although FeS is crucial to make iron biologically accessible, and sulfur compounds were probably very important for the transition from thermal to chemical activation, catalysts that promote two-electron reactions fit better to the chemistry of carbon. We also need to consider the very real possibility that some reactions work better (or at all) at greater depth, products being transported by hydrothermal current to cooler environments near the surface where they can react under milder conditions (Figure 4), and where lower temperature reactions more similar to energetically coupled metabolism (as opposed to synthesis), such as those reported by Muchowska et al. [160] (although under acidic conditions) and Varma et al. [80], or oscillating thioester reactions such as those reported by Semenov et al. [161], come into play. Even at hydrothermal vents, in terms of catalysis, some reactions take place on certain surfaces at higher temperatures, others will only function at lower temperatures with different catalysts. In that sense, the sequence of reactions in the acetyl-CoA pathway [83] from electron bifurcation, to CO, to acyl metal bonds, to thioesters that yield acyl phosphates (Figure 5), which phosphorylate ADP to generate ATP, can be seen as the spatially condensed and enzymatically catalyzed version of a spontaneous geochemical process, the thermodynamic drive for which stems from the natural tendency of CO_2 to be reduced with electrons from H_2. That is, similar to the iron(II) carbonyl of hydrogenases, the chemical reactions of the acetyl-CoA pathway themselves are older than the enzymes that catalyze them.

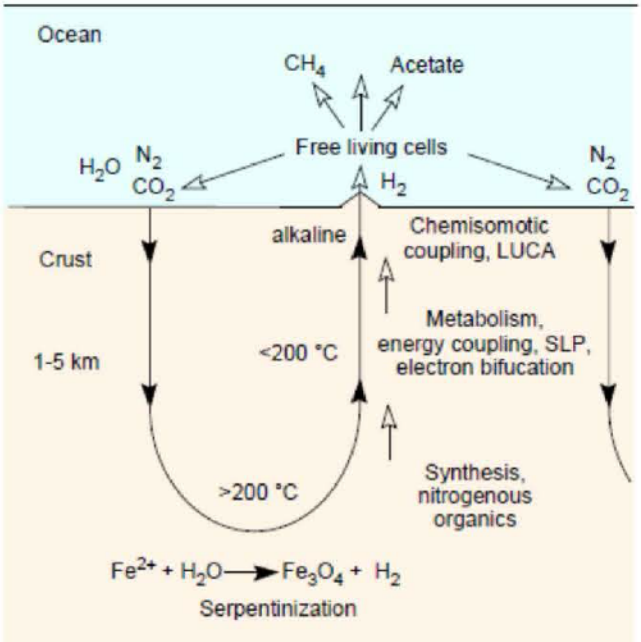


Figure 4. Possible processes at depth and at the ocean floor in serpentinizing systems. See text.

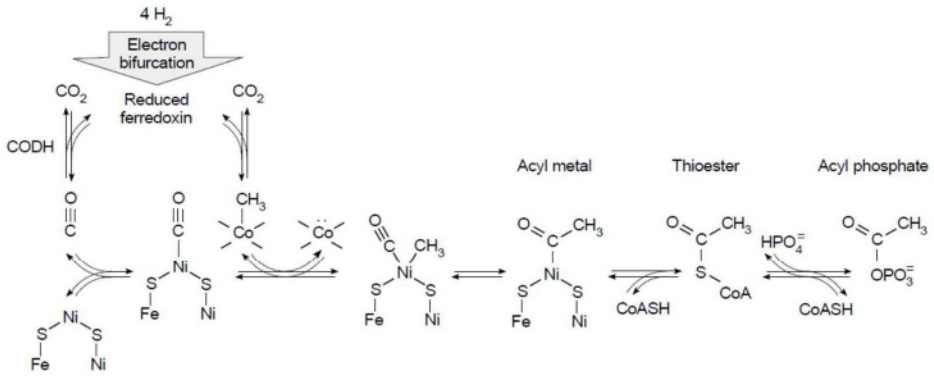


Figure 5. An ancient pathway. The diagram summarizes the biological energy conservation from ferredoxin to acyl phosphate in the acetyl-CoA pathway in an early evolution context [23,78,112]. Note that the reactions shown also occur without enzymes under suitable conditions [152]. See text. For an explanation of electron bifurcation see [106]. CODH: Carbon monoxide dehydrogenase.

Early papers on vents and life [1] voiced similar considerations about the possibility of different processes occurring at different depths. This became more evident as it became clear that serpentinization was relevant to synthesis at vent and origins [50]. A spectrum of processes across depths and conditions needs to be considered. And why, one might ask, have industrial chemists not already exhaustively explored the kinds of reactions sketched at the right of Figure 2? Reacting H_2, N_2 and CO_2 on Fe_3O_4 at temperatures needed for nitrogen reduction is not only Haber–Bosch expensive (energy intense), it is not going to generate any specific product. Chemists have not yet explored all kinds of conditions that might be interesting for prebiotic chemistry [162], the high-pressure–high-temperature reaction parameter space holds potential for investigation. Hydrothermal vents have been supporting microbial life for 3.5 billion years [163], perhaps because it arose there.

There also might be other parallels between bio- and geochemistry. For example, the formation of Ni_3Fe is still a mystery, and it poses an energetic problem similar to flavin-based electron bifurcation: The electrons from H_2 need to go energetically (far) uphill to reduce Fe^{2+} to Fe^0 under the conditions in hydrothermal systems. Might there be a geochemical redox process akin to biological electron bifurcation, that is, might the reduction of Fe^{2+} to Fe^0 via hydrogen be energetically coupled to (and financed by) Ni^0 deposition? It would help to explain why native Ni and Fe are deposited together, and in the case of awaruite in a stoichiometry favoring Ni, even though it is much rarer than Fe in serpentinizing systems.

Some of the organic compounds found on carbonaceous chondrites are apparently formed via FT type reactions [164]. FT type reactions in the presence of NH_3 generate the familiar nucleobases, but at very low conversions (ca. 0.02% of product) and requiring short heating up to ca. 600 °C [164]. An old experiment [165] serves as a reminder: Heating three amino acids (combinations of glycine, alanine, valine, lysine, asparagine and glutamine) without water to temperatures between 160 °C to 200 °C for 4 to 6 h or to temperatures between 180 °C to 350 °C for 1 to 2 h produces, inter alia, the compound shown in Figure 6.

There has to be something innately natural to the chemistry of life that its chemical constituents will assemble effortlessly when the right conditions and catalysts are found. Perhaps reducing CO_2 and N_2 under chemical industry conditions milder than Haber–Bosch, with catalysts derived from H_2 and Fe_3O_4, will deliver a hydrothermal version of the Miller–Urey experiment, with a diversity of relevant nitrogenous compounds. This is an experiment that hydrothermal theories have been missing. The surface structure, morphology, facets, and impurities of catalysts should be also taken into account as crucial parameters. Concomitant reduction of N_2 and CO_2 might yield a diversity of compounds relevant to life more readily than adding NH_3 to reduced carbon compounds. Regardless, by better

understanding the sequence of chemical events within hydrothermal systems we should gain insights into catalysts that can further close the gap between rock–water–carbon interaction in vents, and life.

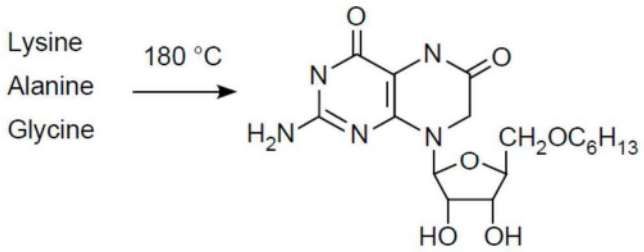


Figure 6. Pterin riboside from amino acids. One of the products obtained by Heinz et al. [165] without catalysts from dry heating of three amino acids is shown. Pterins are important cofactors in the acetyl-CoA pathway [83,86,166]. Note the *N*-glycosidic bond of the heterocyclic to ribose.

Funding: The authors thank the following agencies for funding: The Deutsche Forschungsgemeinschaft (TU 315/8-1 to HT and MA 1426/21-1 to WFM); the ERC (666053 to WFM and a StG to FLS); the VW Foundation (93 046 to WFM); the Wiener Wissenschaft und Technologiefonds (VRG15-007 to FLS); the National Science Foundation (OCE-1536702 to SQL); and the NASA Astrobiology Institute Cooperative Agreement (NNA15BB02A to TMM).

Acknowledgments: We thank Marcella Bonchio (Padua) and Joseph Moran (Strasbourg) for discussions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Corliss, J.B.; Baross, J.A.; Hoffman, S.E. A hypothesis concerning the relationships between submarine hot springs and the origin of life on earth. *Oceanol. Acta* **1981**, 59–69. Available online: http://archimer.ifremer.fr/doc/00245/35661/34170.pdf (accessed on 22 September 2018).
2. Baross, J.A.; Hoffman, S.E. Submarine hydrothermal vents and associated gradient environments as sites for the origin and evolution of life. *Orig. Life Evol. Biosph.* **1985**, *15*, 327–345. [CrossRef]
3. Holm, N.G. Why are Hydrothermal Systems Proposed as Plausible Environments for the Origin of Life? In *Marine Hydrothermal Systems and the Origin of Life*; Holm, N.G., Ed.; Springer: Dordrecht, The Netherlands, 1992; pp. 5–14.
4. Nisbet, E.; Sleep, N.H. The habitat and nature of early life. *Nature* **2001**, *409*, 1083–1091. [CrossRef] [PubMed]
5. McCollom, T.M.; Seewald, J.S. Serpentinites, hydrogen, and life. *Elements* **2013**, *9*, 129–134. [CrossRef]
6. 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*, 377–402. [CrossRef]
7. Kelley, D.S.; Baross, J.A.; Delaney, J.R. Volcanoes, fluids, and life at mid-ocean ridge spreading centers. *Annu. Rev. Earth Planet. Sci.* **2002**, *30*, 385–491. [CrossRef]
8. Freeland, S.J.; Hurst, L.D. The genetic code is one in a million. *J. Mol. Evol.* **1998**, *47*, 238–248. [CrossRef] [PubMed]
9. Tashiro, T.; Ishida, A.; Hori, M.; Igisu, M.; Koike, M.; Méjean, P.; Takahata, N.; Sano, Y.; Komiya, T. Early trace of life from 3.95 Ga sedimentary rocks in Labrador, Canada. *Nature* **2017**, *549*, 516–518. [CrossRef] [PubMed]
10. Ueno, Y.; Yamada, K.; Yoshida, N.; Maruyama, S.; Isozaki, Y. Evidence from fluid inclusions for microbial methanogenesis in the early Archaean era. *Nature* **2006**, *440*, 516–519. [CrossRef] [PubMed]
11. Cleaves, H.J., II. The prebiotic geochemistry of formaldehyde. *Precambrian Res.* **2008**, *164*, 111–118. [CrossRef]
12. Patel, B.H.; Percivalle, C.; Ritson, D.J.; Duffy, C.D.; Sutherland, J.D. Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism. *Nat. Chem.* **2015**, *7*, 301–307. [CrossRef] [PubMed]
13. Russell, M.J.; Nitschke, W.; Branscomb, E. The inevitable journey to being. *Philos. Trans. R. Soc. B Biol. Sci.* **2013**, *368*, 20120254. [CrossRef] [PubMed]
14. Cockell, C.S. The origin and emergence of life under impact bombardment. *Philos. Trans. R. Soc. B* **2006**, *362*, 1845–1856. [CrossRef] [PubMed]

15. Furukawa, Y.; Sekine, T.; Oba, M.; Kakegawa, T.; Nakazawa, H. Biomolecule formation by oceanic impacts on early Earth. *Nat. Geosci.* **2009**, *2*, 62–66. [CrossRef]
16. Russell, M. First Life. *Am. Sci.* **2006**, *94*, 32–39. [CrossRef]
17. Jackson, J.B. Natural pH gradients in hydrothermal alkali vents were unlikely to have played a role in the origin of life. *J. Mol. Evol.* **2016**, *83*, 1–11. [CrossRef] [PubMed]
18. Lane, N. Proton gradients at the origin of life. *BioEssays* **2017**, *39*, 1600217. [CrossRef] [PubMed]
19. Shock, E.L.; Helgeson, H.C. Calculation of the thermodynamic and transport properties of aqueous species at high pressures and temperatures: Correlation algorithms for ionic species and equation of state predictions to 5 kb and 1000 °C. *Geochim. Cosmochim. Acta* **1988**, *52*, 2009–2036. [CrossRef]
20. Shock, E.L. Geochemical constraints on the origin of organic compounds in hydrothermal systems. *Orig. Life Evol. Biosph.* **1990**, *20*, 331–367. [CrossRef]
21. Konn, C.; Charlou, J.L.; Holm, N.G.; Mousis, O. The production of methane, hydrogen, and organic compounds in ultramafic-hosted hydrothermal vents of the Mid-Atlantic Ridge. *Astrobiology* **2015**, *15*, 381–399. [CrossRef] [PubMed]
22. McCollom, T.M.; Sherwood Lollar, B.; Lacrampe-Couloume, G.; Seewald, J.S. The influence of carbon source on abiotic organic synthesis and carbon isotope fractionation under hydrothermal conditions. *Geochim. Cosmochim. Acta* **2010**, *74*, 2717–2740. [CrossRef]
23. Schrenk, M.O.; Brazelton, W.J.; Lang, S.Q. Serpentinization, carbon, and deep life. *Rev. Mineral. Geochem.* **2013**, *75*, 575–606. [CrossRef]
24. Martin, W.; Russell, M.J. On the origin of biochemistry at an alkaline hydrothermal vent. *Philos. Trans. R. Soc. B Biol. Sci.* **2007**, *362*, 1887–1925. [CrossRef] [PubMed]
25. Martin, W.; Baross, J.; Kelley, D.; Russell, M.J. Hydrothermal vents and the origin of life. *Nat. Rev. Microbiol.* **2008**, *6*, 805–814. [CrossRef] [PubMed]
26. Ragsdale, S.W. Nickel-based enzyme systems. *J. Biol. Chem.* **2009**, *284*, 18571–18575. [CrossRef] [PubMed]
27. Can, M.; Armstrong, F.A.; Ragsdale, S.W. Structure, function, and mechanism of the nickel metalloenzyme, CO dehydrogenase, and acetyl-CoA synthase. *Bioinorg. Enzymol.* **2014**, *114*, 4149–4174. [CrossRef] [PubMed]
28. Lancaster, K.M.; Roemelt, M.; Ettenhuber, P.; Hu, Y.; Ribbe, M.W.; Neese, F.; Bergmann, U.; DeBeer, S. X-ray emission spectroscopy evidences a central carbon in the nitrogenase iron-molybdenum cofactor. *Science* **2011**, *334*, 974–977. [CrossRef] [PubMed]
29. Spatzal, T.; Aksoyoglu, M.; Zhang, L.; Andrade, S.L.A.; Schleicher, E.; Weber, S.; Rees, D.C.; Einsle, O. Evidence for interstitial carbon in nitrogenase FeMo cofactor. *Science* **2011**, *334*, 940. [CrossRef] [PubMed]
30. Wei, J.; Ge, Q.; Yao, R.; Wen, Z.; Fang, C.; Guo, L.; Xu, H.; Sun, J. Directly converting CO_2 into a gasoline fuel. *Nat. Commun.* **2017**, *8*, 15174. [CrossRef] [PubMed]
31. Ertl, G. Primary steps in catalytic synthesis of ammonia. *J. Vac. Sci. Technol. A Vac. Surf. Films* **1983**, *1*, 1247–1253. [CrossRef]
32. Kandemir, T.; Schuster, M.E.; Senyshyn, A.; Behrens, M.; Schlögl, R. The Haber–Bosch process revisited: On the real structure and stability of "ammonia iron" under working conditions. *Angew. Chem. Int. Ed.* **2013**, *52*, 12723–12726. [CrossRef] [PubMed]
33. McCollom, T.M.; Seewald, J.S. Abiotic synthesis of organic compounds in deep-sea hydrothermal environments. *Chem. Rev.* **2007**, *107*, 382–401. [CrossRef] [PubMed]
34. Hennet, R.J.-C.; Holm, N.G.; Engel, M.H. Abiotic synthesis of amino acids under hydrothermal conditions and the origin of life: A perpetual phenomenon? *Naturwissenschaften* **1992**, *79*, 361–365. [CrossRef] [PubMed]
35. Shock, E.; Schulte, M.D. Organic synthesis during fluid mixing in hydrothermal systems. *J. Geophys. Res.* **1998**, *103*, 28513–28527. [CrossRef]
36. Amend, J.P.; Shock, E.L. Energetics of amino acid synthesis in hydrothermal ecosystems. *Science* **1998**, *281*, 1659–1662. [CrossRef] [PubMed]
37. Amend, J.P.; McCollom, T.M. Energetics of biomolecule synthesis on early Earth. In *Chemical Evolution II: From the Origins of Life to Modern Soc*; ACS Publications: Washington, DC, USA, 2009; pp. 63–94.
38. Sleep, N.H.; Bird, D.K.; Pope, E.C. Serpentinite and the dawn of life. *Philos. Trans. R. Soc. B Biol. Sci.* **2011**, *366*, 2857–2869. [CrossRef] [PubMed]
39. Arndt, N.T.; Nisbet, E.G. Processes on the young Earth and the habitats of early life. *Annu. Rev. Earth Planet. Sci.* **2012**, *40*, 521–549. [CrossRef]

40. Zahnle, K.; Arndt, N.; Cockell, C.; Halliday, A.; Nisbet, E.; Selsis, F.; Sleep, N.H. Emergence of a habitable planet. *Space Sci. Rev.* **2007**, *129*, 35–78. [CrossRef]
41. Dasgupta, R. Ingassing, storage, and outgassing of terrestrial carbon through geologic time. *Rev. Mineral. Geochem.* **2013**, *75*, 183–229. [CrossRef]
42. Etiope, G.; Schoell, M. Abiotic gas: Atypical, but not rare. *Elements* **2014**, *10*, 291–296. [CrossRef]
43. Bottke, W.F.; Norman, M.D. The late heavy bombardment. *Annu. Rev. Earth Planet. Sci.* **2017**, *45*, 619–647. [CrossRef]
44. Litasov, K.D.; Shatskiy, A.F. Composition of the Earth's core: A review. *Russ. Geol. Geophys.* **2016**, *57*, 22–46. [CrossRef]
45. Kasting, J.F.; Holm, N.G. What determines the volume of the oceans? *Earth Planet. Sci. Lett.* **1992**, *109*, 507–515. [CrossRef]
46. Westall, F.; Brack, A. The importance of water for life. *Space Sci. Rev.* **2018**, *214*, 1–23. [CrossRef]
47. Korenaga, J.; Planavsky, N.J.; Evans, D.A.D. Global water cycle and the coevolution of the Earth's interior and surface environment. *Philos. Trans. R. Soc. A* **2017**, *375*, 20150393. [CrossRef] [PubMed]
48. McCollom, T.M.; Bach, W. Thermodynamic constraints on hydrogen generation during serpentinization of ultramafic rocks. *Geochim. Cosmochim. Acta* **2009**, *73*, 856–875. [CrossRef]
49. Klein, F.; Bach, W.; McCollom, T.M. Compositional controls on hydrogen generation during serpentinization of ultramafic rocks. *Lithos* **2013**, *178*, 55–69. [CrossRef]
50. Holm, N.G. Hydrothermal activity and the volume of the oceans. *Deep. Res. Part II Top. Stud. Oceanogr.* **1996**, *43*, 47–52. [CrossRef]
51. Russell, M.J.; Hall, A.J.; Martin, W. Serpentinization as a source of energy at the origin of life. *Geobiology* **2010**, *8*, 355–371. [CrossRef] [PubMed]
52. Holm, N.G.; Oze, C.; Mousis, O.; Waite, J.H.; Guilbert-Lepoutre, A. Serpentinization and the formation of H_2 and CH_4 on celestial bodies (planets, moons, comets). *Astrobiology* **2015**, *15*, 587–600. [CrossRef] [PubMed]
53. Klein, F.; Bach, W. Fe-Ni-Co-O-S phase relations in peridotite–seawater interactions. *J. Petrol.* **2009**, *50*, 37–59. [CrossRef]
54. Foustoukos, D.I.; Bizimis, M.; Frisby, C.; Shirey, S.B. Redox controls on Ni-Fe-PGE mineralization and Re/Os fractionation during serpentinization of abyssal peridotite. *Geochim. Cosmochim. Acta* **2015**, *150*, 11–25. [CrossRef]
55. Sleep, N.H.; Meibom, A.; Fridriksson, T.; Coleman, R.G.; Bird, D.K. H_2-rich fluids from serpentinization: Geochemical and biotic implications. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 12818–12823. [CrossRef] [PubMed]
56. Lang, S.Q.; Früh-Green, G.; Bernasconi, S.M.; Brazelton, W.J.; Schrenk, M.O.; McGonigle, J.M. Deeply-sourced formate fuels sulfate reducers but not methanogens at Lost City hydrothermal field. *Sci. Rep.* **2018**, *8*, 755. [CrossRef] [PubMed]
57. Janecky, D.R.; Seyfried, W.E. Hydrothermal serpentinization of peridotite within the oceanic crust: Experimental investigations of mineralogy and major element chemistry. *Geochim. Cosmochim. Acta* **1986**, *50*, 1357–1378. [CrossRef]
58. Palandri, J.L.; Reed, M.H. Geochemical models of metasomatism in ultramafic systems: Serpentinization, rodingitization, and sea floor carbonate chimney precipitation. *Geochim. Cosmochim. Acta* **2004**, *68*, 1115–1133. [CrossRef]
59. Charlou, J.L.; Donval, J.P.; Fouquet, Y.; Jean-Baptiste, P.; Holm, N. Geochemistry of high H_2 and CH_4 vent fluids issuing from ultramafic rocks at the Rainbow hydrothermal field (36°14′N,MAR). *Chem. Geol.* **2002**, *191*, 345–359. [CrossRef]
60. Kelley, D.S.; Karson, J.A.; Früh-Green, G.L.; Yoerger, D.R.; Shank, T.M.; Butterfield, D.A.; Hayes, J.M.; Schrenk, M.O.; Olson, E.J.; Proskurowski, G.; et al. A serpentinite-hosted ecosystem: The Lost City Hydrothermal Field. *Science* **2005**, *307*, 1428–1434. [CrossRef] [PubMed]
61. Seyfried, W.E.; Foustoukos, D.I.; Fu, Q. Redox evolution and mass transfer during serpentinization: An experimental and theoretical study at 200 °C, 500 bar with implications for ultramafic-hosted hydrothermal systems at Mid-Ocean Ridges. *Geochim. Cosmochim. Acta* **2007**, *71*, 3872–3886. [CrossRef]
62. McCollom, T.M.; Klein, F.; Robbins, M.; Moskowitz, B.; Berquó, T.S.; Jöns, N.; Bach, W.; Templeton, A. Temperature trends for reaction rates, hydrogen generation, and partitioning of iron during experimental serpentinization of olivine. *Geochim. Cosmochim. Acta* **2016**, *181*, 175–200. [CrossRef]

63. Ueda, H.; Shibuya, T.; Sawaki, Y.; Saitoh, M.; Takai, K.; Maruyama, S. Reactions between komatiite and CO_2-rich seawater at 250 and 350 °C, 500 bars: Implications for hydrogen generation in the Hadean seafloor hydrothermal system. *Prog. Earth Planet. Sci.* **2016**, *3*, 35. [CrossRef]
64. McCollom, T.M.; Seewald, J.S. A reassessment of the potential for reduction of dissolved CO_2 to hydrocarbons during serpentinization of olivine. *Geochim. Cosmochim. Acta* **2001**, *65*, 3769–3778. [CrossRef]
65. Früh-Green, G.L.; Connolly, J.A.D.; Plas, A.; Kelley, D.S.; Grobety, B. Serpentinization of oceanic peridotites: Implications for geochemical cycles and biological activity. In *The Subseafloor Biosphere at Mid-Ocean Ridges Geophysical Monograph Series 144*; Wilcock, W.S.D., Delong, E.F., Kelley, D.S., Baross, J.A., Craig, C.S., Eds.; American Geophysical Union: Washington, DC, USA, 2004; pp. 119–136.
66. Frost, R. On the stability of sulfides, oxides, and native metals in serpentinite. *J. Petrol.* **1985**, *26*, 31–63. [CrossRef]
67. Alt, J.C.; Shanks, W.C. Sulfur in serpentinized oceanic peridotites: Serpentinization processes and microbial sulfate reduction. *J. Geophys. Res. Solid Earth* **1998**, *103*, 9917–9929. [CrossRef]
68. McCollom, T.M. Abiotic methane formation during experimental serpentinization of olivine. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 13965–13970. [CrossRef] [PubMed]
69. Früh-Green, G.L.; Kelley, D.S.; Bernasconi, S.M.; Karson, J.A.; Ludwig, K.A.; Butterfield, D.A.; Boschi, C.; Proskurowski, G. 30,000 Years of hydrothermal activity at the Lost City vent field. *Science* **2003**, *301*, 495–498. [CrossRef] [PubMed]
70. Ludwig, K.A.; Shen, C.C.; Kelley, D.S.; Cheng, H.; Edwards, R.L. U-Th systematics and 230th ages of carbonate chimneys at the Lost City Hydrothermal Field. *Geochim. Cosmochim. Acta* **2011**, *75*, 1869–1888. [CrossRef]
71. Proskurowski, G.; Lilley, M.D.; Seewald, J.S.; Früh-Green, G.L.; Olson, E.J.; Lupton, J.E.; Sylva, S.P.; Kelley, D.S. Abiogenic hydrocarbon production at Lost City hydrothermal field. *Science* **2008**, *319*, 319–607. [CrossRef] [PubMed]
72. Lang, S.Q.; Butterfield, D.A.; Schulte, M.; Kelley, D.S.; Lilley, M.D. Elevated concentrations of formate, acetate and dissolved organic carbon found at the Lost City hydrothermal field. *Geochim. Cosmochim. Acta* **2010**, *74*, 941–952. [CrossRef]
73. Holm, N.G.; Charlou, J.L. Initial indications of abiotic formation of hydrocarbons in the Rainbow ultramafic hydrothermal system, Mid-Atlantic Ridge. *Earth Planet. Sci. Lett.* **2001**, *191*, 1–8. [CrossRef]
74. McDermott, J.M.; Seewald, J.S.; German, C.R.; Sylva, S.P. Pathways for abiotic organic synthesis at submarine hydrothermal fields. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 7668–7672. [CrossRef] [PubMed]
75. Etiope, G.; Sherwood Lollar, B. Abiotic methane on Earth. *Rev. Geophys.* **2013**, *51*, 276–299. [CrossRef]
76. Horita, J.; Berndt, M.E. Abiogenic methane formation and isotopic fractionation under hydrothermal conditions. *Sci. Rep.* **1999**, *285*, 1055–1057. [CrossRef]
77. Heinen, W.; Lauwers, A. The iron-sulfur world and the origins of life: Abiotic thiol synthesis from metallic iron, H_2S and CO_2; a comparison of the thiol generating $FeS/HCl(H_2S)/CO_2$-system and its $Fe^0/H_2S/CO_2$-counterpart. *Proc. K. Ned. Akad. Van Wet.* **1997**, *100*, 11–25.
78. Guan, G.; Kida, T.; Ma, T.; Kimura, K.; Abe, E.; Yoshida, A. Reduction of aqueous CO_2 at ambient temperature using zero-valent iron-based composites. *Green Chem.* **2003**, *5*, 630. [CrossRef]
79. He, C.; Tian, G.; Liu, Z.; Feng, S. A mild hydrothermal route to fix carbon dioxide to simple carboxylic acids. *Org. Lett.* **2010**, *12*, 649–651. [CrossRef] [PubMed]
80. Varma, S.J.; Muchowska, K.B.; Chatelain, P.; Moran, J. Native iron reduces CO_2 to intermediates and endproducts of the acetyl-CoA pathway. *Nat. Ecol. Evol.* **2018**, *2*, 1019–1024. [CrossRef] [PubMed]
81. Thauer, R.K.; Kaster, A.K.; Seedorf, H.; Buckel, W.; Hedderich, R. Methanogenic archaea: Ecologically relevant differences in energy conservation. *Nat. Rev. Microbiol.* **2008**, *6*, 579–591. [CrossRef] [PubMed]
82. Müller, V. Energy conservation in acetogenic bacteria. *Appl. Environ. Microbiol.* **2003**, *69*, 6345–6353. [CrossRef] [PubMed]
83. Fuchs, G. Alternative pathways of carbon dioxide fixation: Insights into the early evolution of life? *Annu. Rev. Microbiol.* **2011**, *65*, 631–658. [CrossRef] [PubMed]
84. Berg, I.A.; Kockelkorn, D.; Buckel, W.; Fuchs, G. A 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway in archaea. *Science* **2007**, *318*, 1782–1786. [CrossRef] [PubMed]
85. Fuchs, G. CO_2 fixation in acetogenic bacteria: Variations on a theme. *FEMS Microbiol. Lett.* **1986**, *39*, 181–213. [CrossRef]

86. Sousa, F.L.; Martin, W.F. Biochemical fossils of the ancient transition from geoenergetics to bioenergetics in prokaryotic one carbon compound metabolism. *Biochim. Biophys. Acta Bioenerg.* **2014**, *1837*, 964–981. [CrossRef] [PubMed]
87. Weiss, M.C.; Sousa, F.L.; Mrnjavac, N.; Neukirchen, S.; Roettger, M.; Nelson-Sathi, S.; Martin, W.F. The physiology and habitat of the last universal common ancestor. *Nat. Microbiol.* **2016**, *1*, 16116. [CrossRef] [PubMed]
88. Weiss, M.C.; Preiner, M.; Xavier, J.C.; Zimorski, V.; Martin, F. The last universal common ancestor between ancient Earth chemistry and the onset of genetics. *PLoS Genet.* **2018**, *14*, e1007518. [CrossRef] [PubMed]
89. Früh-Green, G.L.; Orcutt, B.N.; Green, S.L.; Cotterill, C.; Morgan, S.; Akizawa, N.; Bayrakci, G.; Behrmann, J.-H.; Boschi, C.; Brazelton, W.J.; et al. Northern sites. *Proc. Int. Ocean Discov. Progr.* **2017**, *357*. [CrossRef]
90. Smirnov, A.; Hausner, D.; Laffers, R.; Strongin, D.R.; Schoonen, M.A.A. Abiotic ammonium formation in the presence of Ni-Fe metals and alloys and its implications for the Hadean nitrogen cycle. *Geochem. Trans.* **2008**, *9*, 5–20. [CrossRef] [PubMed]
91. Brandes, J.A.; Boctor, N.Z.; Cody, G.D.; Cooper, B.A.; Hazen, R.M.; Yoder, H.S. Abiotic nitrogen reduction on the early Earth. *Nature* **1998**, *395*, 365–367. [CrossRef] [PubMed]
92. Dörr, M.; Käßbohrer, J.; Grunert, R.; Kreisel, G.; Brand, W.A.; Werner, R.A.; Geilmann, H.; Apfel, C.; Robl, C.; Weigand, W. A possible prebiotic formation of ammonia from dinitrogen on iron sulfide surfaces. *Angew. Chem. Int. Ed.* **2003**, *42*, 1540–1543. [CrossRef] [PubMed]
93. Holm, N.G.; Neubeck, A. Reduction of nitrogen compounds in oceanic basement and its implications for HCN formation and abiotic organic synthesis. *Geochem. Trans.* **2009**, *10*, 9. [CrossRef] [PubMed]
94. Bradley, A.S. The sluggish speed of making abiotic methane. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 13944–13946. [CrossRef] [PubMed]
95. Wang, D.T.; Reeves, E.P.; McDermott, J.M.; Seewald, J.S.; Ono, S. Clumped isotopologue constraints on the origin of methane at seafloor hot springs. *Geochim. Cosmochim. Acta* **2018**, *223*, 141–158. [CrossRef]
96. Seewald, J.S.; Zolotov, M.Y.; McCollom, T. Experimental investigation of single carbon compounds under hydrothermal conditions. *Geochim. Cosmochim. Acta* **2006**, *70*, 446–460. [CrossRef]
97. Miller, H.M.; Mayhew, L.E.; Ellison, E.T.; Kelemen, P.; Kubo, M.; Templeton, A.S. Low temperature hydrogen production during experimental hydration of partially-serpentinized dunite. *Geochim. Cosmochim. Acta* **2017**, *209*, 161–183. [CrossRef]
98. Oparin, A.I. *Origin of Life*; Dover Publications: New York, NY, USA, 1936; ISBN 0-486-60213-3.
99. Dictor, R.A.; Bell, A.T. Fischer–Tropsch synthesis over reduced and unreduced iron oxide catalysts. *J. Catal.* **1986**, *97*, 121–136. [CrossRef]
100. Wu, B.; Tian, L.; Xiang, H.; Zhang, Z.; Li, Y.W. Novel precipitated iron Fischer–Tropsch catalysts with Fe_3O_4 coexisting with α-Fe2O3. *Catal. Lett.* **2005**, *102*, 211–218. [CrossRef]
101. McCollom, T.M. Miller-Urey and beyond: What have we learned about prebiotic organic synthesis reactions in the past 60 years? *Annu. Rev. Earth Planet. Sci.* **2013**, *41*, 207–229. [CrossRef]
102. Holm, N.G.; Dumont, M.; Ivarsson, M.; Konn, C. Alkaline fluid circulation in ultramafic rocks and formation of nucleotide constituents: A hypothesis. *Geochem. Trans.* **2006**, *7*, 14–16. [CrossRef] [PubMed]
103. McCollom, T.M. Laboratory simulations of abiotic hydrocarbon formation in Earth's deep subsurface. *Rev. Mineral. Geochem.* **2013**, *75*, 467–494. [CrossRef]
104. Porosoff, M.D.; Kattel, S.; Li, W.; Liu, P.; Chen, J.G. Identifying trends and descriptors for selective CO_2 conversion to CO over transition metal carbides. *Chem. Commun.* **2015**, *51*, 6988. [CrossRef] [PubMed]
105. Rodriguez, J.A.; Liu, P.; Stacchiola, D.J.; Senanayake, S.D.; White, M.G.; Chen, J.G. Hydrogenation of CO_2 to methanol: Importance of metal-oxide and metal-carbide interfaces in the activation of CO_2. *ACS Catal.* **2015**, *5*, 6696–6706. [CrossRef]
106. Riedel, T.; Schulz, H.; Schaub, G.; Jun, K.W.; Hwang, J.S.; Lee, K.W. Fischer–Tropsch on iron with H_2/CO and H_2/CO_2 as synthesis gases: The episodes of formation of the Fischer–Tropsch regime and construction of the catalyst. *Top. Catal.* **2003**, *26*, 41–54. [CrossRef]
107. Pichler, H. Twenty-five years of synthesis of gasoline by catalytic conversion of carbon monoxide and hydrogen. In *Advances in Catalysis*; Frankenburg, W., Rideal, E., Komarewsky, V., Eds.; Academic Press Inc.: New York, NY, USA, 1952; p. 271. ISBN 9780120078042.
108. Dry, M.E. The Fischer–Tropsch process: 1950–2000. *Catal. Today* **2002**, *71*, 227–241. [CrossRef]

109. Goodrich, C.A.; Bird, J.M. Formation of iron-carbon alloys in basaltic magma at Uivaq, Disko Island: The role of carbon in mafic magmas. *J. Geol.* **1985**, *93*, 475–492. [CrossRef]
110. Shi, N.; Bai, W.; Li, G.; Xiong, M.; Fang, Q.; Yang, J.; Ma, Z.; Rong, H. Yarlongite: A new metallic carbide mineral. *Acta Geol. Sin.* **2008**, *83*, 52–56. [CrossRef]
111. Horita, J.; Polyakov, V.B. Carbon-bearing iron phases and the carbon isotope composition of the deep Earth. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 31–36. [CrossRef] [PubMed]
112. Chen, Q.W.; Bahnemann, D.W. Reduction of carbon dioxide by magnetite: Implications for the primordial synthesis of organic molecules. *J. Am. Chem. Soc.* **2000**, *122*, 970–971. [CrossRef]
113. Russell, M.J.; Martin, W. The rocky roots of the acetyl-CoA pathway. *Trends Biochem. Sci.* **2004**, *29*, 358–363. [CrossRef] [PubMed]
114. Buckel, W.; Thauer, R.K. Flavin-based electron bifurcation, ferredoxin, flavodoxin, and anaerobic respiration with protons (Ech) or NAD+(Rnf) as electron acceptors: A historical review. *Front. Microbiol.* **2018**, *9*, 401. [CrossRef] [PubMed]
115. Schuchmann, K.; Müller, V. Direct and reversible hydrogenation of CO_2 to formate by a bacterial carbon dioxide reductase. *Science* **2013**, *342*, 1382–1385. [CrossRef] [PubMed]
116. Herrmann, G.; Jayamani, E.; Mai, G.; Buckel, W. Energy conservation via electron-transferring flavoprotein in anaerobic bacteria. *J. Bacteriol.* **2008**, *190*, 784–791. [CrossRef] [PubMed]
117. Müller, V.; Chowdhury, N.P.; Basen, M. Electron bifurcation: A long- hidden energy-coupling mechanism. *Annu. Rev. Microbiol.* **2018**, *72*, 331–353. [CrossRef] [PubMed]
118. Kaster, A.-K.; Moll, J.; Parey, K.; Thauer, R.K. Coupling of ferredoxin and heterodisulfide reduction via electron bifurcation in hydrogenotrophic methanogenic archaea. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 2981–2986. [CrossRef] [PubMed]
119. Mall, A.; Sobotta, J.; Huber, C.; Tschirner, C.; Kowarschik, S.; Bačnik, K.; Mergelsberg, M.; Boll, M.; Hügler, M.; Eisenreich, W.; et al. Reversibility of citrate synthase allows autotrophic growth of a thermophilic bacterium. *Science* **2018**, *359*, 563–567. [CrossRef] [PubMed]
120. 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*, 559–563. [CrossRef] [PubMed]
121. 1Martin, W.F.; Thauer, R.K. Energy in ancient metabolism. *Cell* **2017**, *168*, 953–955. [CrossRef] [PubMed]
122. Xavier, J.C.; Preiner, M.; Martin, W.F. Something special about CO. *FEBS J.* **2018**, in press. [CrossRef]
123. Dos Santos, P.C.; Igarashi, R.Y.; Lee, H.; Hoffman, B.M.; Seefeldt, L.C.; Dean, D.R. Substrate interactions with the nitrogenase active site. *Acc. Chem. Res.* **2005**, *38*, 208–214. [CrossRef] [PubMed]
124. Burgess, B.K.; Lowe, D.J. Mechanism of molybdenum nitrogenase. *Chem. Rev.* **1996**, *96*, 2983–3011. [CrossRef] [PubMed]
125. Zhang, L.; Kaiser, J.T.; Meloni, G.; Yang, K.-Y.; Spatzal, T.; Andrade, S.L.A.; Einsle, O.; Howard, J.B.; Rees, D.C. The 16th Fe in the nitrogenase MoFe–protein. *Angew. Chem. Int. Ed.* **2013**, *52*, 10529–10532. [CrossRef] [PubMed]
126. Buckel, W.; Hetzel, M.; Kim, J. ATP-driven electron transfer in enzymatic radical reactions. *Curr. Opin. Chem. Biol.* **2004**, *8*, 462–467. [CrossRef] [PubMed]
127. Wiig, J.A.; Hu, Y.; Ribbe, M.W. Refining the pathway of carbide insertion into the nitrogenase M-cluster. *Nat. Commun.* **2015**, *6*, 8034. [CrossRef] [PubMed]
128. Lee, C.C.; Hu, Y.; Ribbe, M.W. Insights into hydrocarbon formation by nitrogenase cofactor homologs. *MBio* **2015**, *6*, e00307-15. [CrossRef] [PubMed]
129. Boyd, E.S.; Peters, J.W. New insights into the evolutionary history of biological nitrogen fixation. *Front. Microbiol.* **2013**, *4*, 201. [CrossRef] [PubMed]
130. Anderson, J.S.; Rittle, J.; Peters, J.C. Catalytic conversion of nitrogen to ammonia by an iron model complex. *Nature* **2013**, *501*, 84–87. [CrossRef] [PubMed]
131. Martin, W.F. Hydrogen, metals, bifurcating electrons, and proton gradients: The early evolution of biological energy conservation. *FEBS Lett.* **2012**, *586*, 485–493. [CrossRef] [PubMed]
132. Williams, T.A.; Szöllösi, G.J.; Spang, A.; Foster, P.G.; Heaps, S.E.; Boussau, B.; Ettema, T.J.G.; Embley, T.M. Integrative modeling of gene and genome evolution roots the archaeal tree of life. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E4602–E4611. [CrossRef] [PubMed]

133. Takami, H.; Noguchi, H.; Takaki, Y.; Uchiyama, I.; Toyoda, A.; Nishi, S.; Chee, G.; Arai, W.; Nunoura, T.; Itoh, T.; et al. A deeply branching thermophilic bacterium with an ancient acetyl-CoA pathway dominates a subsurface ecosystem. *PLoS ONE* **2012**, *7*, e30559. [CrossRef] [PubMed]
134. Martin, W.F.; Bryant, D.A.; Beatty, J.T. A physiological perspective on the origin and evolution of photosynthesis. *FEMS Microbiol. Rev.* **2018**, *42*, 205–231. [CrossRef] [PubMed]
135. Chapelle, F.H.; Neill, K.O.; Bradley, P.M.; Methe, B.A.; Ciufo, S.A.; Knobel, L.L.; Lovley, D.R. A hydrogen-based subsurface microbial community dominated by methanogens. *Nature* **2002**, *415*, 312–315. [CrossRef] [PubMed]
136. Lever, M.A.; Rouxel, O.; Alt, J.C.; Shimizu, N.; Ono, S.; Coggon, R.M.; Shanks, W.C.; Lapham, L.; Elvert, M.; Prieto-Mollar, X.; et al. Evidence for microbial carbon and sulfur cycling in deeply buried ridge flank basalt. *Science* **2013**, *339*, 1305–1308. [CrossRef] [PubMed]
137. Reeves, E.P.; Mcdermott, J.M.; Seewald, J.S. The origin of methanethiol in midocean ridge hydrothermal fluids. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 5474–5479. [CrossRef] [PubMed]
138. Schrenk, M.O.; Kelley, D.S.; Bolton, S.A.; Baross, J.A. Low archaeal diversity linked to subseafloor geochemical processes at the Lost City Hydrothermal Field, Mid-Atlantic Ridge. *Environ. Microbiol.* **2004**, *6*, 1086–1095. [CrossRef] [PubMed]
139. Brazelton, W.J.; Schrenk, M.O.; Kelley, D.S.; Baross, J.A. Methane- and sulfur-metabolizing microbial communities dominate the Lost City Hydrothermal Field ecosystem. *Appl. Environ. Microbiol.* **2006**, *72*, 6257–6270. [CrossRef] [PubMed]
140. Whitman, W.B.; Coleman, D.C.; Wiebe, W.J. Prokaryotes: The unseen majority. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6578–6583. [CrossRef] [PubMed]
141. Bar-On, Y.M.; Phillips, R.; Milo, R. The biomass distribution on Earth. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 6506–6511. [CrossRef] [PubMed]
142. Heberling, C.; Lowell, R.P.; Liu, L.; Fisk, M.R. Extent of the microbial biosphere in the oceanic crust. *Geochem. Geophys. Geosyst.* **2010**, *11*. [CrossRef]
143. Milucka, J.; Ferdelman, T.G.; Polerecky, L.; Franzke, D.; Wegener, G.; Schmid, M.; Lieberwirth, I.; Wagner, M.; Widdel, F.; Kuypers, M.M.M. Zero-valent sulphur is a key intermediate in marine methane oxidation. *Nature* **2012**, *491*, 541–546. [CrossRef] [PubMed]
144. Miller, S.L.; Bada, J.L. Submarine hot springs and the origin of life. *Nature* **1988**, *334*, 609–611. [CrossRef] [PubMed]
145. Kelley, D.S.; Karson, J.A.; Blackman, D.K.; Früh-Green, G.L.; Butterfield, D.A.; Lilley, M.D.; Olson, E.J.; Schrenk, M.O.; Roe, K.K.; Lebon, G.T.; et al. An off-axis hydrothermal vent field near the Mid-Atlantic Ridge at 30 degrees N. *Nature* **2001**, *412*, 145–149. [CrossRef] [PubMed]
146. Benner, S.A.; Kim, H.-J.; Carrigan, M.A. Asphalt, water, and the prebiotic synthesis of ribose, ribonucleosides, and RNA. *Acc. Chem. Res.* **2012**, *45*, 2025–2034. [CrossRef] [PubMed]
147. Ricardo, A.; Carrigan, M.A.; Olcott, A.N.; Benner, S.A. Borate minerals stabilize ribose. *Science* **2004**, *303*, 196. [CrossRef] [PubMed]
148. Schönheit, P.; Buckel, W.; Martin, W.F. On the origin of heterotrophy. *Trends Microbiol.* **2016**, *24*, 12–25. [CrossRef] [PubMed]
149. Stouthamer, A.H. Energy-yielding pathways. In *The Bacteria Vol 6*; Gunsalus, I.C., Ed.; Academic Press: New York, NY, USA, 1978; pp. 389–462.
150. Baross, J.A.; Martin, W.F. The ribofilm as a concept for life's origins. *Cell* **2015**, *162*, 13–15. [CrossRef] [PubMed]
151. Goldford, J.E.; Hartman, H.; Smith, T.F.; Segrè, D. Remnants of an ancient metabolism without phosphate. *Cell* **2017**, *168*, 1126–1134.e9. [CrossRef] [PubMed]
152. Goldford, J.E.; Segrè, D. Modern views of ancient metabolic networks. *Curr. Opin. Syst. Biol.* **2018**, *8*, 117–124. [CrossRef]
153. Dayhoff, M.O.; Eck, R.V. Evolution of the structure of ferredoxin based on surviving relics of primitive amino acid sequences. *Science* **1966**, *152*, 363–366. [CrossRef]
154. Wächtershäuser, G. Before enzymes and templates: Theory of surface metabolism. *Microbiol. Rev.* **1988**, *52*, 452–484. [PubMed]

155. Thauer, R.K.; Kaster, A.-K.; Goenrich, M.; Schick, M.; Hiromoto, T.; Shima, S. Hydrogenases from methanogenic archaea, nickel, a novel cofactor, and H_2 storage. *Annu. Rev. Biochem.* **2010**, *79*, 507–536. [CrossRef] [PubMed]
156. Thauer, R.K. Hydrogenases and the global H_2 cycle. *Eur. J. Inorg. Chem.* **2011**, *2011*, 919–921. [CrossRef]
157. Wächtershäuser, G. Groundworks for an evolutionary biochemistry: The iron-sulphur world. *Prog. Biophys. Mol. Biol.* **1992**, *58*, 85–201. [CrossRef]
158. Huber, C.; Eisenreich, W.; Wächtershäuser, G. Synthesis of α-amino and α-hydroxy acids under volcanic conditions: Implications for the origin of life. *Tetrahedron Lett.* **2010**, *51*, 1069–1071. [CrossRef]
159. Huber, C.; Wächtershäuser, G. Activated acetic acid by carbon fixation on (Fe,Ni)S under primordial conditions. *Science* **1997**, *276*, 245–248. [CrossRef] [PubMed]
160. 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*, 1716–1721. [CrossRef] [PubMed]
161. Semenov, S.N.; Kraft, L.J.; Ainla, A.; Zhao, M.; Baghbanzadeh, M.; Campbell, V.E.; Kang, K.; Fox, J.M.; Whitesides, G.M. Autocatalytic, bistable, oscillatory networks of biologically relevant organic reactions. *Nature* **2016**, *537*, 656–660. [CrossRef] [PubMed]
162. Schüth, F. A vibrant science lives from within. *Angew. Chem. Int. Ed.* **2016**, *55*, 14878–14879. [CrossRef] [PubMed]
163. Hickman-Lewis, K.; Cavalazzi, B.; Foucher, F.; Westall, F. Most ancient evidence for life in the Barberton greenstone belt: Microbial mats and biofabrics of the ~3.47 Ga Middle Marker horizon. *Precambrian Res.* **2018**, *312*, 45–67. [CrossRef]
164. Hayatsu, R.; Anders, E. Organic compounds in meteorites and their origins. *Top. Curr. Chem.* **1981**, *99*, 1–37. [CrossRef]
165. Heinz, V.B.; Ried, W.; Dose, K. Thermische Erzeugung von Pteridinen und Flavinen aus Aminosauregemischen. *Angew. Chem.* **1979**, *91*, 510–511. [CrossRef]
166. Maden, B.E.H. Tetrahydrofolate and tetrahydromethanopterin compared: Functionally distinct carriers in C1 metabolism. *Biochem. J.* **2000**, *350*, 609–629. [CrossRef] [PubMed]

Publication 7

Title: A hydrogen dependent geochemical analogue of primordial carbon and energy metabolism

Year: 2020

Authors: Martina Preiner, Kensuke Igarashi, Kamila B. Muchowska, Mingquan Yu, Sreejith J. Varma, Karl Kleinermanns, Masaru K. Nobu, Yoichi Kamagata, Harun Tüysüz, Joseph Moran, and William F. Martin

Published in: *Nature Ecology and Evolution.* 10.1038/s41559-020-1125-6.

Contribution: First author.
Major: wrote parts of first draft, assembled supplementary information, edited heavily throughout all review processes and coordinated the distribution and responses with all authors. Drafted, plotted and illustrated all figures except Fig. 6. Responsible for all Ni_3Fe related experiments and the development of the Fe_3O_4 experiments.

Summary: This research article realizes one major goal of all theoretical papers presented to this point throughout this thesis. H_2 dependent CO_2 fixation is shown on three different catalytic minerals, commonly found in hydrothermal systems in the Earth's crust. Direct parallels between the product spectrum of the experiments and the intermediates and products of the Wood-Ljungdahl pathway can be drawn, leading to the conclusion that such abiotic CO_2 fixation could have preceded enzyme dependent metabolic pathways.

nature ecology & evolution

ARTICLES

https://doi.org/10.1038/s41559-020-1125-6

Check for updates

A hydrogen-dependent geochemical analogue of primordial carbon and energy metabolism

Martina Preiner[1,8], Kensuke Igarashi[2,8], Kamila B. Muchowska[3,8], Mingquan Yu[4,8], Sreejith J. Varma[5], Karl Kleinermanns[6], Masaru K. Nobu[7], Yoichi Kamagata[7], Harun Tüysüz[4]✉, Joseph Moran[3]✉ and William F. Martin[1]✉

Hydrogen gas, H_2, is generated by alkaline hydrothermal vents through an ancient geochemical process called serpentinization, in which water reacts with iron-containing minerals deep within the Earth's crust. H_2 is the electron donor for the most ancient and the only energy-releasing route of biological CO_2 fixation, the acetyl-CoA pathway. At the origin of metabolism, CO_2 fixation by hydrothermal H_2 within serpentinizing systems could have preceded and patterned biotic pathways. Here we show that three hydrothermal minerals—greigite (Fe_3S_4), magnetite (Fe_3O_4) and awaruite (Ni_3Fe)—catalyse the fixation of CO_2 with H_2 at 100 °C under alkaline aqueous conditions. The product spectrum includes formate (up to 200 mM), acetate (up to 100 µM), pyruvate (up to 10 µM), methanol (up to 100 µM) and methane. The results shed light on both the geochemical origin of microbial metabolism and the nature of abiotic formate and methane synthesis in modern hydrothermal vents.

Organic synthesis in hydrothermal vents is relevant to life's origin because the reactions involve sustained energy release founded in the disequilibrium between CO_2 and the vast amounts of molecular hydrogen, H_2, generated in the Earth's crust during serpentinization[1–9]. Hydrogen has been a source of electrons and energy since there was liquid water on the early Earth, and it fuelled early anaerobic ecosystems in the Earth's crust[1,8,10]. In biochemistry, the acetyl-CoA pathway of CO_2 fixation uses the electrons and energy of H_2 to simultaneously supply three key requirements for life: reduced carbon in the form of acetyl groups, electrons in the form of reduced ferredoxin and ion gradients for energy conservation in the form of ATP[11,12]. The pathway is linear, not cyclic, it releases energy rather than requiring energy input and its enzymes are replete with primordial metal co-factors[13,14]. It traces to the last universal common ancestor[15], and abiotic, geochemical organic syntheses resembling segments of the pathway occur in modern hydrothermal vents[2,3]. Laboratory simulations of acetyl-CoA pathway reactions include the non-enzymatic synthesis of thioesters from CO and methylsulfide[16], and the synthesis of acetate[17] and pyruvate[18] from CO_2 using native iron or external electrochemical potentials[19] as the electron source. Enzymatic versions of those abiotic reactions occur in core energy metabolism of acetogens and methanogens[11–14], ancient anaerobic autotrophs that live from H_2 and CO_2 via the acetyl-CoA pathway and that still inhabit the crust today[14]. Although the enzymes that catalyse these modern microbial reactions have been widely investigated[11–14], catalysts promoting abiotic reactions in vents today, and that might have been instrumental at life's origin, are poorly understood[2]. A fully abiotic analogue of the acetyl-CoA pathway, from H_2 and CO_2 as it occurs in life, has not been reported to date.

To probe the mechanisms of hydrothermal metabolic reactions emulating ancient pathways, we investigated three different iron minerals that occur naturally in hydrothermal systems: greigite (Fe_3S_4), magnetite (Fe_3O_4) and the nickel iron alloy awaruite (Ni_3Fe). Magnetite and awaruite are common constituents of serpentinizing systems[20] and are more stable under alkaline conditions than greigite[21,22]. Magnetite, like H_2, is a main end product of serpentinization, being formed from water-dependent oxidation of iron(II) silicates[23]. In chemical industry, iron-based materials are the catalysts of choice for diverse industrial processes including Haber–Bosch (fixation of N_2) and Fischer–Tropsch syngas (CO and H_2) conversion to hydrocarbons[7]. Awaruite is an intermetallic compound that forms in serpentinizing systems at high-H_2 partial pressures and very low-H_2S fugacities[5,20], via the reduction of iron(II) and nickel(II) compounds. It is common in Ni-containing serpentinizing systems, where it is usually deposited as small grains[20]. Greigite is formed under conditions of high H_2S activity[5,21] as a transient intermediate in the conversion of mackinawite to pyrite[22,24]; it shares structural similarity with the iron sulfur clusters of many modern enzymes[6]. Iron sulfides can be found at the surface of hydrothermal vents either as small compartments[21] or as nanoparticles in hydrothermal plumes[25], as well as in meteorites[26]. Iron minerals have long been regarded as ancient catalysts[6,16,27], although the key initial reaction connecting the inorganic and organic world—CO_2 fixation with H_2 as the reductant—has not been reported using iron mineral catalysts under biologically relevant conditions[19].

Results

Although very different in structure and composition (Fig. 1), greigite, magnetite and awaruite are geochemically synthesized in hydrothermal systems from pre-existing divalent iron and nickel

[1]Institute for Molecular Evolution, University of Düsseldorf, Düsseldorf, Germany. [2]Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology, Sapporo, Japan. [3]ISIS (UMR 7006), University of Strasbourg, CNRS, Strasbourg, France. [4]Max-Planck-Institut für Kohlenforschung, Mülheim an der Ruhr, Germany. [5]Charité—Universitätsmedizin Berlin, Laboratory 'Biochemistry and System Biology of the Metabolism', Berlin, Germany. [6]Institute for Physical Chemistry, University of Düsseldorf, Düsseldorf, Germany. [7]Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan. [8]These authors contributed equally: Martina Preiner, Kensuke Igarashi, Kamila B. Muchowska, Mingquan Yu. ✉e-mail: tueysuez@kofo.mpg.de; moran@unistra.fr; bill@hhu.de

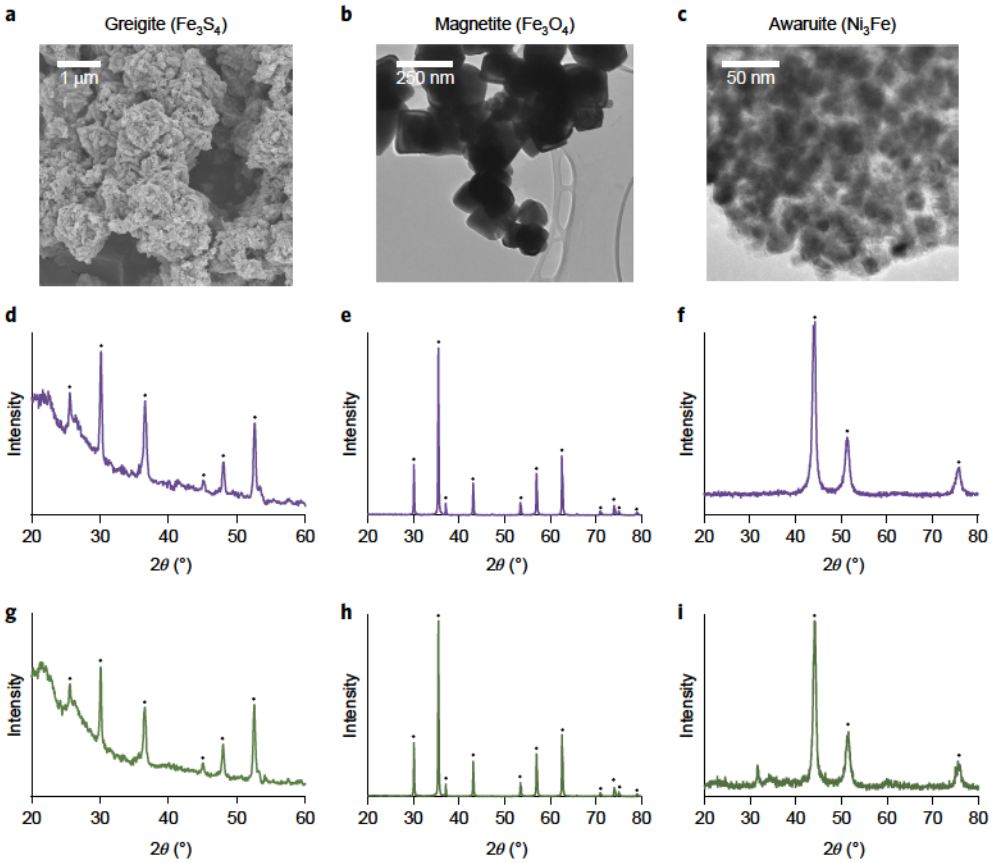


Fig. 1 | Characterization of greigite, magnetite and awaruite catalysts. a–c, The three powders are different in structure and morphology as seen from electron microscopy images, for which greigite (**a**) and awaruite (**c**) were freshly synthesized and magnetite (**b**) was commercially obtained. **d–i**, Comparison of XRD patterns of the minerals greigite (**d**), magnetite (**e**) and awaruite (**f**) before the reaction and after experiments under the following conditions: greigite for 24 h at pH 6.5, stabilized by a phosphate buffer under a H_2/CO_2 atmosphere (**g**); magnetite (**h**) and awaruite (**i**) for 16 h under alkaline conditions (KOH added) under a H_2/CO_2 atmosphere.

minerals during serpentinization[5,8,28]. X-ray diffraction (XRD) applied to our laboratory preparations of colloidal Fe_3S_4 and Ni_3Fe nanoparticles (for details of synthesis, see Methods), as well as commercial Fe_3O_4, reveals their characteristic patterns of crystal structures (Fig. 1).

Building on evidence for catalytic reactivity in previous reports[16–19], we investigated the ability of greigite, magnetite and awaruite to promote the reduction of CO_2 with H_2 in water. Under very mild hydrothermal conditions—at 100 °C under 2 bar H_2/CO_2 (80/20)—formate and acetate synthesis from H_2 and CO_2 occurs readily in nearly neutral and alkaline aqueous solution in the presence of Fe_3S_4 (Fig. 2a). While only formate was detected at 20 °C, formate and acetate were found at 60 °C, which is close to the temperature of vent effluent (ca. 70 °C) in the Lost City hydrothermal field (Fig. 2b)[29]. At 100 bar, Fe_3S_4 catalyses the synthesis of formate and methane from H_2 and CO_2 (Fig. 2c), but not from CO (Extended Data Fig. 4b). Here, methane and formate production is almost stoichiometric relative to hydrogen decrease. For 14 mM H_2 consumed, 1 mM of formate (1 H_2 per molecule of formate) and 2.3 mM of methane (4 H_2 per molecule of methane) are produced, leaving only 3.8 mM of H_2 that might be available for acetate synthesis (which was below detection level in this experimental set-up). At 2 bar, formate accumulates to >2 mM within 4 h while the detection of acetate requires 4–8 h (Fig. 2d). Notably, formate and methane are the main products of abiotic organic synthesis observed in the effluent of modern serpentinizing hydrothermal systems[9,30–33].

We found that magnetite, like greigite, catalyses the aqueous synthesis of formate and acetate in the range of 10 µM to 1 mM from H_2 and CO_2, but also the formation of methanol and pyruvate under mild (25 bar H_2/CO_2, 40/60 ratio and 100 °C) hydrothermal conditions (Fig. 3a). Pyruvate is a crucial intermediate of carbon and energy metabolism in nearly all microbes, and the main product of CO_2 fixation in autotrophs that use the acetyl-CoA pathway[11]. It accumulates at 5–10 µM in the presence of Fe_3O_4 across the pH range 6.0–10.0, when either native iron (Fe) or H_2 is used as the reductant (Fig. 3a). Magnetite generates a generally uniform product distribution across conditions tested, and also when smaller amounts of catalyst are used (Extended Data Fig. 6b). Additionally we investigated different amounts of Fe as a reductant, showing that its impact on product concentrations is low even when a large excess of Fe was used. Both Fe and Fe_3O_4 formed a solid disc after the reaction, which probably hindered further oxidation of Fe and thus further accumulation of reduced carbon compounds (Extended Data Fig. 7a).

At 100 °C, awaruite catalyses the synthesis of acetate and methanol in the range 10–100 µM at pH 5.0–8.0 whereby either the native alloy itself, H_2 or native Fe can function as the reductant, albeit with differing efficiency and product distribution (Fig. 3b). Under alkaline conditions, with either native Fe or H_2 as reductant, formate accumulates in the 200 mM range with 1 mmol of metal atoms as catalyst. Physical contact between awaruite and native iron is not required for product formation (Extended Data Fig. 7b). In the case of awaruite, lower temperatures improved pyruvate synthesis (Fig. 4a), similar to previous studies[18]. Pyruvate is formed under alkaline

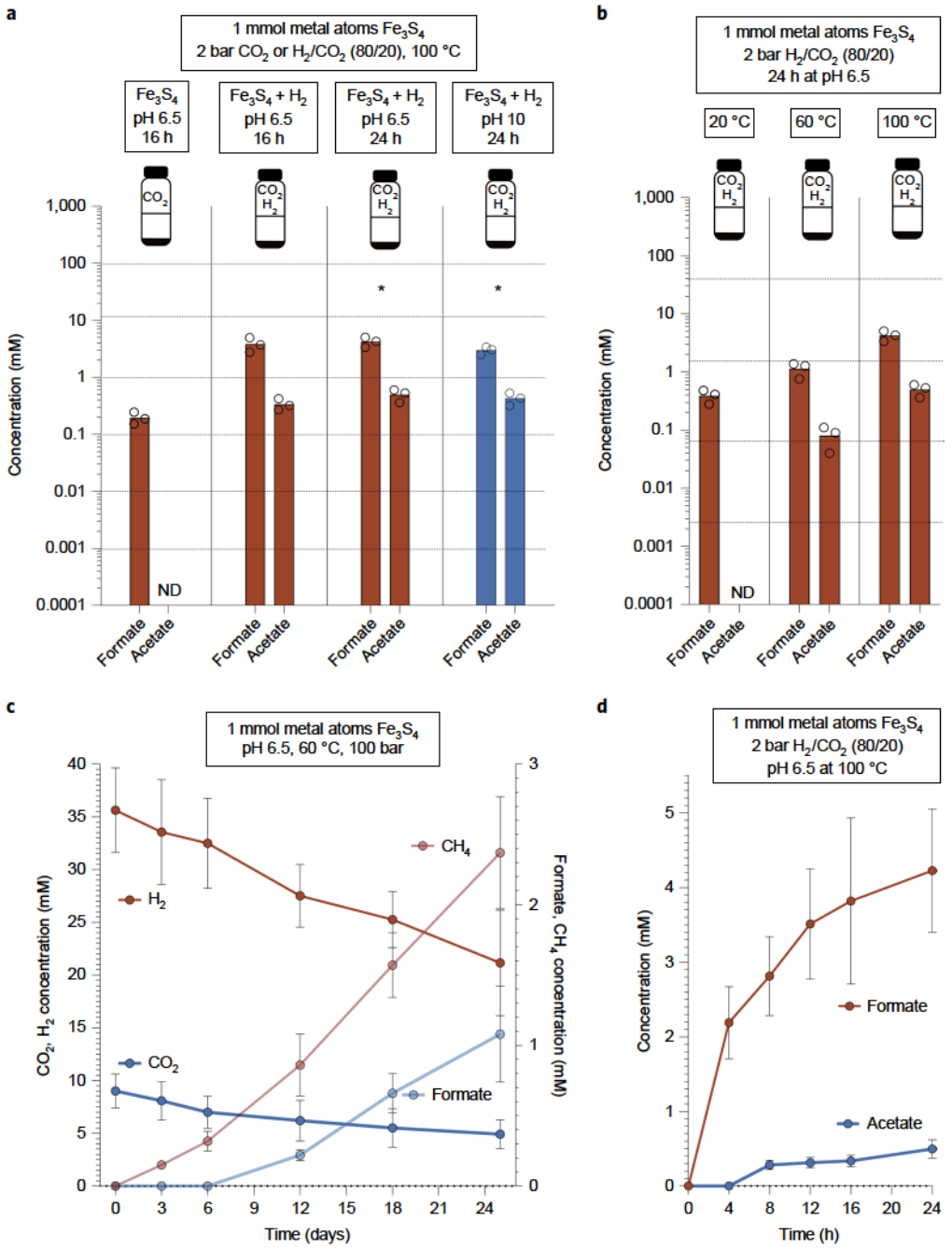


Fig. 2 | Fixation of CO_2 with H_2, catalysed by greigite. a, Catalysis by greigite at 100 °C. **b**, Effect of temperature on greigite catalysis. **c**, Time course experiment of high-pressure methane and formate production from CO_2 and H_2 under greigite catalysis (liquid phase, 150 ml) at 100 bar and 60 °C. **d**, Reaction progress over time under a 2-bar H_2/CO_2 atmosphere and 100 °C. All reactions were performed in water containing a phosphate buffer (**a**,**b**,**d**, 3 ml, **c**, 150 ml). **a**,**b**, Flasks summarize the reaction parameters, with greigite depicted in black. Amounts of catalyst are normalized by the number of moles of metal atoms per mole of mineral compound: 0.33 mmol of greigite is equivalent to 1 mmol of metal atoms in each catalyst. Individual experiments were performed under either CO_2 or H_2/CO_2 atmosphere. Red bars, pH < 7.0; blue bars, pH > 7.0. ND, not detected (no product was formed, or product concentration was below the detection limit). Circles correspond to the values of individual experiments. Values of 0 are not shown on the logarithmic scale. Asterisks indicate experiments for which Gibbs free energy was calculated in Table 1. Experimental concentration values and s.d. are listed in Supplementary Table 1, and control experiments are shown in Extended Data Fig. 2a. The influence of pH (4.0–10.0) on reactions catalysed by greigite is shown in Extended Data Fig. 4a.

conditions at 70 °C (Fig. 4a), even at lower amounts of catalyst than previously used (0.5 mmol of metal atoms), and reaches 10 µM when higher amounts are used (Fig. 4b). This suggests that pyruvate production in reactions with smaller amounts of awaruite probably occurs, but is below the detection limit of the ^{1}H-nuclear magnetic resonance (NMR) spectroscopy used here. Using even less Ni_3Fe (0.05 mmol metal atoms) is still effective for formate, acetate and methanol formation in thermal gradients from 100 to 30 °C (Extended Data Fig. 6a,c), conditions similar to those of natural alkaline hydrothermal vents[29]. Catalysts are required for the

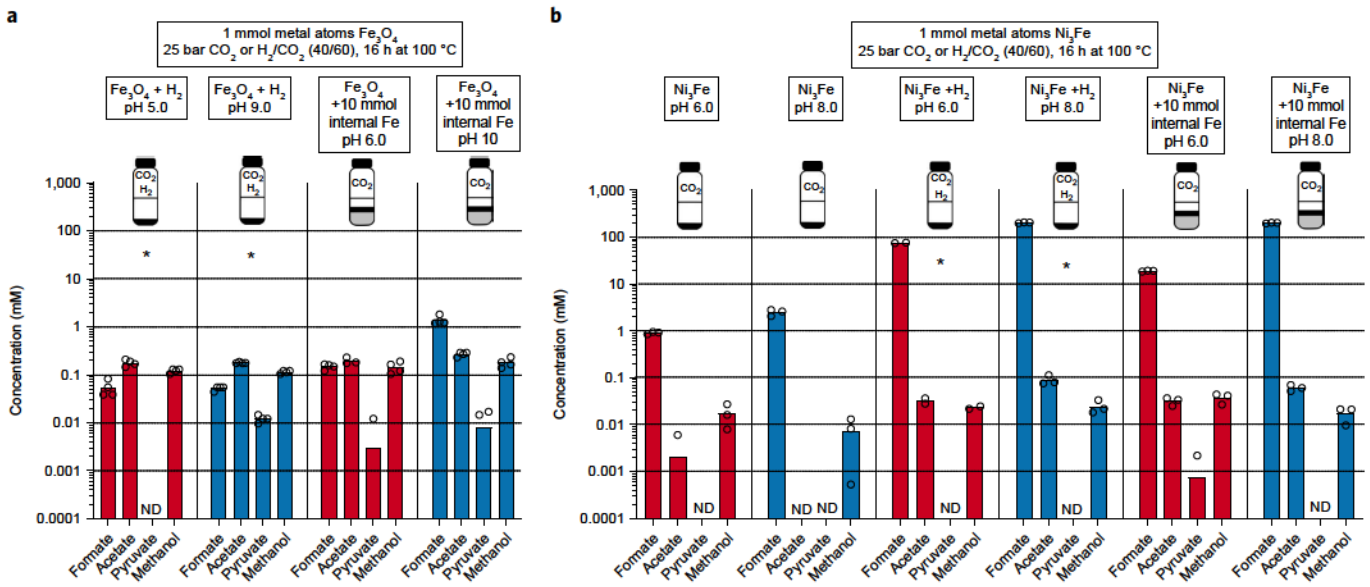


Fig. 3 | Fixation of CO_2 with H_2, catalysed by magnetite and awaruite. a,**b**, Catalysis by magnetite (**a**) and awaruite (**b**). All reactions were performed in water (1 ml). Flasks in each panel summarize the reaction parameters: hydrothermal minerals are depicted in black, additional iron powder in grey. Amounts of catalyst are normalized by the number of moles of metal atoms per mole of mineral compound: 0.33 mmol of magnetite, as well as 0.25 mmol of awaruite, are equivalent to 1 mmol of metal atoms in each catalyst. Individual experiments were performed under either CO_2 atmosphere, H_2/CO_2 atmosphere or CO_2 atmosphere with Fe powder as an electron source (also for H_2 formation from H_2O). Experiments without native Fe were performed with decontaminated stir bars; those containing native Fe were performed without stir bars due to solidification of Fe powder during the process. Red bars, pH < 7.0; blue bars, pH > 7.0. ND, not detected (no product was formed, or product concentration was below the detection limit). Experiments performed at pH < 7.0 were treated with KOH after the reaction, as in Varma et al.[18]. Circles correspond to the values of individual experiments. Values of 0 are not shown on the logarithmic scale. Asterisks indicate experiments for which Gibbs free energy was calculated in Table 1. Concentration values and s.d. of the experiments are listed in Supplementary Table 1, and control experiments are shown in Extended Data Figs. 2b,c (awaruite) and 3 (Fe^0) and Supplementary Tables 4–7. Background levels of formate at least three orders of magnitude below experimental product concentrations (awaruite); background levels of acetate (ca. 10–20 μM) were observed in controls using awaruite as catalyst. All background levels were subtracted before plotting (see Supplementary Information for all background values).

reaction—controls without catalysts yielded only trace levels of product (Extended Data Fig. 2b,c and Supplementary Tables 6 and 7).

In some experiments using Ni_3Fe, we detected ethanol in concentrations up to >100 μM (Extended Data Fig. 5b). We observed trace amounts of methane (ca. 19 ppm) in awaruite-catalysed reactions (Extended Data Fig. 8), which is substantially less than that detected in an earlier report using H_2 and CO_2 for 1–2 weeks at 500 bar and 200–400 °C with awaruite as the catalyst[34]. The hydrothermal conditions we found for the synthesis of organics from H_2 and CO_2 over 16 h with awaruite as catalyst are sufficiently mild in terms of temperature and energetics to permit microbial growth. Of the catalysts employed, only awaruite showed minor alteration after reaction, probably due to mild oxidation (Fig. 1g–i). Formate accumulation catalysed by awaruite reflects the near-equilibrium interconversion of H_2–CO_2 and formate[35].

To avoid contamination, no organic buffers were employed in any of our experiments. Because greigite is sensitive to high pH, phosphate buffer was employed here. In the experiments with magnetite and awaruite, no buffers were used. In Figs. 3 and 4, blue bars indicate reactions where the starting pH was ~11.0 through the addition of KOH to generate alkaline vent conditions; the pH measured at completion is dependent on the amount of mineral used and metal, in addition to the amount of CO_2 dissolved and organic acid synthesized. No water loss, which would potentially distort the product concentrations, was detected in any of our experiments.

Sustained synthesis of reactive organic compounds was essential at the origin of metabolism and had to be thermodynamically favourable. Equations (1–5) show the redox reactions taking place between CO_2 and H_2 to form formate (equation (1)), methanol (equation (2)), acetate (equation (3)), pyruvate (equation (4)) and methane (equation (5)).

Table 1 | ΔG for CO_2 fixation product formation (kJ mol^{-1})

	Greigite		Magnetite		Awaruite	
Product	**1**	**2**	**3**	**4**	**5**	**6**
Formate	0.31	−25.58	−19.56	−48.14	−2.56	−15.26
Methanol	ND	ND	−46.60	−46.60	−51.49	−50.33
Acetate	−71.00	−96.69	−108.59	−137.16	−120.03	−132.17
Pyruvate	ND	ND	ND	−57.18	ND	ND

Values of ΔG refer to the reactions as shown in equations (1–5). The conditions are those for reactions marked with asterisks in Figs. 2 and 3. Details of reaction conditions for columns 1–6: (1) 0.33 mmol Fe_3S_4, 100 °C, 24 h, pH 6.5, 2 bar H_2/CO_2 (80/20); (2) 0.33 mmol Fe_3S_4, 100 °C, 24 h, pH 10.0, 2 bar H_2/CO_2 (80/20); (3) 0.33 mmol, Fe_3O_4, 100 °C, 16 h, pH 6.0, 25 bar H_2/CO_2 (40/60); (4) 0.33 mmol, Fe_3O_4, 100 °C, 16 h, pH 9.0, 25 bar H_2/CO_2 (40/60); (5) 0.25 mmol, Ni_3Fe, 100 °C, 16 h, pH 6.0, 25 bar H_2/CO_2 (40/60); (6) 0.25 mmol, Ni_3Fe, 100 °C, 16 h, pH 8.0, 25 bar H_2/CO_2 (40/60). Columns 1, 3, 5: pH < 7.0; columns 2, 4, 6: pH > 8.0. ND, not detected (no product was formed or product concentration was below the detection limit). Values of ΔG for product accumulation at 100 nM in these experiments (below the detection level) are given in Supplementary Table 3.

$$H_2 + CO_2 \rightarrow HCOO^- + H^+ \quad (1)$$

$$3H_2 + CO_2 \rightarrow CH_3OH + H_2O \quad (2)$$

$$4H_2 + 2CO_2 \rightarrow CH_3COO^- + 2H_2O + H^+ \quad (3)$$

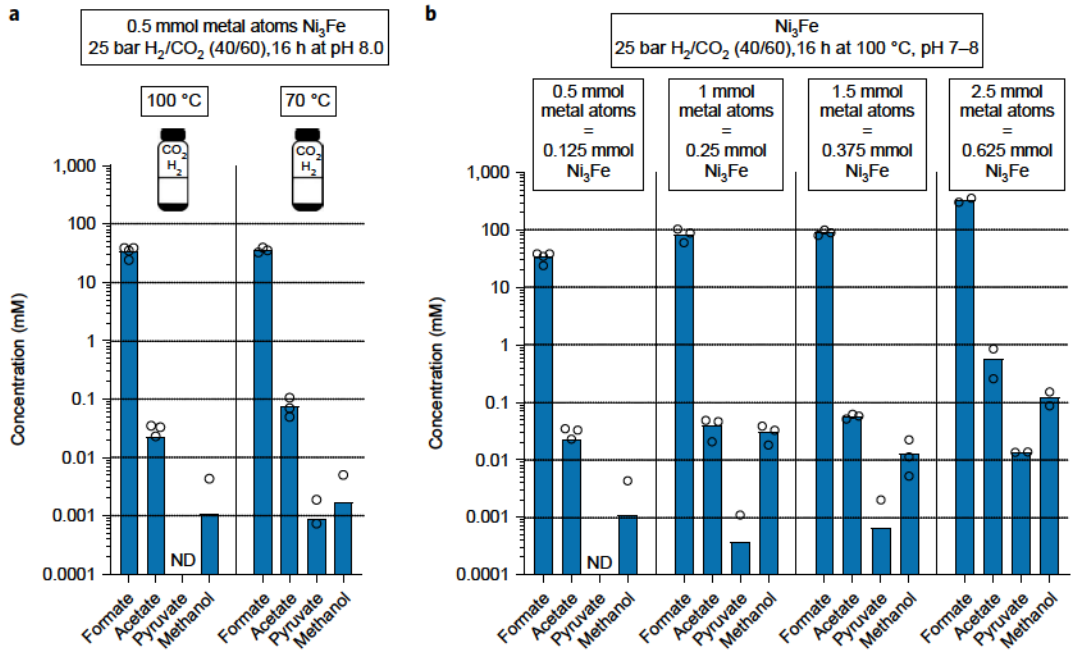


Fig. 4 | Awaruite catalyses synthesis of formate, acetate, and pyruvate. a, Effect of temperature on awaruite catalysis. **b**, Impact of amount of awaruite catalyst used. All reactions were performed in water (1 ml). Amounts of catalyst used are normalized to the number of moles of metal atoms per mole of mineral compound: 0.25 mmol of awaruite is equivalent to 1 mmol of metal atoms. Individual experiments were performed under H_2/CO_2 (40/60) atmosphere. All experiments were conducted without stir bars. Values of 0 are not shown on the logarithmic scale. All measurements were performed in at least triplicate (2.5 mmol of awaruite in duplicate). ND, not detected (no product was formed or product concentration was below the detection limit). Circles correspond to values obtained in individual experiments. Experimental concentration values and s.d. are listed in Supplementary Table 1, and control experiments are shown in Extended Data Fig. 2b,c and Supplementary Tables 6 and 7.

$$5H_2 + 3CO_2 \rightarrow CH_3(CO)COO^- + 3H_2O + H^+ \quad (4)$$

$$4H_2 + CO_2 \rightarrow CH_4 + 2H_2O \quad (5)$$

The changes in Gibbs free energy, ΔG, for six of the H_2-dependent reactions reported here are given in Table 1 (detailed datasets are shown in Supplementary Tables 2 and 3). The synthesis of observed products is close to equilibrium or exergonic. For most compounds and conditions, product generation did not reach equilibrium, indicating the kinetic inhibition of reactions. Only H_2-dependent reduction of CO_2 to formate approached equilibrium in the presence of greigite or awaruite (according to experiments in Figs. 2a and 3b). Pyruvate and CH_4 production were detected only under specific conditions despite being exergonic in nearly all treatments. In treatments with H_2 and magnetite, for example, pyruvate generation was detected only under alkaline conditions (Fig. 3a), while in treatments with H_2 and awaruite it was detected only under alkaline conditions and when the amount of mineral was increased (Fig. 4b). H_2-dependent reduction of formate to acetate (equation (3) – equation (1)); $3H_2 + CHOO^- + CO_2 \rightarrow CH_3COO^- + 2H_2O$) consistently reached similar ΔG values for each mineral, regardless of pH and mineral content (approximately −72, −89 and −115 kJ mol^{-1} at 100 °C for greigite, magnetite and awaruite, respectively), suggesting the possibility of shared features between the minerals' catalytic mechanisms. None of the three minerals catalysed acetate synthesis to completion ($\Delta G << 0$), suggesting the possible presence of kinetic barriers and an opportunity for energetic coupling. For those reactions in which no H_2 was added, only native metals were available as reductant (Extended Data Figs. 3, 5a and 7), probably generating intermediate H_2 from water.

Discussion

When greigite, magnetite or awaruite is used as a catalyst, the synthesis of formate, acetate, methanol and pyruvate from H_2 and CO_2 under hydrothermal conditions is facile. The synthesis of formate and acetate is furthermore robust to the catalyst employed. The main product we observed was formate (Figs. 2–4), which is also the main organic product of abiotic organic synthesis found in alkaline hydrothermal vent effluent[9,31,36,37]. We propose a mechanism for the catalysed two-electron reduction of CO_2 to formate for all three minerals (Extended Data Fig. 10).

Formate synthesis from H_2 and CO_2 was anticipated by earlier studies[38,39], and formate synthesis from CO_2 has been reported at high temperatures (> 250 °C) and pressures (> 300 bar) with hydrothermal minerals[40]. However, the amounts of formate we observed with Ni_3Fe at moderate temperature and pressure (70–100 °C and 25 bar H_2/CO_2 atmosphere), as well as the accumulation of acetate and pyruvate, reveal an unexpected correspondence between spontaneous H_2-dependent CO_2 reduction and metabolism. We see a clear tendency of Ni-containing compounds to preferentially produce formate in high concentrations[18], while pyruvate accumulation is preferentially observed with Fe. These product-catalyst specificities are reflected in the active site metals of corresponding enzymes of the modern acetyl-CoA pathway[11–13,41–46].

Under physiological conditions, the reducing power of H_2 is insufficient to reduce CO_2. Microbes studied to date reduce CO_2 with electrons from H_2, employing flavin-based electron bifurcation to synthesize reduced iron sulfur clusters in ferredoxin for CO_2 fixation[12,47]. This biological CO_2 fixation usually also entails ion gradients[47,48]. The reactions reported here require neither electron bifurcation nor ion gradients. With suitable inorganic catalysts that activate both H_2 and CO_2 to enable their reaction, products of the

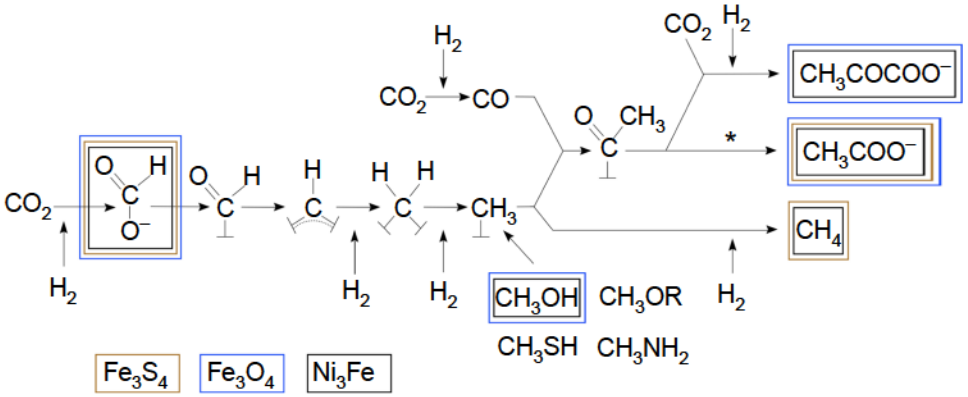


Fig. 5 | Congruence between the acetyl-CoA pathway and reactions catalysed by three iron minerals found in hydrothermal vents. The chemical reactions shown summarize the acetyl-CoA pathway as it occurs in hydrogenotrophic bacteria and archaea, as depicted in ref. [11], with the exception of free formate later discovered in the archaeal pathway[65]. The methenyl (=CH–), methylene (–CH_2–) and methyl (–CH_3) groups of the bacterial and archaeal pathways are bound to tetrahydrofolate and tetrahydromethanopterin, respectively, generically indicated here as catalysts (⊥). Colour-edged boxes indicate products observed in reactions using iron mineral catalysts. Asterisk indicates the reaction sequence in which energy is conserved as ATP via substrate-level phosphorylation in the biological pathway (acyl-nickel, thioester and acyl-phosphate intermediates employed by the enzymatic pathway for stepwise conservation of free energy in exergonic conversion of the nickel-bound acyl group to ATP[11] are not shown). All products shown were observed at temperatures ≤ 100 °C and obtained within < 24 h, except methane in the case of greigite, which was observed over the course of 25 d (Fig. 2c). Methanol, methyl sulfide, methyl amines and methoxy groups from coal can serve as methyl donors for the pathway[11,66].

acetyl-CoA pathway (Figs. 2, 3 and 4) are formed without the addition of organic co-factors.

With the exception of ethanol, the reaction products we observe correspond exactly to those of the biological acetyl-CoA pathway to pyruvate[11] (Fig. 5). No other reaction products were observed. That is, the mineral-catalysed H_2-dependent reduction of CO_2 delivers a very discrete subset of possible chemical structures: one that constructs the backbone of carbon and energy metabolism in primitive anaerobic autotrophs[11–15]. The acetyl-CoA pathway[11,14] entails 11 main enzymes totalling ~15,000 amino acid residues[13,41–45] plus six organic co-factors, each with its own complex biosynthesis[14]. The bacterial and archaeal versions of the pathway involve evolutionarily unrelated enzymes but chemically similar methyl synthesis routes[6,11,14]. The reactions of the acetyl-CoA pathway employed in modern metabolism (Fig. 5) involve the stepwise conservation of chemical energy during CO_2 fixation as acetyl-nickel, acetyl-thioester, acetyl-phosphate and ATP synthesis via substrate-level phosphorylation (marked with an asterisk in Fig. 5)[11–14]. Although the nature of catalyst-bound intermediates of the biological pathway from H_2 and CO_2 to methane, acetate and pyruvate is known[11–14], the identity of catalyst-bound intermediates of mineral-catalysed reactions is not.

Proposals for the nature of primordial CO_2 fixation and energy conservation at biochemical origins typically posit the participation of external energy sources[49], such as UV light[50], heat, impact, pressure, electrical currents or ion gradients[28], to push organic synthesis forward. The reactions reported here require no additional energy source for the unfolding of a protometabolic acetyl-CoA pathway from H_2 and CO_2 other than the natural reactivity of two gases and metal catalysts, indicating that neither membranes, though essential for the emergence of free-living cells[6,51–53], nor external potentials[19,54] were required for primordial CO_2 fixation along an exergonic, H_2-dependent, non-enzymatic pathway to C3 products. The energy for the synthesis of compounds capable of phosphorylating ADP via substrate-level phosphorylation[6,11,12]—for reactions reported here, and for those of the enzymatically catalysed acetyl-CoA pathway—stems from the exergonic synthesis of biologically relevant organic compounds from H_2 and CO_2. Our findings suggest that abiotic, geochemical versions of energy-releasing reactions underlying the acetyl-CoA pathway very likely preceded the enzymes that catalyse it today[11,14,18,55]. The simplicity and primordial nature of these reactions furthermore suggests that metabolism elsewhere could initiate by a similar route.

Methods

General information. An overview of the experiments performed can be found in Extended Data Fig. 1, and relevant controls in Extended Data Figs. 2 and 3 and Supplementary Tables 4–7. The quantity of each transition metal reagent tested as a carbon fixation catalyst was normalized so that it would contain the same number of millimoles of metal atoms across the experiments. For example, 1 mmol of metal atoms corresponds to 0.33 mmol of greigite (99 mg), 0.33 mmol of magnetite (77 mg) and 0.25 mmol of awaruite (58 mg). Each reaction was performed in at least triplicate. Information on suppliers, grade and purity of all reagents used is listed in the Supplementary Information.

Synthesis of greigite. Every piece of apparatus used in greigite synthesis was stored in an anaerobic chamber (Coy Laboratory Products) under a gas mixture of N_2/H_2/CO_2 (80/5/15) for at least 48 h before use, to remove residual oxygen. Reagents for greigite synthesis were purged with N_2 before use unless otherwise stated. Amorphous FeO(OH) was synthesized as reported previously[56] and suspended in Milli-Q water (0.30 mol l^{-1}) under air atmosphere. After purging with N_2, this suspension was stored in a glass bottle under N_2/H_2/CO_2 (80/5/15). Solutions of Na_2S (1.0 M) and H_2SO_4 (2.0 M) were prepared as reported previously[57] and stored in a glass bottle under N_2. Greigite was synthesized in a solid–gas reaction system as reported previously[57] with slight modifications. In brief, amorphous FeO(OH) (0.66 mmol, 2.2 ml of water suspension) was aliquoted to a glass reaction vessel, and a test tube containing 1.0 ml of the Na_2S solution was placed in the vessel inside the anaerobic chamber. The vessel was sealed with an ethylene tetrafluoroethylene (ETFE)-coated butyl rubber stopper and an aluminium seal. The vessel was then removed from the anaerobic chamber and the headspace gas replaced with Ar. After returning the vessel to the anaerobic chamber, H_2S gas was generated inside the vessel by injecting 0.5 ml of the H_2SO_4 solution into the Na_2S solution in the test tube using a disposable Myjector syringe (Terumo). The vessel was incubated at 80 °C for 3 h. The resulting greigite suspension was collected by pipetting from several reaction vessels, washed with 0.5 M HCl and then rinsed with N_2-purged Milli-Q water in the anaerobic chamber as described previously[57].

CO_2 fixation catalysed by greigite. Synthesized greigite (0.33 mmol) was resuspended in 3 ml of potassium phosphate buffer (20 mM) of designated pH. The greigite suspension was placed in a fresh glass reaction vessel, which was then sealed with an ETFE-coated butyl rubber stopper and an aluminium seal. The vessel was then removed from the chamber and the headspace gas replaced with H_2/CO_2 (80/20) or CO_2 outside the chamber. Vessels were incubated at 100 °C over 4–24 h.

High-performance liquid chromatography (HPLC) analysis (greigite experiments). Liquid-phase components were analysed on a D-2000 LaChrom

Elite HPLC system (Hitachi), equipped with an Aminex HPX-87H column (300 mm, 7.8 mm internal diameter (i.d.); Bio-Rad Laboratories) and an L-2400 UV detector at 240 nm and L-2490 RI detector as described previously[58]. Supernatants obtained in the CO_2 reduction experiments were collected after centrifugation inside the anaerobic chamber, then 10 µl of the supernatants obtained were directly injected into the HPLC circuit and chromatographed under an isocratic flow of 0.7 ml min^{-1} (eluent, 10 mM H_2SO_4 in H_2O). Column temperature was maintained at 50 °C. Identities of the analytes detected were determined by the liquid chromatography–tandem mass spectrometry system: Agilent 1200 HPLC (Agilent Technologies) coupled to an HCT Ultra mass spectrometer (Bruker Daltonics) using a Shodex HILICpak VG-50 2D column (150 mm, 2 mm i.d.; Showa Denko). The supernatant prepared as above was mixed with an equal amount of the eluent, then 5 µl of the mixture was injected into the HPLC circuit and chromatographed under an isocratic flow of 0.1 ml min^{-1} (eluent, a mixture of acetonitrile and 0.25% ammonia water at 80/20). Column temperature was maintained at 30 °C.

High-pressure measurements (greigite experiments). A previously developed[59] high-pressure incubation system was used for the high-pressure CO_2 (Fig. 2c) and CO (Extended Data Fig. 4b) reduction reactions in this study. The system consisted of an incubation vessel (stainless steel with Sulfinert coating on its internal wall, volume 150 cm^3; Swagelok), inflow/outflow tubes with valves (Swagelok) and a 500D automated syringe pump (Teledyne Isco). Greigite suspension was placed in the reaction vessel inside the anaerobic chamber. After sealing the vessel with inflow and outflow tubes, the headspace gas was replaced with $H_2 + CO_2$ (80/20) through a rubber septum equipped with an inflow tube, via a needle. This vessel was then connected to the syringe pump via the inflow tube to complete the incubation system. Potassium phosphate buffer was injected by the syringe pump to reach a hydrostatic pressure of 100 bar. Incubation at 60 °C started after H_2 and CO_2 were completely dissolved in the liquid phase (verified by gas chromatography analysis). Samples were periodically collected via the outflow tube while keeping the same hydrostatic pressure through automated pressure control of the syringe pump.

Gas analysis (greigite experiments). Gas-phase measurements were carried out on a gas chromatograph GC-2014 (Shimadzu) as described previously[58]. Depending on the target gas component, different columns and detectors were used: an Rt-QPLOT (30 m, 0.32 mm i.d., 10 µm F.T.; Restek) with flame ionization detector (FID) for CH_4, molecular sieve 13× column (2 m, 3 mm i.d.; Shimadzu) with a thermal conductivity detector (TCD) for H_2 and CO, and activated charcoal column (2.0 m, 3 mm, 60/80 mesh; Shinwa Chemical Industries) with TCD for CO_2. Pure He and Ar were used as carrier gases for FID and TCD, respectively. Gases were identified by gas chromatography–mass spectrometry (GC–MS) using two systems: (1) TQ8040 NX GC–MS (Shimadzu) equipped with a polar capillary column (TC-70, 30 m, 0.25 mm i.d., 0.25 µm F.T.; GL Sciences); (2) QP2010 Plus GC–MS (Shimadzu) equipped with Rt-Q-BOND (15 m, 0.32 mm i.d., 10 µm F.T.; Restek). Carrier gas in both systems was pure He.

Synthesis of awaruite (Ni_3Fe) nanoparticles. As previously reported[60,61], spent tea leaves can be used as sustainable hard template to synthesize native metal nanoparticles at the desired composition. For the synthesis of nanoparticular Ni_3Fe, washed and dried tea leaves were added to an aqueous solution of $Ni(NO_3)_2{\cdot}6H_2O$ and $Fe(NO_3)_3{\cdot}9H_2O$ (molar ratio 3/1) and stirred at room temperature for 2 h. The mass ratio of tea leaves and metal precursors was set at 2/1. Due to the low decomposition temperature of the metal nitrate salt (< 200 °C), metal oxide nanoparticles can be formed in the pore confinement of the template before its structural damage/combustion. The carbon-based tea leaf template was burned out in air atmosphere (550 °C for 4 h), and the resulting Ni_3Fe oxide was washed with 0.1 M HCl solution for 2 h and cleaned with deionized water. Finally, the product was treated in a reductive 10% H_2/Ar flow (100 ml min^{-1}) at 500 °C for 2 h to generate the intermetallic Ni_3Fe compound.

CO_2 fixation catalysed by magnetite and awaruite. Awaruite and magnetite powders (commercial) were placed in a 1.5-ml glass vial. In the case of the magnetite and awaruite experiments shown in Fig. 3, a clean polytetrafluoroethylene (PTFE)-coated stir bar was added to the vial. All further awaruite experiments were conducted without stir bars. The reaction vials were then filled with 1.0 ml of Milli-Q water. Whenever the effect of increased pH of the reaction mixtures was tested, solid KOH was added to Milli-Q water before the reaction (45 mg ml^{-1}). KOH had previously been tested for contaminants by ^{1}H-NMR analysis (Extended Data Fig. 9a). To prevent cross-contamination while allowing for ready access of the gas to the reaction mixture, the vials were closed by caps with punctured PTFE septa. The reaction vials (3–12) were placed in a stainless-steel pressure reactor (Berghof or Parr) which was then sealed, flushed three times with ca. 5 bar CO_2, pressurized to a final value of 25 bar CO_2 (unless noted otherwise) and heated at the desired temperature (an external heating mantle was used) for 16 h. At a reaction temperature of 100 °C, a maximum pressure of ca. 30 bar was reached. After the reaction, the reactor was allowed to cool to room temperature (3–4 h from 100 °C, 2–3 h from 70 °C) before sample analysis[18,55].

Experiments with iron powder or hydrogen gas. These experiments were performed according to the general procedure described above, except that 10 mmol (560 mg) of Fe^0 powder was first placed in the reaction vials followed by the mineral tested and no stir bars were added. Further experiments exploring the impact of the amount of Fe^0 powder are shown in Extended Data Fig. 7a. Whenever H_2 was used in the experiments, the pressure reactor was first flushed with CO_2, pressurized with 10 bar of H_2 and then brought to 25 bar by further addition of CO_2 (H_2/CO_2 approximately 40/60).

Work-up procedure for reaction mixtures (magnetite and awaruite). The pH of individual reaction mixtures was determined via TRITEST L pH 1.0–11.0 papers (Macherey-Nagel) directly after the reaction. The values of Ni_3Fe experiments were confirmed with a pH metre (Lab 875, SI Analytics) and a pH combination micro-electrode (A 157 IDS, SI Analytics). CO_2 dissolved in the reaction mixture during the reaction decreased reaction pH values due to the formation of carbonic acid. Reaction mixtures that did not contain KOH were either treated with ca. 45 mg of solid KOH per 1-ml reaction mixture to precipitate the metal ions as hydroxides (in the case of Fe_3O_4 and Ni_3Fe experiments, shown in Fig. 3), or left untreated (in the case of Ni_3Fe). The treatment of individual experimental rows was also dependent on the visible concentration of metal ions in solution (since these ions have to be removed by precipitation as hydroxides before NMR measurements), and is additionally described in the corresponding figure legends. All samples were then centrifuged at 13,000 r.p.m. for 10 min. The supernatant was then separated from the precipitate (catalyst) and stored at 4 °C overnight or longer before NMR or HPLC analysis.

NMR analysis (magnetite and awaruite). Concentrations of formate, acetate, pyruvate and methanol (as methoxide) were determined by ^{1}H-NMR, following the protocol established in Varma et al.[18]. The supernatant of the centrifuged samples was mixed with sodium 3-(trimethylsilyl)-1-propanesulfonate D_2O-solution as the internal standard (CH_3 peak at 0 ppm). NMR spectra were acquired on either a Bruker Avance III 600 or 300 spectrometer at 297 K, using a ZGESGP pulse programme. Thirty-two scans were acquired for each sample and the relaxation delay was set to 40 s (600 MHz) and 87 s (300 MHz), with a spectral width of either 12,315 ppm (600 MHz) or 11,963 ppm (300 MHz). Analysis and integration were performed using MestReNova (v.10.0.2) software. Shifts of the measured products are depicted in Extended Data Fig. 9b.

Powder X-ray diffraction. Power XRD analysis was performed for pre- and post-reaction catalysts. For greigite, XRD specimes were prepared as described previously[57]. In brief, sampls were collected by centrifugation and the pellet obtained directly mounted in slurry form on a silicon holder (SanyuShoko), then sealed using polyimide film (Nilaco Corporation) and vacuum grease (JEOL) to avoid possible desiccation and oxidation during analysis. Specimens were analysed using a RINT2000 X-ray diffractometer (Rigaku) at room temperature for $CuK\alpha_{1,2}$ radiation scanning at a step interval of 0.02° 2θ and a counting time of 2 s with a 2θ range of 20–60°, operating at an accelerating voltage of 40 kV at 30 mA. To prepare specimens for magnetite and awaruite experiments, samples were collected, washed with Milli-Q water and dried under vacuum. XRD patterns of these specimens were collected at room temperature using a theta-theta diffractometer (Stoe) in BraggBrentano geometry for $CuK\alpha_{1,2}$ radiation scanning, at a step interval of 0.04° 2θ and a counting time of 6 s with a 2θ range of 20–80°.

Electron microscopy. Electron microscopic observation was conducted for pre-reaction catalysts to check their morphology. For greigite, specimens were prepared as described previously[57]. Briefly, in the anaerobic chamber, greigite was rinsed at least three times with N_2-purged Milli-Q water, dried at room temperature and then mounted on an aluminium stab using carbon tape. Specimens were removed from the anaerobic chamber, coated with platinum/palladium alloy with an ion-sputter E102 (Hitachi) and observed on either a JSM-6330F (JEOL) or JSM-7800F (JEOL) field-emission scanning electron microscope at an acceleration voltage of 5 kV. Magnetite samples were deposited on lacy carbon film-coated Cu grids (400 mesh) and observed on an H-7100 (Hitachi) transmission electron microscope at an acceleration voltage of 100 kV. Awaruite samples were collected and embedded in Spurr resin (hard mixture). Resin blocks thus obtained were trimmed using an EM TRIM milling system (Leica). Thin sections were cut from the resin blocks by microtome with a 35° diamond knife (Reichert Ultra-Cut), dispersed in Milli-Q water, transferred from the water surface on lacy carbon film-coated Cu grids (400 mesh) and observed on an S-5500 (Hitachi) scanning transmission electron microscope at an acceleration voltage of 30 kV.

Thermodynamic calculations. For ΔG calculations, published values of ΔH (reaction enthalpy) and ΔG values were used[62,63]. The effect of temperature on Gibbs free energy yield was calculated using the Gibbs–Helmholtz equation. Equilibrium constants at different temperatures were adjusted using the van't Hoff equation (detailed equations given in Supplementary Information). Corrections based on non-standard pressures were estimated using partial molar volume changes of the reactions[64]. For any organic compounds not detected, an aqueous concentration of 0.1 µM was assumed. For CH_4, a partial pressure of 10^{-7} bar was

assumed when not detected. In reactions containing Fe^0 as an electron donor (Supplementary Table 2), H_2 concentration was estimated by assuming that H_2-dependent CO_2 reduction to formate reached equilibrium. Final H_2 and CO_2 concentrations were estimated based on the measured products (subtracting 1 mol of H_2 per mole formate detected).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All data are available in the main text, Extended Data Figs. 1–10 and the Supplementary Information (Supplementary Methods, Supplementary Tables 1–7, Supplementary Figs. 1–29 and Supplementary Equations).

Received: 9 July 2019; Accepted: 23 January 2020;

References

1. Baross, J. A. & Hoffman, S. E. Submarine hydrothermal vents and associated gradient environments as sites for the origin and evolution of life. *Orig. Life Evol. Biosph.* **15**, 327–345 (1985).
2. McCollom, T. M. Abiotic methane formation during experimental serpentinization of olivine. *Proc. Natl Acad. Sci. USA* **113**, 13965–13970 (2016).
3. McDermott, J. M., Seewald, J. S., German, C. R. & Sylva, S. P. Pathways for abiotic organic synthesis at submarine hydrothermal fields. *Proc. Natl Acad. Sci. USA* **112**, 7668–7672 (2015).
4. Ménez, B. et al. Abiotic synthesis of amino acids in the recesses of the oceanic lithosphere. *Nature* **564**, 59–63 (2018).
5. Klein, F. & Bach, W. Fe-Ni-Co-O-S phase relations in peridotite–seawater interactions. *J. Petrol.* **50**, 37–59 (2009).
6. Martin, W. F. & Russell, M. J. On the origin of biochemistry at an alkaline hydrothermal vent. *Philos. Trans. R. Soc. B* **362**, 1887–1925 (2007).
7. Preiner, M. et al. Serpentinization: connecting geochemistry, ancient metabolism and industrial hydrogenation. *Life* **8**, 41 (2018).
8. Sleep, N. H., Bird, D. K. & Pope, E. C. Serpentinite and the dawn of life. *Philos. Trans. R. Soc. B* **366**, 2857–2869 (2011).
9. Schrenk, M. O., Brazelton, W. J. & Lang, S. Q. Serpentinization, carbon, and deep life. *Rev. Miner. Geochem.* **75**, 575–606 (2013).
10. Arndt, N. T. & Nisbet, E. G. Processes on the young Earth and the habitats of early life. *Annu. Rev. Earth Planet. Sci.* **40**, 521–549 (2012).
11. Fuchs, G. Alternative pathways of carbon dioxide fixation: insights into the early evolution of life? *Annu. Rev. Microbiol.* **65**, 631–658 (2011).
12. Müller, V., Chowdhury, N. P. & Basen, M. Electron bifurcation: a long-hidden energy-coupling mechanism. *Annu. Rev. Microbiol.* **72**, 331–353 (2018).
13. Ragsdale, S. W. & Pierce, E. Acetogenesis and the Wood–Ljungdahl pathway of CO_2 fixation. *Biochim. Biophys. Acta* **1784**, 1873–1898 (2008).
14. Sousa, F. L. & Martin, W. F. Biochemical fossils of the ancient transition from geoenergetics to bioenergetics in prokaryotic one carbon compound metabolism. *Biochim. Biophys. Acta* **1837**, 964–981 (2014).
15. Weiss, M. C. et al. The physiology and habitat of the last universal common ancestor. *Nat. Microbiol.* **1**, 16116 (2016).
16. Huber, C. & Wächtershäuser, G. Activated acetic acid by carbon fixation on (Fe,Ni)S under primordial conditions. *Science* **276**, 245–248 (1997).
17. He, C., Tian, G., Liu, Z. & Feng, S. A mild hydrothermal route to fix carbon dioxide to simple carboxylic acids. *Org. Lett.* **12**, 649–651 (2010).
18. Varma, S. J., Muchowska, K. B., Chatelain, P. & Moran, J. Native iron reduces CO_2 to intermediates and endproducts of the acetyl-CoA pathway. *Nat. Ecol. Evol.* **2**, 1019–1024 (2018).
19. Roldan, A. et al. Bio-inspired CO_2 conversion by iron sulfide catalysts under sustainable conditions. *Chem. Commun.* **51**, 7501–7504 (2015).
20. Rajendran, S. & Nasir, S. Hydrothermal altered serpentinized zone and a study of Ni-magnesioferrite–magnetite–awaruite occurrences in Wadi Hibi, Northern Oman Mountain: discrimination through ASTER mapping. *Ore Geol. Rev.* **62**, 211–226 (2014).
21. 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. London* **154**, 377–402 (1997).
22. Rickard, D. & Luther, G. W. Chemistry of iron sulfides. *Chem. Rev.* **107**, 514–562 (2007).
23. McCollom, T. M. & Seewald, J. S. Serpentinites, hydrogen, and life. *Elements* **9**, 129–134 (2013).
24. Hunger, S. & Benning, L. G. Greigite: a true intermediate on the polysulfide pathway to pyrite. *Geochem. Trans.* **8**, 1 (2007).
25. Findlay, A. J. et al. Iron and sulfide nanoparticle formation and transport in nascent hydrothermal vent plumes. *Nat. Commun.* **10**, 1597 (2019).
26. Schmitt-Kopplin, P. et al. High molecular diversity of extraterrestrial organic matter in Murchison meteorite revealed 40 years after its fall. *Proc. Natl Acad. Sci. USA* **107**, 2763–2768 (2010).
27. Dayhoff, M. O. & Eck, R. V. Evolution of the structure of ferredoxin based on surviving relics of primitive amino acid sequences. *Science* **152**, 363–366 (1966).
28. White, L. M., Bhartia, R., Stucky, G. D., Kanik, I. & Russell, M. J. Mackinawite and greigite in ancient alkaline hydrothermal chimneys: identifying potential key catalysts for emergent life. *Earth Planet. Sci. Lett.* **430**, 105–114 (2015).
29. Kelley, D. S. et al. An off-axis hydrothermal vent field near the Mid-Atlantic Ridge at 30° N. *Nature* **412**, 145–149 (2001).
30. Kelley, D. S., Baross, J. A. & Delaney, J. R. Volcanoes, fluids, and life at mid-ocean ridge spreading centers. *Annu. Rev. Earth Planet. Sci.* **30**, 385–491 (2002).
31. Lang, S. Q., Butterfield, D. A., Schulte, M., Kelley, D. S. & Lilley, M. D. Elevated concentrations of formate, acetate and dissolved organic carbon found at the Lost City hydrothermal field. *Geochim. Cosmochim. Acta* **74**, 941–952 (2010).
32. Lang, S. Q. et al. Deeply-sourced formate fuels sulfate reducers but not methanogens at Lost City hydrothermal field. *Sci. Rep.* **8**, 755 (2018).
33. Etiope, G. & Sherwood Lollar, B. Abiotic methane on Earth. *Rev. Geophys.* **51**, 276–299 (2013).
34. Horita, J. & Berndt, M. E. Abiogenic methane formation and isotopic fractionation under hydrothermal conditions. *Sci. Rep.* **285**, 1055–1057 (1999).
35. Schuchmann, K. & Müller, V. Direct and reversible hydrogenation of CO_2 to formate by a bacterial carbon dioxide reductase. *Science* **342**, 1382–1385 (2013).
36. Eickenbusch, P. et al. Origin of short-chain organic acids in serpentinite mud volcanoes of the Mariana convergent margin. *Front. Microbiol.* **10**, 1729 (2019).
37. Etiope, G. & Schoell, M. Abiotic gas: atypical, but not rare. *Elements.* **10**, 291–296 (2014).
38. McCollom, T. M. & Seewald, J. S. Experimental constraints on the hydrothermal reactivity of organic acids and acid anions: I. Formic acid and formate. *Geochim. Cosmochim. Acta* **67**, 3625–3644 (2003).
39. McCollom, T. M. & Seewald, J. S. Carbon isotope composition of organic compounds produced by abiotic synthesis under hydrothermal conditions. *Earth Planet. Sci. Lett.* **243**, 74–84 (2006).
40. McCollom, T. M. & Seewald, J. S. A reassessment of the potential for reduction of dissolved CO_2 to hydrocarbons during serpentinization of olivine. *Geochim. Cosmochim. Acta* **65**, 3769–3778 (2001).
41. Menon, S. & Ragsdale, S. W. Unleashing hydrogenase activity in carbon monoxide dehydrogenase/acetyl-CoA synthase and pyruvate:ferredoxin oxidoreductase. *Biochemistry* **35**, 15814–15821 (1996).
42. Jeoung, J.-H. & Dobbek, H. Carbon dioxide activation at the Ni,Fe-cluster of anaerobic carbon monoxide dehydrogenase. *Conserv. Exhib.* **318**, 1461–1464 (2007).
43. Dobbek, H., Svetlitchnyi, V., Gremer, L., Huber, R. & Meyer, O. Crystal structure of a carbon monoxide dehydrogenase reveals a [Ni-4Fe-5S] cluster. *Science* **293**, 1281–1285 (2001).
44. Chabrière, E. et al. Crystal structures of the key anaerobic enzyme pyruvate ferredoxin oxidoreductase free and in complex with pyruvate. *Nat. Struct. Biol.* **6**, 182–190 (1999).
45. Volbeda, A. et al. Crystal structure of the nickel-iron hydrogenase from *Desulfovibrio gigas*. *Nature* **373**, 580–587 (1995).
46. Martin, W. F. Carbon-metal bonds: rare and primordial in metabolism. *Trends Biochem. Sci.* **44**, 807–818 (2019).
47. Buckel, W. & Thauer, R. K. Flavin-based electron bifurcation, ferredoxin, flavodoxin, and anaerobic respiration with protons (Ech) or NAD+(Rnf) as electron acceptors: a historical review. *Front. Microbiol.* **9**, 401 (2018).
48. Vasiliadou, R., Dimov, N., Szita, N., Jordan, S. & Lane, N. Possible mechanisms of CO_2 reduction by H_2 via prebiotic vectorial electrochemistry. *Interface Focus* **9**, 20190073 (2018).
49. Kaufmann, M. On the free energy that drove primordial anabolism. *Int. J. Mol. Sci.* **10**, 1853–1871 (2009).
50. Patel, B. H., Percivalle, C., Ritson, D. J., Duffy, C. D. & Sutherland, J. D. Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism. *Nat. Chem.* **7**, 301–307 (2015).
51. Lane, N. & Martin, W. F. The origin of membrane bioenergetics. *Cell* **151**, 1406–1416 (2012).
52. Jordan, S. F., Nee, E. & Lane, N. Isoprenoids enhance the stability of fatty acid membranes at the emergence of life potentially leading to an early lipid divide. *Interface Focus* **9**, 20100067 (2019).
53. Jordan, S. F. et al. Promotion of protocell self-assembly from mixed amphiphiles at the origin of life. *Nat. Ecol. Evol.* **3**, 1705–1714 (2019).
54. Kitadai, N. et al. Metals likely promoted protometabolism in early ocean alkaline hydrothermal systems. *Sci. Adv.* **5**, eaav7848 (2019).
55. Muchowska, K. B. et al. Metals promote sequences of the reverse Krebs cycle. *Nat. Ecol. Evol.* **1**, 1716–1721 (2017).
56. Lovley, D. R. & Phillips, E. J. Organic matter mineralization with reduction of ferric iron in anaerobic sediments. *Appl. Environ. Microbiol.* **51**, 683–689 (1986).

57. Igarashi, K., Yamamura, Y. & Kuwabara, T. Natural synthesis of bioactive greigite by solid–gas reactions. *Geochim. Cosmochim. Acta* **191**, 47–57 (2016).
58. Kato, S., Yumoto, I. & Kamagata, Y. Isolation of acetogenic bacteria that induce biocorrosion by utilizing metallic iron as the sole electron donor. *Appl. Environ. Microbiol.* **81**, 67–73 (2015).
59. Mayumi, D. et al. Carbon dioxide concentration dictates alternative methanogenic pathways in oil reservoirs. *Nat. Commun.* **4**, 1998 (2013).
60. Deng, X., Chan, C. K. & Tüysüz, H. Spent tea leaf templating of cobalt-based mixed oxide nanocrystals for water oxidation. *ACS Appl. Mater. Interfaces* **8**, 32488–32495 (2016).
61. Yu, M., Moon, G., Bill, E. & Tüysüz, H. Optimizing Ni–Fe oxide electrocatalysts for oxygen evolution reaction by using hard templating as a toolbox. *ACS Appl. Energy Mater.* **2**, 1199–1209 (2019).
62. Hanselmann, K. W. Microbial energetics applied to waste repositories. *Experientia* **47**, 645–687 (1991).
63. Amend, J. P. & Shock, E. L. Energetics of overall metabolic reactions of thermophilic and hyperthermophilic archaea and bacteria. *FEMS Microbiol. Rev.* **25**, 175–243 (2001).
64. Wang, G., Spivack, A. J. & Hondt, S. D. Gibbs energies of reaction and microbial mutualism in anaerobic deep subseafloor sediments of ODP Site 1226. *Geochim. Cosmochim. Acta* **74**, 3938–3947 (2010).
65. Wagner, T., Ermler, U. & Shima, S. The methanogenic CO_2 reducing-and-fixing enzyme is bifunctional and contains 46 [4Fe-4S] clusters. *Science* **354**, 114–117 (2015).
66. Mayumi, D. et al. Methane production from coal by a single methanogen. *Science* **354**, 222–225 (2016).
67. Hoffman, B. M. et al. Mechanism of nitrogen fixation by nitrogenase: the next stage. *Chem. Rev.* **114**, 4041–4062 (2014).

Acknowledgements

We thank Y. Dai for setting up gas analysis for the awaruite experiments, A. do Nascimento Vieira for performing parts of the revision experiments, A. Bähr and P. Lim for scientific support and J. C. Xavier for discussions. For funding, J.M., W.F.M. and H.T. thank the VW foundation (no. 96_742). W.F.M. and H.T. thank Deutsche Forschungsgemeinschaft (no. MA-1426/21-1/TU 315/8-1) and W.F.M. thanks the European Research Council (no. ERC 666053). This work is partly supported by IMPRS-RECHARGE and MAXNET Energy consortium of the Max Planck Society. K.I. and Y.K. thank JSPS KAKENHI Grant-in-Aid for Scientific Research on Innovative Areas (K.I., no. JP17H05240; Y.K., no. 26106004). K.I. is also supported by Grant-in-Aid for Young Scientists B (no. JP17K15255). J.M. thanks the European Research Council (no. ERC 639170) and ANR LabEX (no. ANR-10-LABX-0026 CSC). This work was also partly supported by Nanotechnology Platform Program (Molecule and Material Synthesis) of the Ministry of Education, Culture, Sports, Science and Technology, Japan.

Author contributions

W.F.M. wrote the initial draft of the main text and all authors edited the manuscript. W.F.M., H.T., J.M. and M.P. designed the awaruite experiments. M.P. performed the awaruite experiments and assembled the results for the main text and Supplementary Information material. K.B.M. designed and performed the magnetite experiments. S.J.V. performed exploratory experiments with magnetite. Design of the greigite experiments was done by K.I. and Y.K. K.I. performed the experiments. H.T. and M.Y. designed and synthesized the awaruite nanoparticles and performed XRD and transmission electron microscopy measurements for the magnetite and awaruite experiments. M.K.N. performed and interpreted the thermodynamics calculations. K.K., J.M., H.T. and M.P. formulated the H_2 reduction mechanism shown in the Supplementary Information.

Competing interests

The authors declare no competing interests.

Additional information

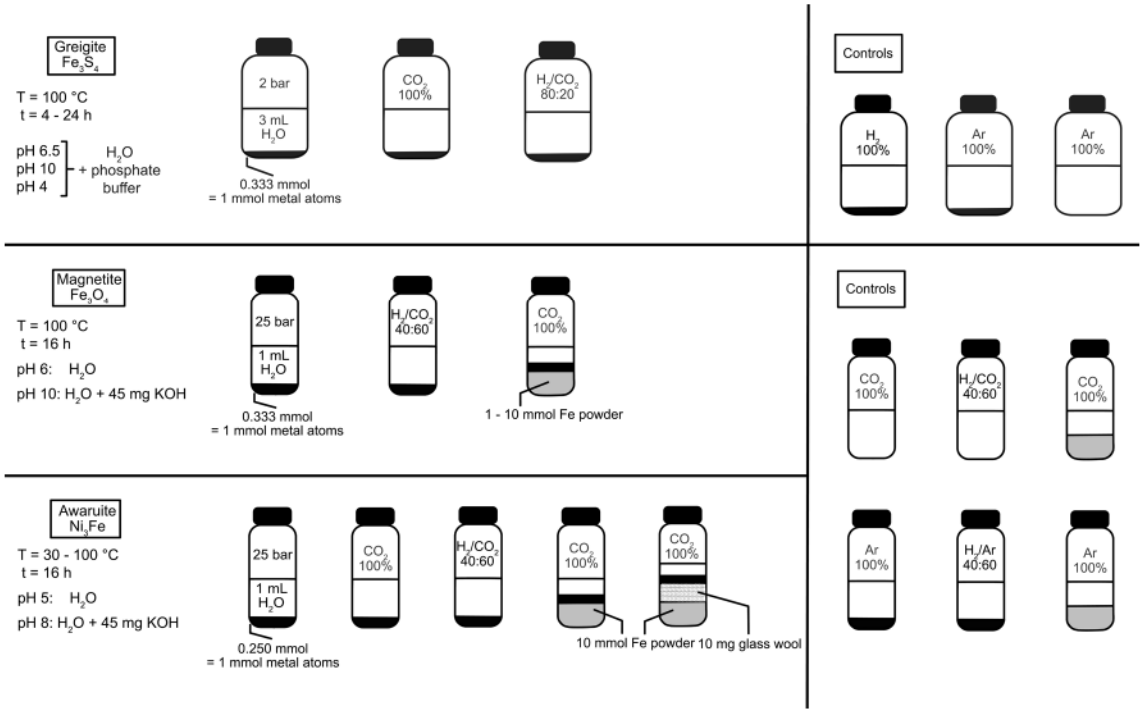
Extended data is available for this paper at https://doi.org/10.1038/s41559-020-1125-6.

Supplementary information is available for this paper at https://doi.org/10.1038/s41559-020-1125-6.

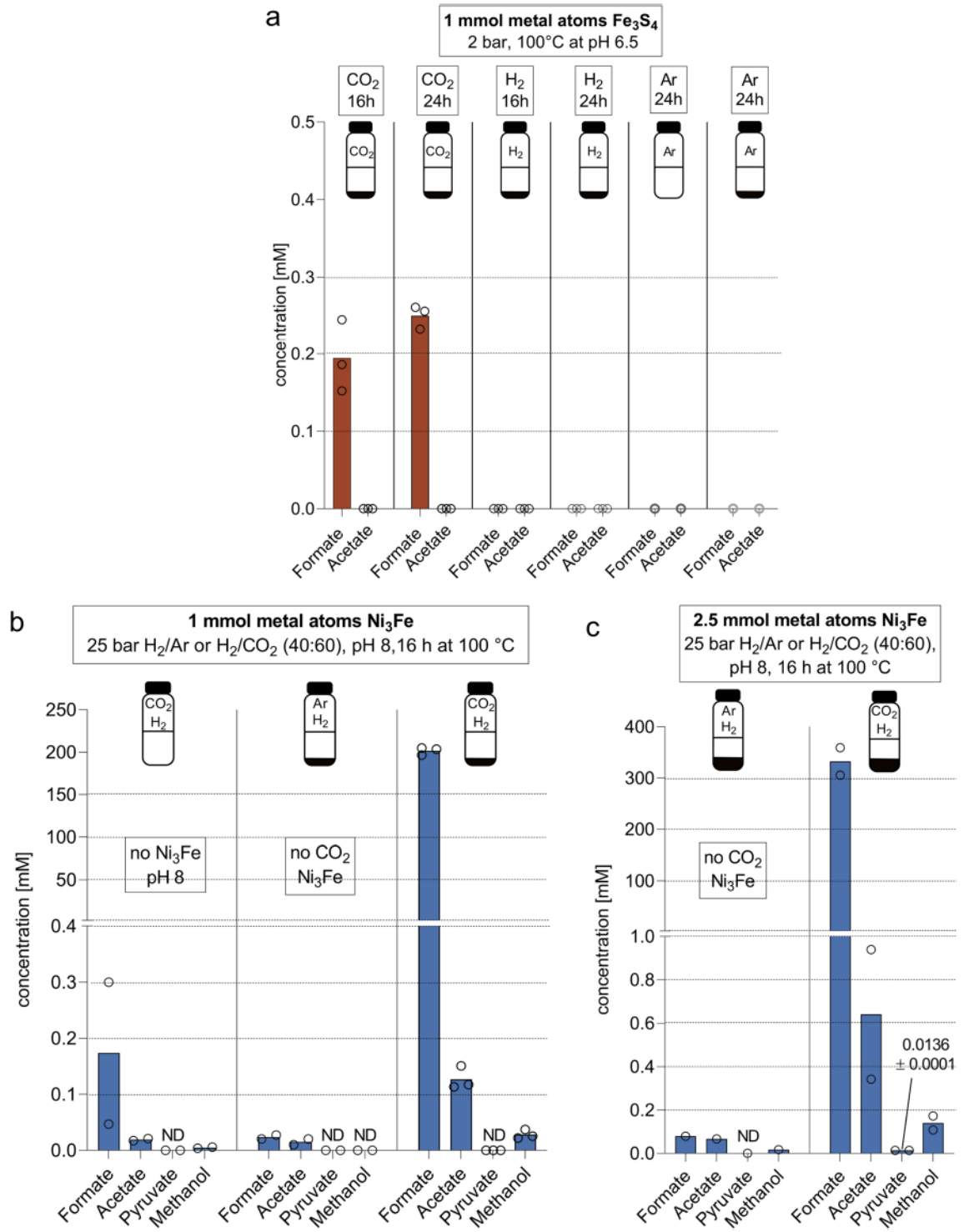
Correspondence and requests for materials should be addressed to H.T., J.M. or W.F.M.

Reprints and permissions information is available at www.nature.com/reprints.

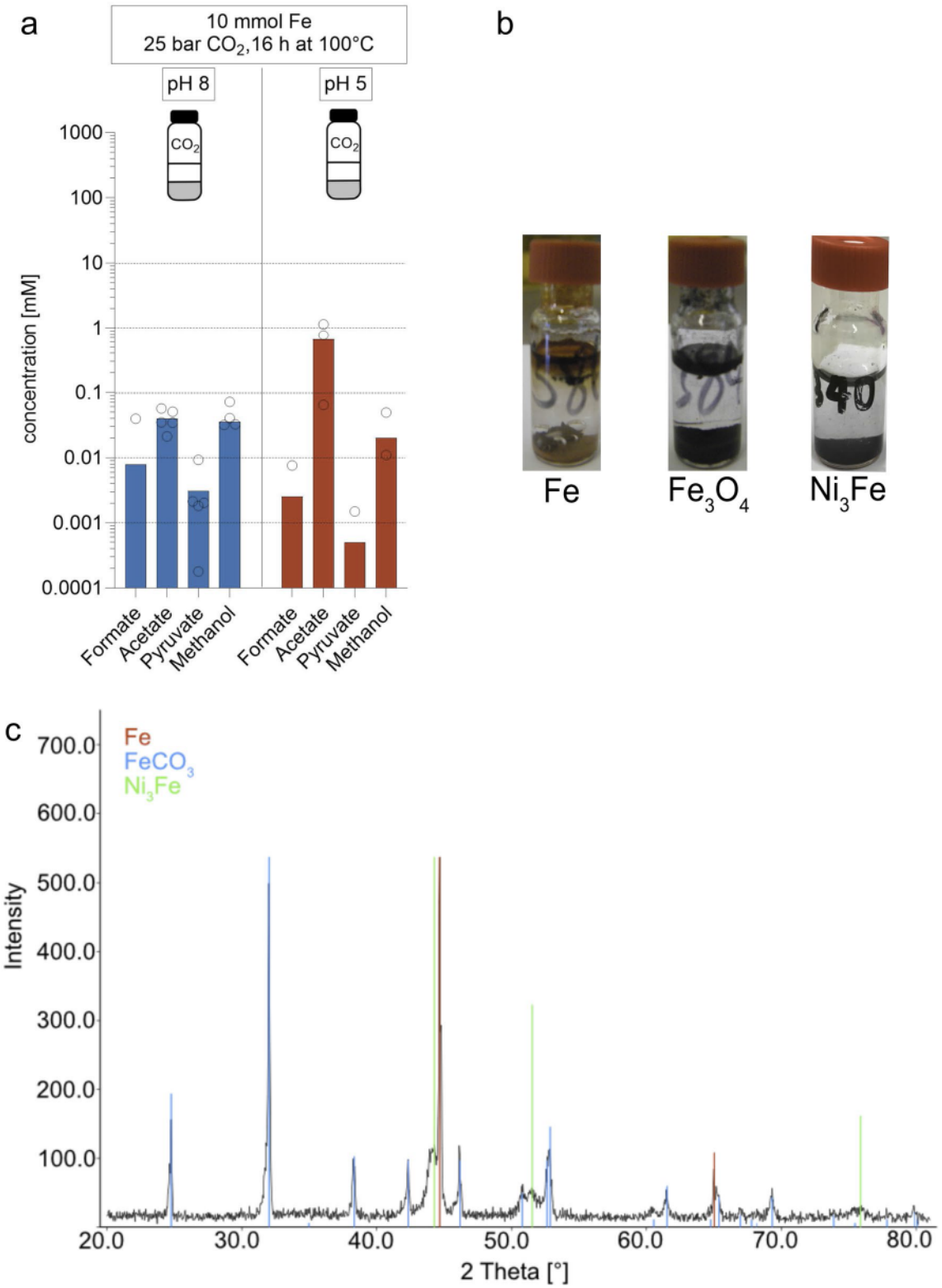
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



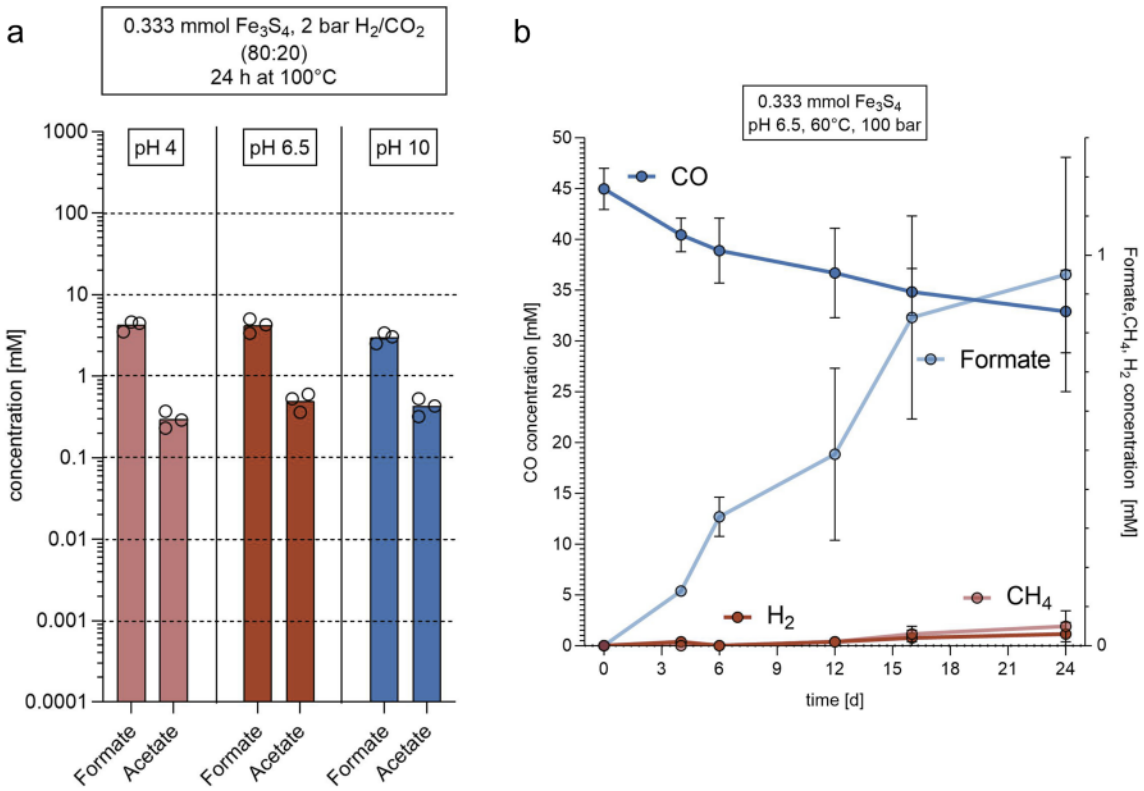
Extended Data Fig. 1 | Overview of the main experiments performed in this study. Three different iron-containing hydrothermal minerals were tested for their ability to catalyse the reaction between CO_2 and H_2: greigite (Fe_3S_4), magnetite (Fe_3O_4), and awaruite (Ni_3Fe).



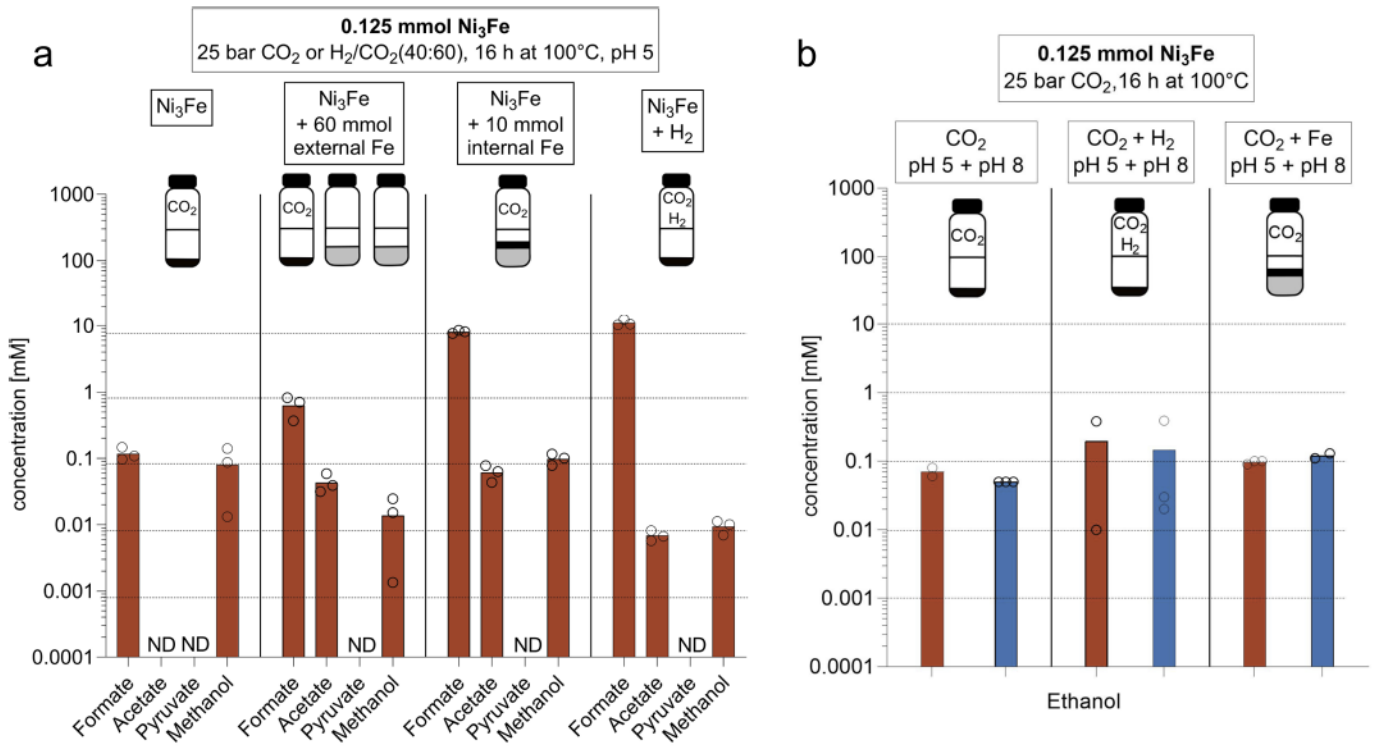
Extended Data Fig. 2 | Background controls for CO_2 fixation. a Control runs for greigite experiments—each run was performed either in the absence of H_2 or CO_2, or both. Circles show individual measurements. CO_2-only and H_2-only controls were performed at least in triplicate – values of 0 are not shown by the logarithmic scale. Four types of control experiments under the conditions of greigite experiments were performed with the catalyst: one under 100% CO_2 atmosphere, one under 100% H_2 atmosphere and two under Ar atmosphere (one with, one without catalyst). The mass spectra of the argon-controls are listed in the supplemental material. The CO_2-only experiments show that formate can be formed in small amounts without H_2 gas, suggesting that the electrons necessary for CO_2 reduction can also come from greigite's Fe^{2+} ions, either by forming H_2 from water or by a direct reduction of CO_2. The latter seems less probable as the step from CO_2 to formate is a 2-electron reaction, which electrons Fe^{2+} cannot provide (see also the proposed mechanism in Extended Data Fig. 10). **b** Comparison between background and product concentration in awaruite experiments with 1 mmol metal atoms Ni_3Fe (16 h at 100 °C, 25 bar, pH > 7). Both CO_2 fixation background without Ni_3Fe and the background of Ni_3Fe itself under an Ar/H_2 atmosphere are significantly lower than after H_2-dependent CO_2 reduction with Ni_3Fe. **c** Comparing background and product concentration in awaruite experiments with 2.5 mmol metal atoms Ni_3Fe (16 h at 100 °C, 25 bar, pH > 7). More details on the background contamination in awaruite and magnetite experiments are listed in Supplementary Tables 3–6.



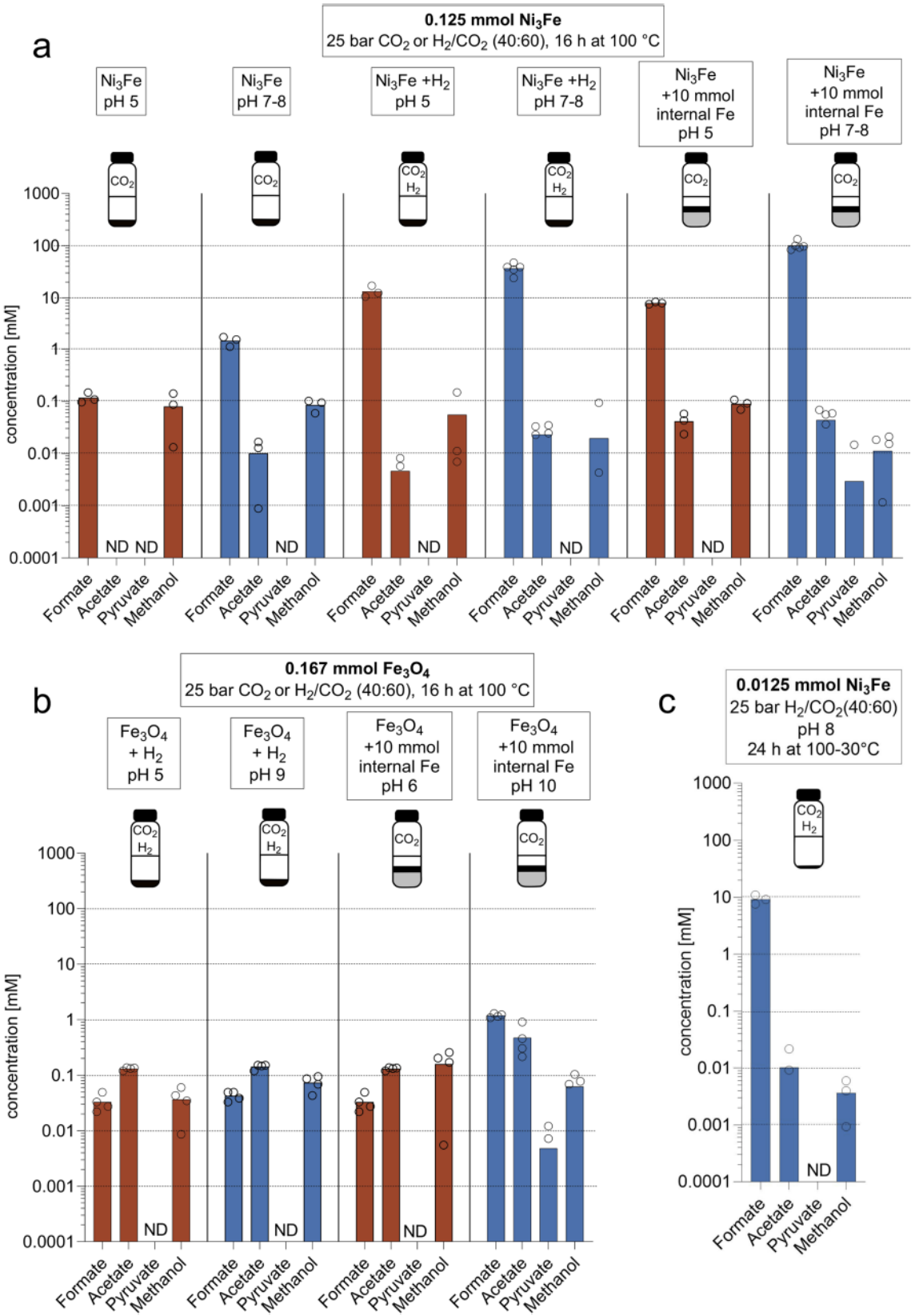
Extended Data Fig. 3 | Controls with added Fe^0 powder. a Control for internal Fe^0 runs (relevant to magnetite and awaruite experiments). Red: pH < 7; Blue: pH > 7. Circles show individual measurements, all controls were performed in at least triplicate – values of 0 are not shown by the logarithmic scale. As shown in Varma et al.[18], a KOH workup after the reaction was necessary prior to analysis. Here we show that KOH can also be added before the reaction (blue bars) in order to liberate the products into the solution. For reactions at lower pH (red bars), KOH was added afterwards. In both cases Fe powder promotes the reduction of CO_2 to formate, acetate, pyruvate and methanol. The exact mechanism remains unclear, but it is probable that the Fe powder is being oxidized, leading to the production of H_2 from H_2O. **b** Pictures after the reaction at pH 8, 16 h, 100 °C, 25 bar CO_2/H_2, showing the visual level of oxidation of Fe^0, Fe_3O_4 and Ni_3Fe. **c** XRD of iron powder after a reaction with 0.125 mmol Ni_3Fe on top (25 bar CO_2, 16 h, pH 8, 100 °C), the results of the CO_2 fixation are shown in Extended Data Fig. 6. The XRD spectrum shows that a major part of the iron surface is converted into iron(II) carbonate (siderite, $FeCO_3$), thus confirming the oxidation of the iron powder and the precipitation of Fe^{2+} ions with carbonate at the same time.



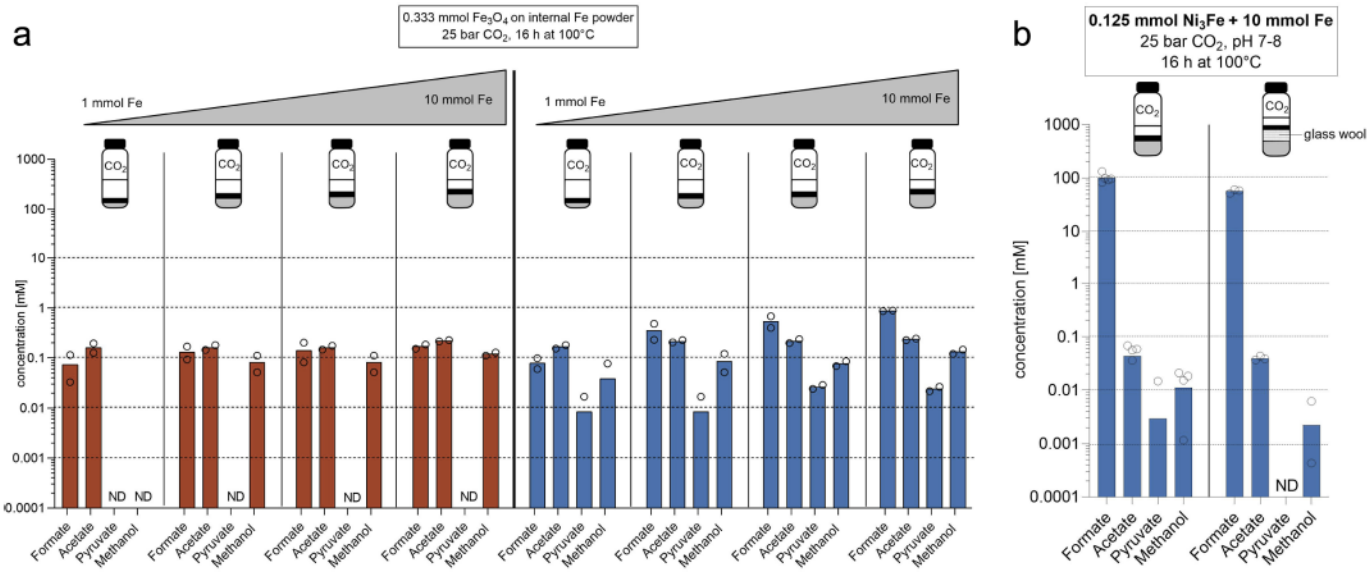
Extended Data Fig. 4 | Influence of pH and high-pressure CO experiments for greigite. a Influence of pH on greigite reactions. Red: pH < 7; Blue: pH > 7. Circles show individual measurements, all measurements performed in at least triplicate. The distribution of carbon fixation products in greigite reactions remains stable with changing pH. **b** Time course experiment of high-pressure methane and formate production from CO under greigite catalysis (liquid phase, 150 mL) at 60 °C and 100 bar. Using CO gas instead of CO_2 and H_2 does not explain the amount of methane produced (up to 0.04 mM). The rationale for using CO as a sole reactant stems from previous reports where small organics were obtained in appreciable quantities[16]. In reactions of CO with greigite and water, no organic products were found other than formate, whose carbon has the same redox state as CO.



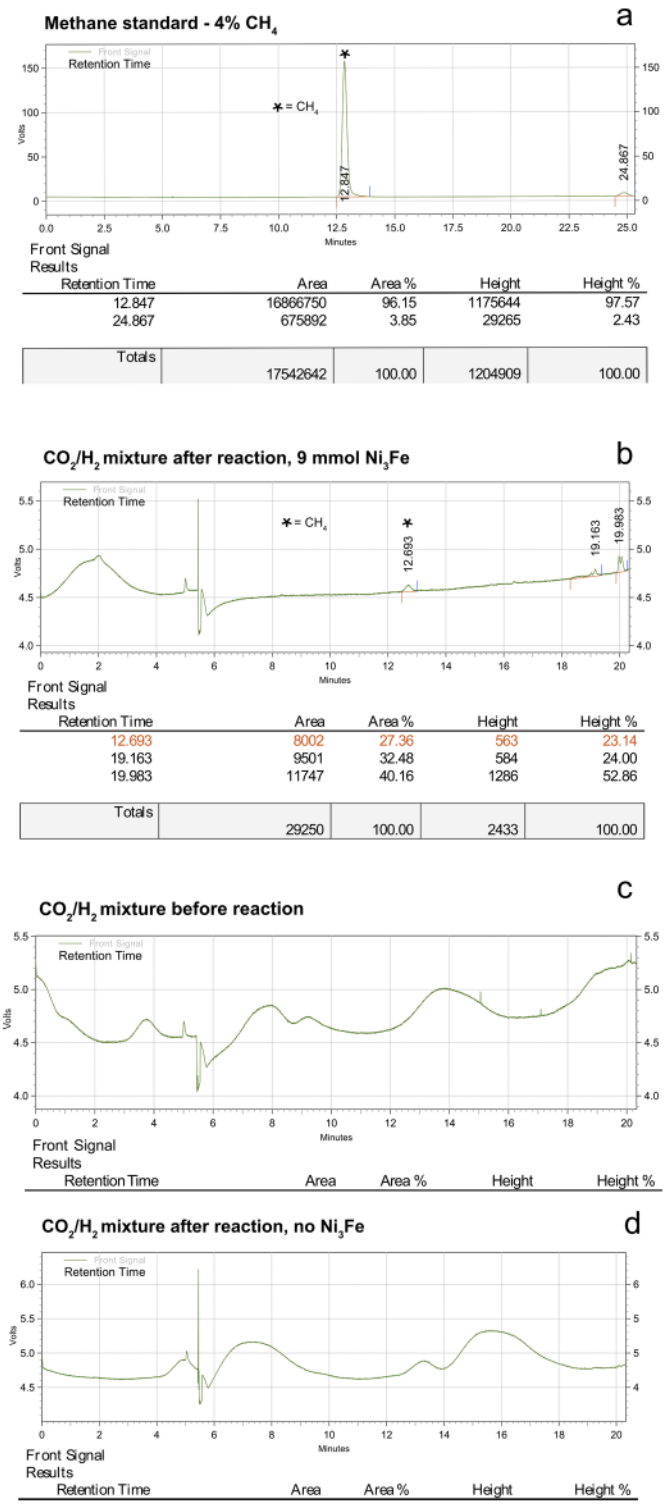
Extended Data Fig. 5 | Fe^0 as external and internal H_2 source and ethanol occurrence in awaruite experiments. a Screen of the quantity of Fe^0 as H_2 source. As a native metal compound, awaruite could serve as an electron donor and/or a catalyst. To establish the catalytic activity of the nanoparticular awaruite and the ability of Fe^0 powder to serve as an H_2 delivery system, the concentration of hydrogen gas within the reactor runs was gradually increased. First, 6 neighbouring vials filled with 10 mmol (560 mg) of iron powder ('external iron') each provided a small amount of H_2 to the atmosphere inside the pressure reactor. By adding 10 mmol iron powder into reaction vials ("internal iron") and placing the nanoparticular awaruite on top of it, especially formate concentration was pushed towards very high yields (8 mM at pH 5; 100 mM at pH 8, depicted in main Fig. 3b). **b** Occurrence of ethanol in Ni_3Fe reactions. Red: pH < 7; Blue: pH > 7. Circles show individual measurements; all measurements were performed in at least duplicate – values of 0 are not shown by the logarithmic scale. Ethanol was detected in concentrations up to over 0.1 mM. Its concentration fluctuated over the course of experiments, leading to the conclusion that it might be a product of decarboxylation of pyruvate or other side reactions.



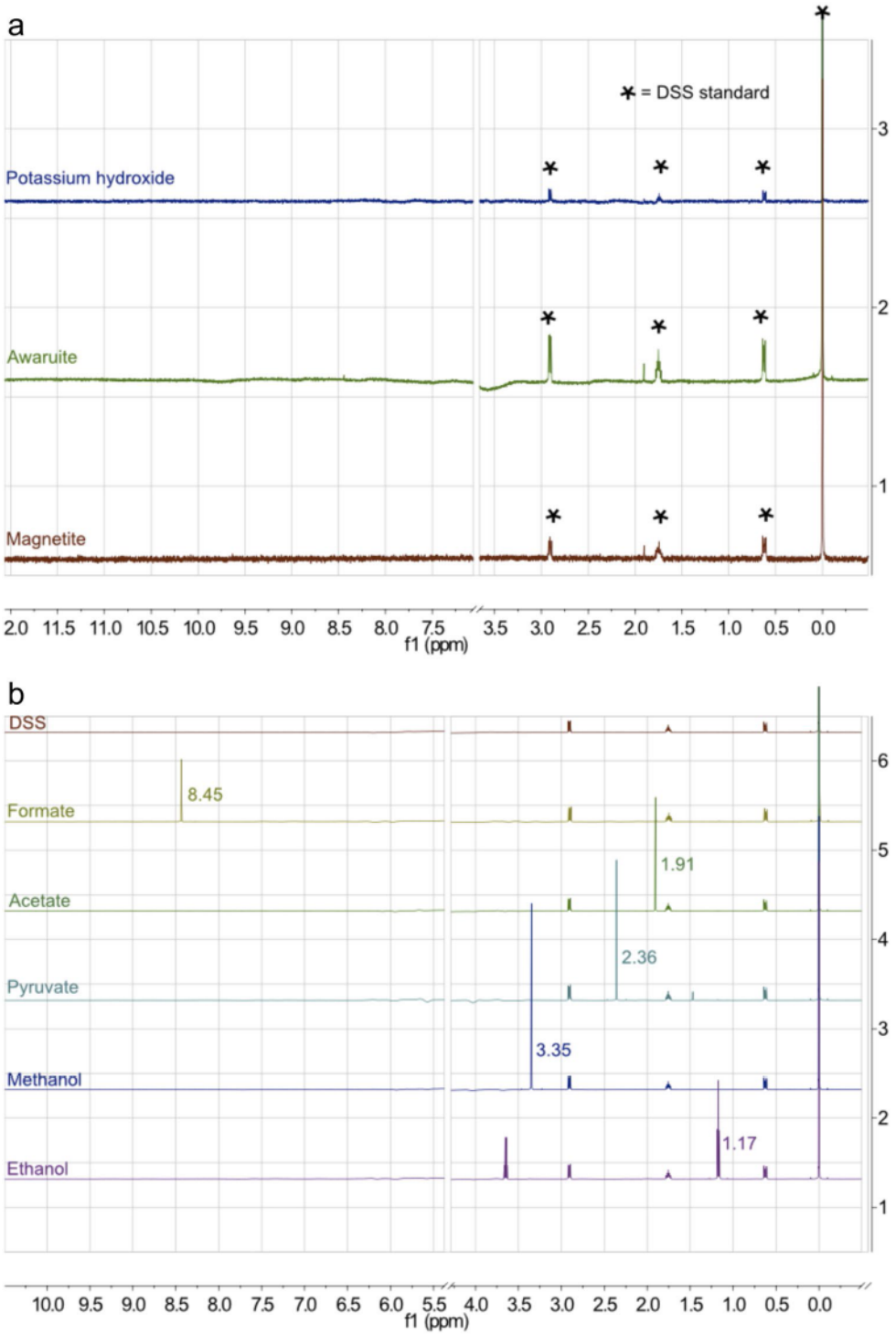
Extended Data Fig. 6 | Lower amounts of catalyst in magnetite and awaruite experiments. a Ni_3Fe catalysis of CO_2 fixation with 0.125 mmol catalyst (0.5 mmol metal atoms). Reactions mixtures with pH < 7 were not treated with KOH after the reaction. **b** Fe_3O_4 catalysis of CO_2 fixation with 0.167 mmol catalyst (0.5 mmol metal atoms). Reactions mixtures with pH < 7 were treated with KOH after reaction. Using less magnetite and less awaruite still yields a noticeable amount of product. **c** Ni_3Fe catalysis of CO_2 fixation at 100–30 °C. Red: pH < 7; Blue: pH > 7.Circles show individual measurements, all measurements were performed in at least triplicate – values of 0 are not shown by the logarithmic scale. Under this thermal gradient (8 h at 100 °C, 8 h at 70 °C, 8 h at 30 °C), even very small amounts (12.5 µmol = 50 µmol of metal atoms) of Ni_3Fe suffice to produce notable amounts of formate, acetate and methanol.



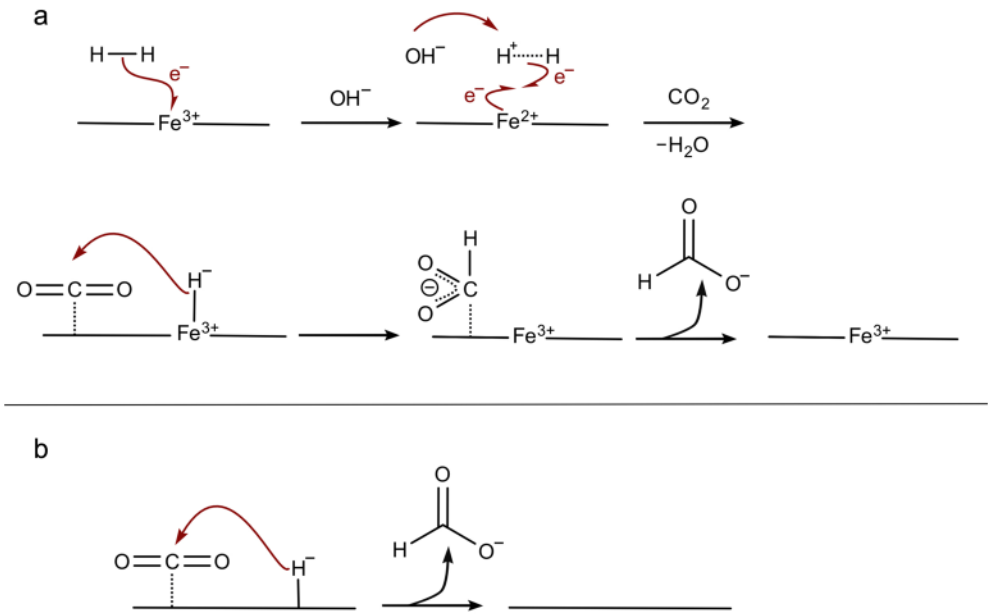
Extended Data Fig. 7 | Effect of physical contact between Fe^0 as internal H_2 source and catalyst and effect of Fe^0 amount. a Effect of Fe^0 as internal H_2 source with magnetite as catalyst. Red: pH < 7; Blue: pH > 7. Circles show individual measurements, all measurements were performed in duplicate – values of 0 are not shown by the logarithmic scale. In the absence of H_2 and with Fe^0 as reductant, CO_2 can be reduced in water in the presence of Fe_3O_4. Pyruvate accumulates at detectable levels at pH 10 but not at pH 6. The product increase with increasing iron quickly reaches an unexpected maximum, considering the minor differences in product concentration between the two samples with the highest Fe amount. A likely reason is that not the entire bulk of the iron powder reacts/becomes oxidized, but only the surface, thus inhibiting further interaction of the water molecules with the unreacted iron underneath the iron oxide and/or iron carbonate layer formed. **b** Separating awaruite (0.125 mmol Ni_3Fe = 0.5 mmol metal atoms) and internal iron powder. Circles show individual measurements, all measurements in triplicate – values of 0 are not shown by the logarithmic scale. Physical contact between the mineral and native iron is not required for product formation. As shown in Extended Data Fig. 3, placing the catalysts directly on the iron powder in order to "harvest" the nascent hydrogen ascending from the iron powder has the disadvantage that products on the surface of the iron will mix with the ones from awaruite (or magnetite). By tightly packing ca. 10 mg of decontaminated glass wool between iron and awaruite, this effect can be decreased. Although in the initial experiments, no pyruvate could be detected, the other products were formed in appreciable amounts.



Extended Data Fig. 8 | Detection of methane by GC-FID for awaruite experiments. a Gas analysis of 4% (40,000 ppm) methane standard. **b** CO_2/H_2 reaction with Ni_3Fe (awaruite) as catalyst. These measurements show that CH_4 is formed, although only in very small amounts. According to the methane standard measurement, the methane outcome of a typical experiment (16 h, 100 °C, 60:40 CO_2/H_2 atmosphere, 25 bar) containing 9 mmol of awaruite within one reactor run is roughly 19 ppm (determined by a one-point calibration). Two controls were performed: **c** one with the reactor pressurized (25 bar) with the 60:40 CO_2/H_2 gas mixture, and **d** one after subjecting the CO_2/H_2 gas mixture to the typical experimental conditions (16 h, 100 °C) within the otherwise empty reactor. Neither control experiment showed traces of methane.



Extended Data Fig. 9 | ^{1}H-NMR controls for catalysts and reagents and shifts of product peaks in the magnetite and awaruite experiments. a ^{1}H-NMR material controls for catalysts and reagents used for experiments with NMR product detection. Awaruite and magnetite were tested for surface contamination by treating it with a potassium hydroxide solution to cleave potential contaminants from the surface. Also potassium hydroxide was tested for contaminations. **b** ^{1}H NMR chemical shifts of product peaks observed and quantified in the magnetite and awaruite experiments. Quantification was achieved using two-point linear regression analysis.



Extended Data Fig. 10 | Proposed mechanisms for CO_2 reduction with H_2 catalysed by hydrothermal minerals. We propose an ionic mechanism for the catalysed two-electron reduction of CO_2 to formate in water for all three minerals **a** Proposed mechanism for CO_2 reduction with H_2 catalysed by Fe_3O_4 (or Fe_3S_4). An H_2 molecule approaches the Fe_3O_4 (or Fe_3S_4) surface and reduces Fe^{3+} to Fe^{2+}. The generated H_2^+ is unstable and decomposes to H^+, assisted by OH^- (accounting for increased product formation at pH > 7 in magnetite experiments) – and to a hydrogen atom (H·) which picks up an electron from Fe^{2+} to become a hydride (H^-). Hydride mechanisms were described in previous literature[67]. CO_2 on the other hand physisorbs on the magnetite surface and reacts with the hydride to yield $HCOO^-$ which is displaced from the surface by OH^-. Observations from experiments not discussed in this publication show, that experiments with minerals only containing ferrous iron (FeO and FeS) give far lower yields of CO_2 fixation products. **b** Proposed mechanism of CO_2 reduction catalysed by Ni_3Fe. As Ni_3Fe only consists of zero-valent metals, H_2 can dissociate on the metal surface. Then, the H atoms diffuse into awaruite where they can capture mobile ("free") electrons from the conduction band of the metal alloy.

Publication 8

Title: Catalysts, autocatalysis and the origin of metabolism

Year: 2019

Authors: Martina Preiner, Joana C. Xavier, Andrey do Nascimento Vieira, Karl Kleinermanns, John F. Allen, and William F. Martin

Published in: *Interface Focus,* 9(6), 20190072.
doi: 10.1098/rsfs.2019.0072.

Contribution: First and shared corresponding author.
Major: wrote substantial sections of the paper, edited heavily throughout all review processes and coordinated the distribution and responses with all authors. Drafted and illustrated Fig. 1.

Summary: This review article, besides giving an overview of the hydrothermal vent theory of the OoL, tries to connect simple inorganic catalysis to the emergence of complex autocatalytic networks. Furthermore, the parallel catalytic activation of CO_2, H_2 and N_2 is introduced as a possible, not yet attempted, hydrothermal route towards amino acids and nucleobases. This route provides a blueprint for further experimental research on heterogeneous mineral catalysts in OoL research.

INTERFACE FOCUS

royalsocietypublishing.org/journal/rsfs

Review

Cite this article: Preiner M, Xavier JC, Vieira do NA, Kleinermanns K, Allen JF, Martin WF. 2019 Catalysts, autocatalysis and the origin of metabolism. *Interface Focus* **9**: 20190072. http://dx.doi.org/10.1098/rsfs.2019.0072

Accepted: 30 August 2019

One contribution of 14 to a theme issue 'The origin of life: the submarine alkaline vent theory at 30'.

Subject Areas:
biogeochemistry, astrobiology, chemical biology

Keywords:
hydrothermal vents, catalysis, activation, energy, origin of life, prebiotic chemistry

Authors for correspondence:
Martina Preiner
e-mail: preiner@hhu.de
Joana C. Xavier
e-mail: xavier@hhu.de

THE ROYAL SOCIETY PUBLISHING

Catalysts, autocatalysis and the origin of metabolism

Martina Preiner[1], Joana C. Xavier[1], Andrey do Nascimento Vieira[1], Karl Kleinermanns[2], John F. Allen[3] and William F. Martin[1]

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

MP, 0000-0001-5137-5556; JCX, 0000-0001-9242-8968; ANV, 0000-0003-4468-7724; JFA, 0000-0002-0950-0429; WFM, 0000-0003-1478-6449

If life on Earth started out in geochemical environments like hydrothermal vents, then it started out from gasses like CO_2, N_2 and H_2. 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 H_2–CO_2 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 H_2 and CO_2 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 physico-chemical 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 H_2-rich effluent over geological timescales [17]. The main redox reaction they harbour is the H_2–CO_2 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 H_2 with CO_2.

The H_2–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 CO_2 with 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 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 CO_2 and H_2: the door to CO_2 fixation

In biology, acetogens and methanogens fix CO_2 via the H_2-dependent reduction of CO_2 to a methyl group and CO, followed by condensation of the methyl moiety and 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 H_2-dependent reduction of CO_2 to acetyl-CoA. Under standard physiological conditions, the synthesis of the thioester is exergonic by about –59 kJ mol^{-1} [27], while there is not enough energy to generate thioesters and synthesize 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 H_2 to have played a role in early chemical evolution, it required activation—it required catalysis. It is noteworthy that 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

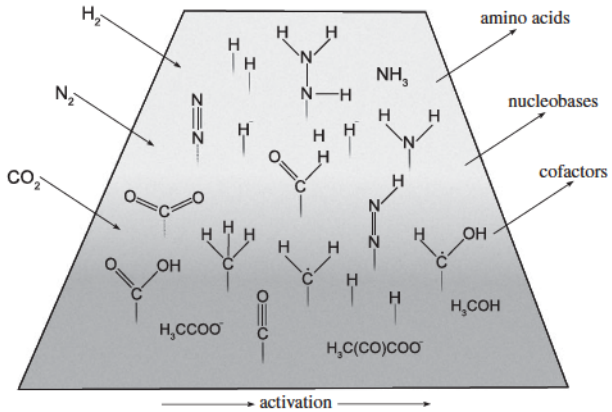


Figure 1. Simultaneous activation of H_2, CO_2 and N_2 on mineral surfaces leading to the formation of a variety of biologically relevant molecules, such as amino acids, nucleic acid bases and cofactors. Molecules, such as pyruvate, acetate, methanol and ammonia, are known to form on transition metal containing surfaces [15,45]. Little is known about the products obtained when the separation of N and C fixation is revoked. Heterogeneous catalysis may have been the key for early processes of protometabolism. Dashed lines indicate physisorption, non-dashed lines indicate chemisorption on the surface.

[32,33], all three harbour carbon metal bonds at their active site [26]. The central enzyme of the acetyl-CoA pathway, the only exergonic 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 H_2 [36–38]. Flavin-based electron bifurcation thus accounts for the thermodynamics of 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 CO_2 fixation with H_2.

The H_2-dependent reaction from the most oxidized form of carbon, CO_2, to its most reduced form, methane (CH_4), is thermodynamically favourable under reducing conditions. However, in serpentinizing, alkaline hydrothermal systems [39] the direct transfer of electrons from H_2 to 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 catalysis stems from kinetic barriers in the sequence of reactions from CO_2 to 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 CH_4. While high temperatures, high pressures and long reaction times lead to the accumulation of 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 can occur on mineral surfaces (figure 1)—which are not unlike the catalysts used in industry to produce hydrocarbons [15,25]. The activation of molecules on mineral surfaces is likely to have preceded

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 (N_2) into biosynthetic pathways. From a chemical point of view, N_2 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 N_2 would have led to an ocean with dissolved N_2 and thus—via sequestration through the Earth's crust—to a nitrogen source in serpentinizing systems [49,50]. Looking at biology, N_2 fixation is considered ancient [50,51]. There is only one way for N_2 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 Fe_7S_9 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 N_2 molecule [53,54]. The following steps involve sequential hydrogenations of the nitrogen molecule. There, as for CO_2 fixation, hydrogenase activity is needed to deliver electrons from H_2 to N_2. This hydrogenase activity is promoted by the FeS clusters of the nitrogenase complex [53,55], which is the sole entry point of N_2 into metabolism. As CODH and hydrogenase, nitrogenase also traces back to LUCA [26,56].

Biology operates within constraints of temperature and pressure. Biological N_2 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 N_2. In industry, the reduction of N_2 might resemble prebiotic FeS-based nitrogen fixation [45]. The greatest impediment to N_2 reduction is its activation energy. N_2 is very stable at normal atmospheric temperatures and pressures. Thus, few processes are capable of activating N_2 sufficiently in order to form N-rich molecules. Industrial N_2 conversion to NH_3 via the Haber–Bosch process (H_2-dependent) requires Fe-based catalysts such as Fe_3O_4, 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 N_2 fixation [57].

Not unlike the stepwise use of Fe atoms found in the active sites of the nitrogenase complex, industrial N_2 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 Fe_xN [59].

Can serpentinization reduce N_2? Although there is abundant evidence for abiotic CO_2 reduction in serpentinizing systems [60,61], evidence for abiotic N_2 reduction is so far lacking. Laboratory simulations suggest that N_2 can be reduced to ammonia (NH_3) with mineral catalysts under mild hydrothermal conditions [45,62]. Incorporation of N from N_2 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 NH_3 [25], but this remains to be shown experimentally.

There are nevertheless very curious parallels between industrial hydrogenation processes and geochemical H_2-dependent reactions. Serpentinization not only reduces H_2O to H_2 and CO_2 to formate and CH_4, it also generates inorganic catalysts within the Earth's crust [25]. These include magnetite, Fe_3O_4, which is the catalyst of choice for the industrial Haber–Bosch process (H_2-dependent N_2 reduction) and for Fischer–Tropsch (CO_2 reduction) applications [59,63] and awaruite, Ni_3Fe, which catalyses the H_2-dependent reduction of CO_2 to methane at high pressures and temperatures [40]. While H_2 and CO_2 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 H_2, CO_2 and N_2 simultaneously. If so, amino acids or even bases and cofactors might be obtained via such routes. It has been reported that Fe^{2+} and Fe^0 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 synthesize pyruvate from H_2 and CO_2 [15]. Pyruvate now appears to be a much more readily synthesized prebiotic compound than previously assumed.

If N_2 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 CO_2 (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 CO_2 and N from aspartate.

There is a clear record of geochemical origins preserved in metabolism [26]. This record can be resurrected in the

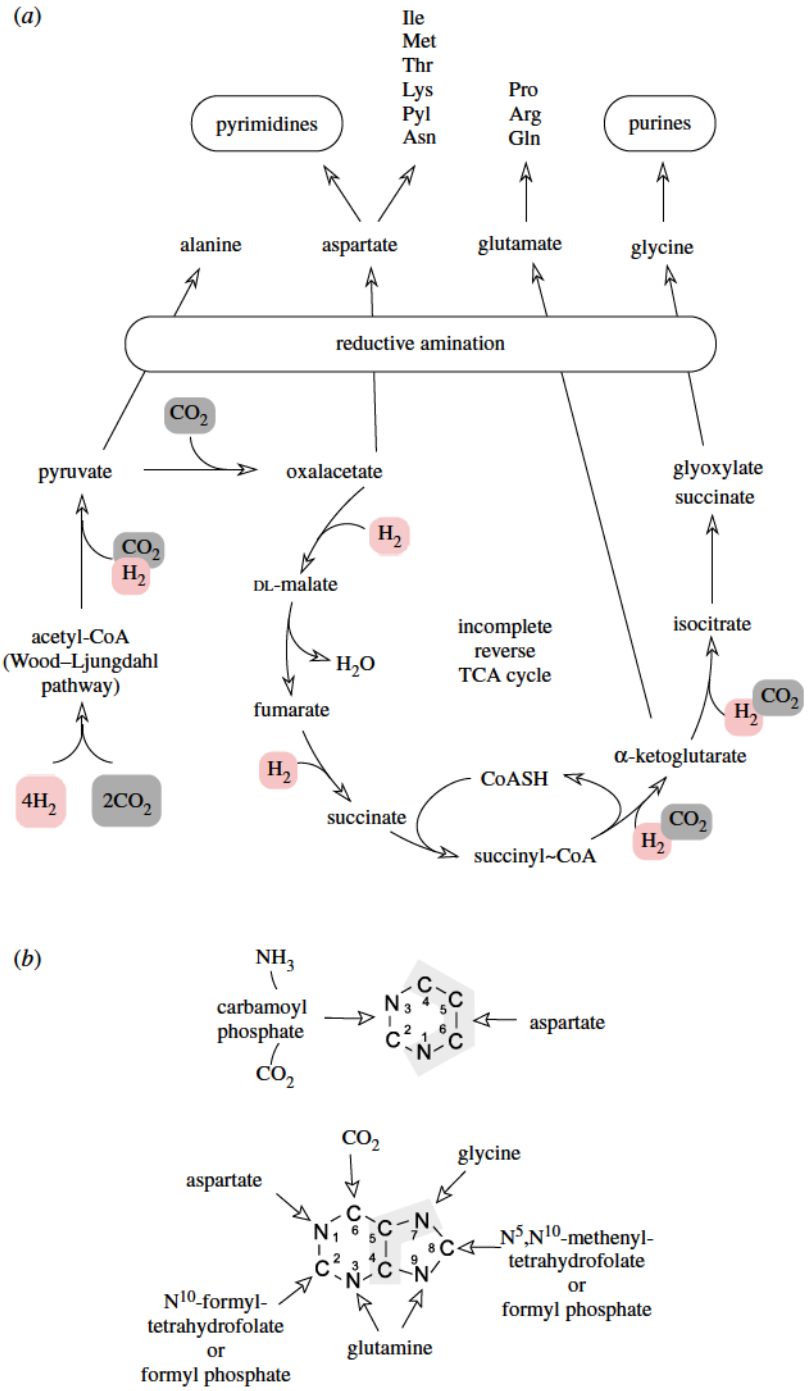


Figure 2. A path from H_2 and CO_2 to nucleic acid bases (adapted from figs. 3, 4 and 6 of [9]. (*a*) The lower portion of the panel shows the biosynthesis of carbon backbones in microbes that use the acetyl-CoA pathway and the incomplete (horseshoe) reverse citric acid cycle. Reductive steps of CO_2 fixation are indicated as H_2-dependent, though reduced ferredoxin or NAD(P)H are the reductants in metabolism. The first four 2-oxoacids to arise via the route shown, and, if reductively aminated, generate Ala, Asp, Glu and Gly (upper portion of the panel). Muchowska *et al*. [64] showed that pyruvate, oxaloacetate, 2-oxoglutarate and glyoxylate are readily reduced to the corresponding amino acids by hydroxylamine under mild conditions in the presence of native iron. Asp is the starting point for biosynthesis of five other canonical amino acids and pyrrolysine, Glu is the starting point for synthesis of Gln, Arg and Pro. (*b*) Asp and Gly are central to pyrimidine and purine biosynthesis, respectively (modified from fig. 4 of [9]). The involvement of CO_2 in purine and pyrimidine synthesis is noteworthy, as is the involvement of folate-bound C1 intermediates of the acetyl-CoA pathway in purine synthesis, which are replaced by the simpler intermediate formyl phosphate in methanogens. This suggests the possibility of a small prebiotic biochemical network linking CO_2 reduction to nucleic acid base synthesis. (Online version in colour.)

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 N_2, H_2 and CO_2 can lead to thermodynamically stable products that include

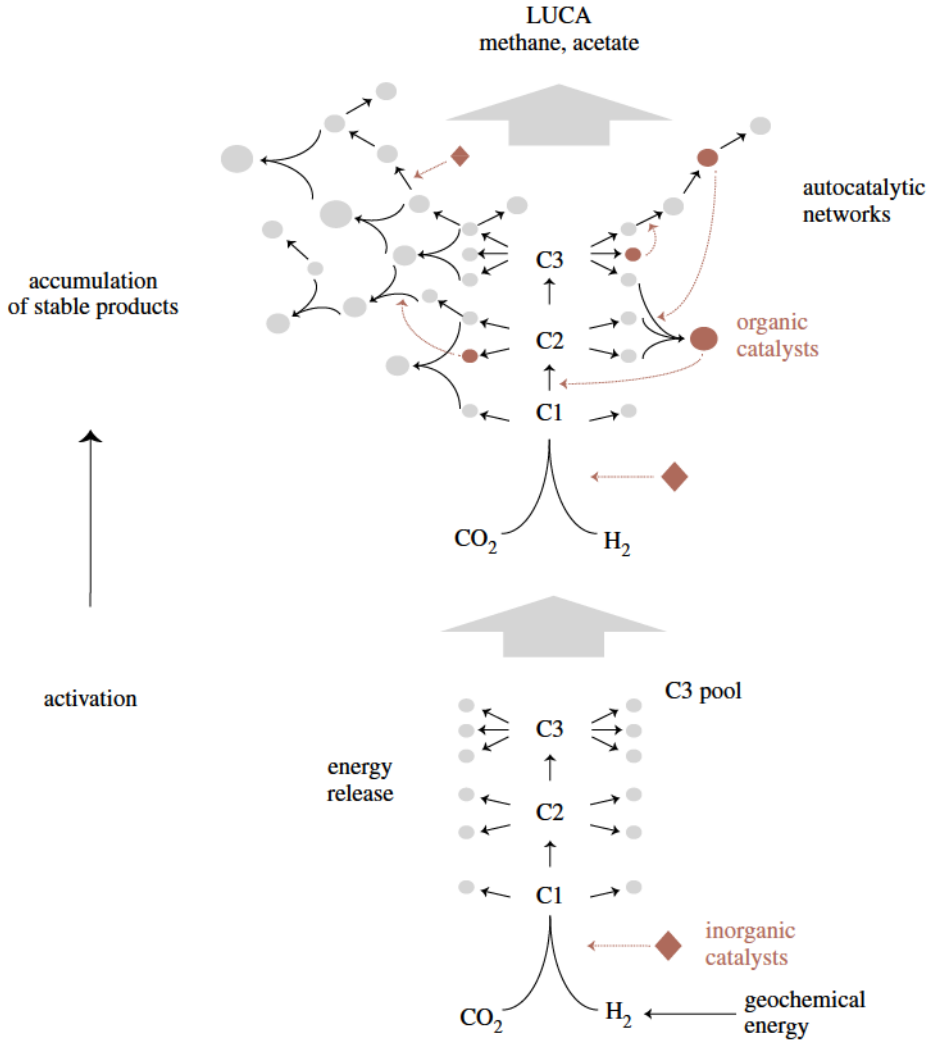


Figure 3. Purely geochemical reactions such as CO_2 fixation with H_2 can give rise to autocatalytic networks and protometabolism, as long as energy is released. Kinetically controlled reactions build up a specific set of products which interact further to form an autocatalytic network that serves as 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.)

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.

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^- 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_2S (HS^- 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

1. Zahnle K, Arndt N, Cockell C, Halliday A, Nisbet E, Selsis F, Sleep NH. 2007 Emergence of a habitable planet. *Space Sci. Rev.* **129**, 35–78. (doi:10.1007/s11214-007-9225-z)
2. Tashiro T, Ishida A, Hori M, Igisu M, Koike M, Méjean P, Takahata N, Sano Y, Komiya T. 2017 Early trace of life from 3.95 Ga sedimentary rocks in Labrador, Canada. *Nature* **549**, 516–518. (doi:10.1038/nature24019)
3. Baross JA, Hoffman SE. 1985 Submarine hydrothermal vents and associated gradient environments as sites for the origin and evolution of life. *Orig. Life Evol. Biosph.* **15**, 327–345. (doi:10.1007/BF01808177)
4. Russell MJ, Hall AJ. 1997 The emergence of life from iron monosulphide bubbles at a submarine hydrothermal redox and pH front. *J. Geol. Soc. London.* **154**, 377–402. (doi:10.1.1.520.7821)
5. Nisbet E, Sleep NH. 2001 The habitat and nature of early life. *Nature* **409**, 1083–1091. (doi:10.1038/35059210)
6. Amdt NT, Nisbet EG. 2012 Processes on the young Earth and the habitats of early life. *Annu. Rev. Earth Planet. Sci.* **40**, 521–549. (doi:10.1146/annurev-earth-042711-105316)
7. Sleep NH. 2018 Geological and geochemical constraints on the origin and evolution of life. *Astrobiology* **18**, 1199–1219. (doi:10.1089/ast.2017.1778)
8. McCollom TM, Seewald JS. 2006 Carbon isotope composition of organic compounds produced by abiotic synthesis under hydrothermal conditions.

Earth Planet. Sci. Lett. **243**, 74–84. (doi:10.1016/j.epsl.2006.01.027)

9. Martin W, Russell MJ. 2007 On the origin of biochemistry at an alkaline hydrothermal vent. *Phil. Trans. R. Soc. B* **362**, 1887–1925. (doi:10.1098/rstb.2006.1881)
10. McDermott JM, Seewald JS, German CR, Sylva SP. 2015 Pathways for abiotic organic synthesis at submarine hydrothermal fields. *Proc. Natl Acad. Sci. USA* **112**, 7668–7672. (doi:10.1073/pnas.1506295112)
11. Camprubi E, Jordan SF, Vasiliadou R, Lane N. 2017 Iron catalysis at the origin of life. *IUBMB Life* **69**, 373–381. (doi:10.1002/iub.1632)
12. Baross JA. 2018 The rocky road to biomolecules. *Nature* **564**, 42–43. (doi:10.1038/d41586-018-07262-8)
13. Ménez B, Pisapia C, Andreani M, Jamme F, Vanbellingen QP, Brunelle A, Richard L, Dumas P, Réfrégiers M. 2018 Abiotic synthesis of amino acids in the recesses of the oceanic lithosphere. *Nature* **564**, 59–63. (doi:10.1038/s41586-018-0684-z)
14. Varma SJ, Muchowska KB, Chatelain P, Moran J. 2018 Native iron reduces CO_2 to intermediates and endproducts of the acetyl-CoA pathway. *Nat. Ecol. Evol.* **2**, 1019–1024. (doi:10.1038/s41559-018-0542-2)
15. Preiner M *et al*. 2019 A hydrogen dependent geochemical analogue of primordial carbon and energy metabolism. *BioRxiv* (doi:10.1101/682955)
16. Martin W, Baross J, Kelley D, Russell MJ. 2008 Hydrothermal vents and the origin of life. *Nat. Rev. Microbiol.* **6**, 805–814. (doi:10.1038/nrmicro1991)
17. Baross JA, Martin WF. 2015 The ribofilm as a concept for life's origins. *Cell* **162**, 13–15. (doi:10.1016/j.cell.2015.06.038)
18. Thauer RK, Jungermann K, Decker K. 1977 Energy conservation in chemotropic anaerobic bacteria. *Bacteriol. Rev.* **41**, 100–180.
19. McCollom TM, Amend JP. 2005 A thermodynamic assessment of energy requirements for biomass synthesis by chemolithoautotrophic micro-organisms in oxic and anoxic environments. *Geobiology* **3**, 135–144. (doi:10.1111/j.1472-4669.2005.00045.x)
20. Schrenk MO, Brazelton WJ, Lang SQ. 2013 Serpentinization, carbon, and deep life. *Rev. Mineral. Geochemistry* **75**, 575–606. (doi:10.2138/rmg.2013.75.18)
21. Hickman-Lewis K, Cavalazzi B, Foucher F, Westall F. 2018 Most ancient evidence for life in the Barberton greenstone belt: microbial mats and biofabrics of the ~3.47 Ga Middle Marker horizon. *Precambrian Res.* **312**, 45–67. (doi:10.1016/j.precamres.2018.04.007)
22. Brazelton WJ, Schrenk MO, Kelley DS, Baross JA. 2006 Methane- and sulfur-metabolizing microbial communities dominate the Lost City Hydrothermal Field ecosystem. *Appl. Environ. Microbiol.* **72**, 6257–6270. (doi:10.1128/AEM.00574-06)
23. Weiss MC, Sousa FL, Mrnjavac N, Neukirchen S, Roettger M, Nelson-Sathi S, Martin WF. 2016 The physiology and habitat of the last universal common ancestor. *Nat. Microbiol.* **1**, 16116. (doi:10.1038/nmicrobiol.2016.116)
24. Sousa FL, Martin WF. 2014 Biochemical fossils of the ancient transition from geoenergetics to bioenergetics in prokaryotic one carbon compound metabolism. *Biochim. Biophys. Acta - Bioenerg.* **1837**, 964–981. (doi:10.1016/j.bbabio.2014.02.001)
25. Preiner M *et al*. 2018 Serpentinization: connecting geochemistry, ancient metabolism and industrial hydrogenation. *Life* **8**, 41. (doi:10.3390/life8040041)
26. Martin WF. 2019 Carbon–metal bonds: rare and primordial in metabolism. *Trends Biochem. Sci.* **44**, 807–818. (doi:10.1016/j.tibs.2019.04.010)
27. Fuchs G. 2011 Alternative pathways of carbon dioxide fixation: insights into the early evolution of life? *Annu. Rev. Microbiol.* **65**, 631–658. (doi:10.1146/annurev-micro-090110-102801)
28. Martin WF. 2012 Hydrogen, metals, bifurcating electrons, and proton gradients: the early evolution of biological energy conservation. *FEBS Lett.* **586**, 485–493. (doi:10.1016/j.febslet.2011.09.031)
29. Schuchmann K, Müller V. 2013 Direct and reversible hydrogenation of CO_2 to formate by a bacterial carbon dioxide reductase. *Science* **342**, 1382–1386. (doi:10.1126/science.1244758)
30. Thauer RK, Kaster AK, Seedorf H, Buckel W, Hedderich R. 2008 Methanogenic archaea: ecologically relevant differences in energy conservation. *Nat. Rev. Microbiol.* **6**, 579–591. (doi:10.1038/nrmicro1931)
31. Goldford JE, Hartman H, Smith TF, Segrè D. 2017 Remnants of an ancient metabolism without phosphate. *Cell* **168**, 1126–1134. (doi:10.1016/j.cell.2017.02.001)
32. Thauer RK, Kaster A-K, Goenrich M, Schick M, Hiromoto T, Shima S. 2010 Hydrogenases from methanogenic archaea, nickel, a novel cofactor, and H_2 storage. *Annu. Rev. Biochem.* **79**, 507–536. (doi:10.1146/annurev.biochem.030508.152103)
33. Thauer RK. 2011 Hydrogenases and the global H_2 cycle. *Eur. J. Inorg. Chem.* **2011**, 919–921. (doi:10.1002/ejic.201001255)
34. Berg IA. 2011 Ecological aspects of the distribution of different autotrophic CO_2 fixation pathways. *Appl. Environ. Microbiol.* **77**, 1925–1936. (doi:10.1128/AEM.02473-10)
35. Xavier JC, Preiner M, Martin WF. 2018 Something special about CO-dependent CO_2 fixation. *FEBS Lett.* **285**, 4181–4195. (doi:10.1111/febs.14664)
36. Buckel W, Thauer RK. 2013 Energy conservation via electron bifurcating ferredoxin reduction and proton/Na^+ translocating ferredoxin oxidation. *Biochim. Biophys. Acta - Bioenerg.* **1827**, 94–113. (doi:10.1016/j.bbabio.2012.07.002)
37. Schoelmerich MC, Müller V. 2019 Energy conservation by a hydrogenase-dependent chemiosmotic mechanism in an ancient metabolic pathway. *Proc. Natl Acad. Sci. USA* **116**, 6329–6334. (doi:10.1073/pnas.1818580116)
38. Peters JW, Miller AF, Jones AK, King PW, Adams MWW. 2016 Electron bifurcation. *Curr. Opin. Chem. Biol.* **31**, 146–152. (doi:10.1016/j.cbpa.2016.03.007)
39. Klein F, Bach W. 2009 Fe-Ni-Co-O-S phase relations in peridotite–seawater interactions. *J. Petrol.* **50**, 37–59. (doi:10.1093/petrology/egn071)
40. Horita J, Berndt ME. 1999 Abiogenic methane formation and isotopic fractionation under hydrothermal conditions. *Sci. Rep.* **285**, 1055–1057. (doi:10.1126/science.285.5430.1055)
41. Maden BEH. 2000 Tetrahydrofolate and tetrahydromethanopterin compared: functionally distinct carriers in C1 metabolism. *Biochem. J.* **350**, 609–629. (doi:10.1042/0264-6021:3520935)
42. Appel AM *et al*. 2013 Frontiers, opportunities, and challenges in biochemical and chemical catalysis of CO_2 fixation. *Chem. Rev.* **113**, 6621–6658. (doi:10.1021/cr300463y)
43. McCollom TM, Seewald JS. 2003 Experimental study of the hydrothermal reactivity of organic acids and acid anions: II. Acetic acid, acetate, and valeric acid. *Geochim. Cosmochim. Acta* **67**, 3645–3664. (doi:10.1016/S0016-7037(03)00135-2)
44. Wolfenden R. 2011 Benchmark reaction rates, the stability of biological molecules in water, and the evolution of catalytic power in enzymes. *Annu. Rev. Biochem.* **80**, 645–667. (doi:10.1146/annurev-biochem-060409-093051)
45. Dörr M *et al*. 2003 A possible prebiotic formation of ammonia from dinitrogen on iron sulfide surfaces. *Angew. Chemie Int. Ed.* **42**, 1540–1543. (doi:10.1002/anie.200250371)
46. Wächtershäuser G. 1988 Before enzymes and templates: theory of surface metabolism. *Microbiol. Rev.* **52**, 452–484.
47. Russell MJ, Hall AJ, Gize AP. 1990 Pyrite and the origin of life. *Nature* **344**, 387. (doi:10.1038/344387b0)
48. Sousa FL, Thiergart T, Landan G, Nelson-Sathi S, Pereira IAC, Allen JF, Lane N, Martin WF. 2013 Early bioenergetic evolution. *Phil. Trans. R. Soc. B* **368**, 20130088. (doi:10.1098/rstb.2013.0088)
49. Voss M, Bange HW, Dippner JW, Middelburg JJ, Montoya JP, Ward B. 2013 The marine nitrogen cycle: recent discoveries, uncertainties and the potential relevance of climate change. *Phil. Trans. R. Soc. B* **368**, 20130121. (doi:10.1098/rstb.2013.0121)
50. Stüeken EE, Kipp MA, Koehler MC, Buick R. 2016 The evolution of Earth's biogeochemical nitrogen cycle. *Earth-Sci. Rev.* **160**, 220–239. (doi:10.1016/j.earscirev.2016.07.007)
51. Nishizawa M, Miyazaki J, Makabe A, Koba K, Takai K. 2014 Physiological and isotopic characteristics of nitrogen fixation by hyperthermophilic methanogens: key insights into nitrogen anabolism of the microbial communities in Archean hydrothermal systems. *Geochim. Cosmochim. Acta* **138**, 117–135. (doi:10.1016/j.gca.2014.04.021)
52. Hoffman BM, Lukoyanov D, Dean DR, Seefeldt LC. 2013 Nitrogenase: a draft mechanism. *Acc. Chem. Res.* **46**, 587–595. (doi:10.1021/ar300267m)
53. Hoffman BM, Lukoyanov D, Yang Z-Y, Dean DR, Seefeldt LC. 2014 Mechanism of nitrogen fixation by nitrogenase: the next stage. *Chem. Rev.* **114**, 4041–4062. (doi:10.1021/cr400641x)
54. Barney BM, Lee HI, Dos Santos PC, Hoffman BM, Dean DR, Seefeldt LC. 2006 Breaking the N_2 triple bond: insights into the nitrogenase mechanism.

Dalt. Trans. **21**, 2277–2284. (doi:10.1039/b517633f)

55. Siegbahn PEM. 2019 The mechanism for nitrogenase including all steps. *Phys. Chem. Chem. Phys.* **21**, 15 747–15 759. (doi:10.1039/C9CP02073J)
56. Weiss MC, Preiner M, Xavier JC, Zimorski V, Martin F. 2018 The last universal common ancestor between ancient Earth chemistry and the onset of genetics. *PLoS Genet.* **14**, e1007518. (doi:10.1371/journal.pgen.1007518)
57. Cherkasov N, Ibhadon AO, Fitzpatrick P. 2015 A review of the existing and alternative methods for greener nitrogen fixation. *Chem. Eng. Process. Process Intensif.* **90**, 24–33. (doi:10.1016/j.cep.2015.02.004)
58. Leigh GJ. 2004 Haber–Bosch and other industrial processes. In *Catalysts for nitrogen fixation: nitrogenases, relevant chemical models, and commercial processes* (eds BE Smith, RL Richards, WE Newton), pp. 33–54. Dordrecht, The Netherlands: Kluwer Academic Publishers.
59. Kandemir T, Schuster ME, Senyshyn A, Behrens M, Schlögl R. 2013 The Haber–Bosch process revisited: on the real structure and stability of 'ammonia iron' under working conditions. *Angew. Chemie - Int. Ed.* **52**, 12 723–12 726. (doi:10.1002/anie.201305812)
60. Etiope G, Ionescu AM. 2015 Low-temperature catalytic CO_2 hydrogenation with geological quantities of ruthenium: a possible abiotic CH_4 source in chromitite-rich serpentinized rocks. *Geofluids* **15**, 438–452. (doi:10.1111/gfl.1210)
61. McCollom TM, Seewald JS. 2013 Serpentinites, hydrogen, and life. *Elements* **9**, 129–134. (doi:10.2113/gselements.9.2.129)
62. Smirnov A, Hausner D, Laffers R, Strongin DR, Schoonen MAA. 2008 Abiotic ammonium formation in the presence of Ni–Fe metals and alloys and its implications for the Hadean nitrogen cycle. *Geochem. Trans.* **9**, 5–20. (doi:doi:10.1186/1467-4866-9-5)
63. Porosoff MD, Yan B, Chen JG. 2016 Catalytic reduction of CO_2 by H_2 for synthesis of CO, methanol and hydrocarbons: challenges and opportunities. *Energy Environ. Sci.* **9**, 62–73. (doi:10.1039/c5ee02657a)
64. Muchowska KB, Varma SJ, Moran J. 2019 Synthesis and breakdown of universal metabolic precursors promoted by iron. *Nature* **569**, 104–107. (doi:10.1038/s41586-019-1151-1)
65. Stryer L. 1975 *Biochemistry*. San Francisco, CA: Freeman.
66. Ownby K, Xu H, White RH. 2005 A *Methanocaldococcus jannaschii* archaeal signature gene encodes for a 5-formaminoimidazole-4-carboxamide-1-β-D-ribofuranosyl 5′-monophosphate synthetase: a new enzyme in purine biosynthesis. *J. Biol. Chem.* **280**, 10 881–10 887. (doi:10.1074/jbc.M413937200)
67. Copley SD, Smith E, Morowitz HJ. 2007 The origin of the RNA world: co-evolution of genes and metabolism. *Bioorg. Chem.* **35**, 430–443. (doi:10.1016/j.bioorg.2007.08.001)
68. Wong JTF. 2005 Coevolution theory of genetic code at age thirty. *Bioessays* **27**, 416–425. (doi:10.1002/bies.20208)
69. Trifonov EN, Volkovich Z, Frenkel ZM. 2012 Multiple levels of meaning in DNA sequences, and one more. *Ann. NY Acad. Sci.* **1267**, 35–38. (doi:10.1111/j.1749-6632.2012.06589.x)
70. Kauffman SA. 1986 Autocatalytic sets of proteins. *J. Theor. Biol.* **119**, 1–24. (doi:10.1016/S0022-5193(86)80047-9)
71. Steel M, Hordijk W, Xavier JC. 2019 Autocatalytic networks in biology: structural theory and algorithms. *J. R. Soc. Interface* **16**, 20180808. (doi:10.1098/rsif.2018.0808)
72. Vaidya N, Lehman N. 2009 One RNA plays three roles to provide catalytic activity to a group I intron lacking an endogenous internal guide sequence. *Nucleic Acids Res.* **37**, 3981–3989. (doi:10.1093/nar/gkp271)
73. Hordijk W, Steel M. 2004 Detecting autocatalytic, self-sustaining sets in chemical reaction systems. *J. Theor. Biol.* **227**, 451–461. (doi:10.1016/j.jtbi.2003.11.020)
74. Schönheit P, Buckel W, Martin WF. 2016 On the origin of heterotrophy. *Trends Microbiol.* **24**, 12–25. (doi:10.1016/j.tim.2015.10.003)
75. Sousa FL, Hordijk W, Steel M, Martin WF. 2015 Autocatalytic sets in *E. coli* metabolism. *J. Syst. Chem.* **6**, 4. (doi:10.1186/s13322-015-0009-7)
76. Xavier JC, Hordijk W, Kauffman SA, Steel M, Martin WF. 2019 Autocatalytic chemical networks preceded proteins and RNA in evolution. *BioRxiv* (doi:10.1101/693879)
77. Sousa FL, Preiner M, Martin WF. 2018 Native metals, electron bifurcation, and CO_2 reduction in early biochemical evolution. *Curr. Opin. Microbiol.* **43**, 77–83. (doi:10.1016/j.mib.2017.12.010)
78. Kitani A, Tsunetsugu S, Sasaki K. 1991 Fe^{III}-ion-catalysed non-enzymatic transformation of ADP into ATP. *J. Chem. Soc. Perkin Trans. 2* **3**, 329–331. (doi:10.1039/P29910000329)
79. Kitani A, Tsunetsugu S, Suzuki A, Ito S, Sasaki K. 1995 Fe(III)-ion-catalysed non-enzymatic transformation of adenosine diphosphate into adenosine triphosphate part II. Evidence of catalytic nature of Fe ions. *Bioelectrochem. Bioenerg.* **36**, 47–51. (doi:10.1016/0302-4598(94)01751-L)
80. Martin WF, Thauer RK. 2017 Energy in ancient metabolism. *Cell* **168**, 953–955. (doi:10.1016/j.cell.2017.02.032)
81. Liu Y, Beer LL, Whitman WB. 2012 Sulfur metabolism in archaea reveals novel processes. *Environ. Microbiol.* **14**, 2632–2644. (doi:10.1111/j.1462-2920.2012.02783.x)
82. Kelley DS, Baross JA, Delaney JR. 2002 Volcanoes, fluids, and life at mid-ocean ridge spreading centers. *Annu. Rev. Earth Planet. Sci.* **30**, 385–491. (doi:10.1146/annurev.earth.30.091201.141331)
83. Pasek MA. 2008 Rethinking early Earth phosphorus geochemistry. *Proc. Natl Acad. Sci. USA* **105**, 853–858. (doi:10.1073/pnas.0708205105)
84. Konhauser KO. 2007 Was there really an Archean phosphate crisis. *Science* **315**, 1234. (doi:10.1126/science.1136328)

Concluding remarks

It is still hard to believe how many researchers, outside and inside OoL related research were and are working on the ins and outs of CO_2 fixation (Guan *et al.*, 2003; He *et al.*, 2010; Klein and McCollom, 2013; McCollom, 2016; Varma *et al.*, 2018; Preiner, Igarashi, *et al.*, 2019). CO_2 fixation seems like a simple enough task, especially when high yield is not the main goal, merely detecting the actual products is sufficient. But the task is not simple at all. Often, the electron donors chosen are sulfide minerals, which on their own (i.e. without electrochemical gradients) cannot reduce CO_2, mainly because they are only capable of one-electron reduction which, as explained, is an unfavourable reduction step for CO_2 (see Chapter 2.2). Zero valent metal compounds were tested by us and by our collaborators as electron donors. Although CO_2 was reduced this way, with time it became clear that H_2 is the most efficient electron donor provided the right catalysts are around. There is good reason to believe that H_2 is also formed in the zero valent metals and acts as the actual reducing agent.

The results presented in 'A hydrogen dependent geochemical analogue of primordial carbon and energy metabolism' (Publication 7) show that abiotic and biotic H_2 dependent CO_2 fixation under mild hydrothermal conditions apparently follow very similar routes. This alone cannot lead to the conclusion that the mechanisms are also similar, especially because one is heterogeneous and the other enzymatic. But heterogeneous hydrogenation mechanisms (see Fig. 4 in 2.3) indicates that the principles of hydrogenation are very similar either way, for instance as in the enzymatic mechanisms of formate dehydrogenase (Maia *et al.*, 2016). The actual mechanism in our reactions, which are not taking place in a dry gas phase as with large-scale hydrogenations, but in water, remains to be identified.

This brings up another issue that needs to be addressed in the future. Indeed, most prebiotic CO_2 fixation experiments are performed in aqueous solution which is of course appropriate for all homogeneous catalytic experiments, but can block heterogeneous catalysts. From industrial processes it is known that water, which is often a side product of hydrogenation, poisons the catalysts through hydroxyl formation on their surfaces, thus blocking the prospective catalysis sites (Porosoff, Yan and Chen, 2016). This would also explain why the yields of all aqueous CO_2/H_2 experiments are generally very low in comparison to those of purposive chemical syntheses. Especially under our alkaline conditions, hydroxyl group formation can occur *in situ* and would not easily be visible in the XRD spectra. On the other hand, studies with water vapour in hydrogenation processes have shown that H_2O, although lowering the output of industrially relevant products including methane and larger hydrocarbons, increases the percentage of C1–C4 'oxygenates'—oxidized carbon

compounds—possibly including organic acids central in metabolism (Satterfield *et al.*, 1986). Water poisoning might thus be a problem, but a manageable one.

Furthermore, it could be that the water activity in serpentinizing systems is low, due to water–rock interactions. Water is either bound as hydroxides or reduced to hydrogen via serpentinization (Frost and Beard, 2007). We started working with clays and zeolites to meet such conditions in new experiments. Decreasing water activity in the system might also promote the formation of more complex carbon compounds, especially when other elements such as nitrogen are added to the reaction. The water-tolerant catalysts that are developed for industrial purposes are metal clusters encapsulated in carbon (Liu *et al.*, 2003) or in porous SiO_2 shells (Qiao *et al.*, 2014). Both forms come close to a porous serpentinizing system.

What is missing in the realm of OoL theories built around serpentinizing systems is an equivalent of the Miller–Urey experiment: a one-pot synthesis of amino acids, nucleic bases and cofactors from simple gasses. This thesis has identified an avenue for future pursuit.

Our CO_2 fixation results are robust to different catalysts and conditions, opening up the possibility that any serpentinizing system—also outside of deep-sea hydrothermal vents—could be a site for the early steps of abiogenesis. Besides mid-ocean ridges, serpentinization of ultramafic rocks also occurs in forearc regions between an ocean trench and the associated volcanic arc or terrestrial ophiolites (*i.e.* obducted/accretionary oceanic crust) (Holm *et al.*, 2015). The conditions found in such regions should be investigated in an OoL context.

What became evident working on this project, in which bioinformatics, network specialists, microbiologists, surface and organic chemists exchanged much information, is the value of interdisciplinarity in order to form a deeper understanding of the mechanisms in both biology and (geo)chemistry. Biology needs this knowledge to find out how the transition from abiotic to biotic took place 4 billion years ago, and also to understand what happens at geochemical sites mechanistically. Chemistry can improve the desperately needed process of CO_2 fixation by learning from nature's very long experience with the issue (Appel *et al.*, 2013). One main problem for industrial CO_2 fixation is the lack of pure H_2 gas. Most technologically used H_2 comes from the compounds that are meant to be produced by it: hydrocarbons. Better means of H_2 synthesis can improve CO_2 fixation and potentially decrease CO_2 emissions. Serpentinization solved this problem for geochemical settings and might have been an important prerequisite for life as it would have been able to provide the electron and energy source that is H_2. Geochemical processes could also inspire industrial hydrogen production.

References

Álvarez, A. *et al.* (2017) 'CO_2 activation over catalytic surfaces', *ChemPhysChem*, 18, pp. 3135–3141. doi: 10.1002/cphc.201700782.

Appel, A. M. *et al.* (2013) 'Frontiers, opportunities, and challenges in biochemical and chemical catalysis of CO_2 fixation', *Chemical Reviews*, 113(8), pp. 6621–6658. doi: 10.1021/cr300463y.

Bada, J. L. and Lazcano, A. (2003) 'Prebiotic Soup—Revisiting the Miller Experiment', *Science*, 300, pp. 745–746. doi: 10.1126/science.1085145.

Baross, J. A. and Hoffman, S. E. (1985) 'Submarine hydrothermal vents and associated gradient environments as sites for the origin and evolution of life', *Origins of Life and Evolution of the Biosphere*, 15, pp. 327–345. doi: 10.1007/BF01808177.

Bassham, J. A. (1979) 'The reductive pentose phosphate cycle and its regulation', in Gibbs, M. and Latzko, E. (eds) *Photosynthesis II. Encyclopedia of Plant Physiology (New Series), vol 6.* Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 9–30. doi: https://doi.org/10.1007/978-3-642-67242-2_2.

Behrens, M. *et al.* (2012) 'The active site of methanol synthesis over $Cu/ZnO/Al_2O_3$ industrial catalysts', *Science*, 336(6083), pp. 893–898. doi: 10.1126/science.1219831.

Bell, A. T., Gates, B. C. and Ray, D. (2007) *Basic research needs: Catalysis for Energy*. Bethesda, Maryland. doi: 10.2172/927492.

Berg, I. A. *et al.* (2007) 'A 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway in archaea', *Science*, 318(5857), pp. 1782–1786. doi: 10.1126/science.1149976.

Berg, I. A. *et al.* (2010) 'Autotrophic carbon fixation in archaea', *Nature Reviews Microbiology*, 8(6), pp. 447–460. doi: 10.1038/nrmicro2365.

Böller, B., Durner, K. M. and Wintterlin, J. (2019) 'The active sites of a working Fischer–Tropsch catalyst revealed by operando scanning tunnelling microscopy', *Nature Catalysis*, 2, pp. 1027–1034. doi: 10.1038/s41929-019-0360-1.

Buckingham, A. D. and Disch, R. L. (1963) 'The Quadrupole Moment of the Carbon Dioxide Molecule', *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 273, pp. 275–289. doi: 10.1098/rspa.1963.0088.

Bungenberg de Jong, H. G. (1932) 'Die Koazervation und ihre Bedeutung für die Biologie', *Protoplasm*, 15, pp. 110–173.

Butlerow, A. (1861) 'Bildung einer zuckerartigen Substanz durch Synthese', *Justus Liebigs Annalen der Chemie*, 120(3), pp. 295–298.

Camprubi, E. *et al.* (2017) 'Iron catalysis at the origin of life', *IUBMB Life*, 69(6), pp. 373–381. doi: 10.1002/iub.1632.

Can, M. *et al.* (2017) 'Bond in the CO-treated form of acetyl-CoA synthase'. doi: 10.1021/acs.biochem.6b00983.

Chabrière, E. *et al.* (1999) 'Crystal structures of the key anaerobic enzyme pyruvate ferredoxin oxidoreductase free and in complex with pyruvate', *Nature Structural Biology*, 6(2), pp. 182–190. doi: 10.1038/5870.

Chalk, A. J. and Halpern, J. (1959) 'Medium effects in the homogeneous catalytic activation of molecular hydrogen by metal salts. I. Cupric and cuprous heptanoates in heptanoic acid', *Physical and Inorganic Chemistry*, 81(22), pp. 5846–5852. doi: 10.1021/ja01531a002.

Chen, Q. W. and Bahnemann, D. W. (2000) 'Reduction of carbon dioxide by magnetite: Implications for the primordial synthesis of organic molecules', *Journal of American Chemical Society*, 122(10), pp. 970–971. doi: 10.1021/ja991278y.

Choi, H.-S. and Suh, M. P. (2009) 'Highly selective CO2 capture in flexible 3D coordination polymer', *Angewandte Chemie - International Edition*, 48, pp. 6865–6869. doi: 10.1002/anie.200902836.

Corliss, J. B. *et al.* (1979) 'Submarine thermal springs on the Galapágos rift', 203(4385), pp. 1073–1083.

Corliss, J. B., Baross, J. A. and Hoffman, S. E. (1981) 'An hypothesis concerning the relationships between submarine hot springs and the origin of life on earth', in *Oceanologica Acta, Special issue*, pp. 59–69.

Cornell, R. and Schwertmann, U. (2003) 'Formation', in *The iron oxides: Structure, properties, reactions, occurences and uses, second edition*. 2nd edn. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA, pp. 345–364. doi: 10.1002/3527602097.

Crabtree, R. H. (1997) 'Where smokers rule', *Science*, 276, p. 2515. doi: 10.1126/science.276.5310.222.

Cvjetko Bubalo, M. *et al.* (2015) 'Green solvents for green technologies', *Journal of Chemical Technology & Biotechnology,* 90(9), pp. 1631–1639. doi: 10.1002/jctb.4668.

D'Alessandro, D. M., Smit, B. and Long, J. R. (2010) 'Carbon dioxide capture carbon dioxide capture: prospects for new materials', *Angewandte Chemie - International Edition*, 49, pp. 6058–6082. doi: 10.1002/anie.201000431.

Darwin, C. (1859) *On the origin of species by means of natural selection*. Edited by D. and C. Appelton. New York.

Darwin, C. (1871) *Correspondence Project*, *Letter No. 7471*. Available at: https://www.darwinproject.ac.uk/letter/DCP-LETT-7471.xml.

Dayhoff, M. O. and Eck, R. V. (1966) 'Evolution of the structure of ferredoxin based on surviving relics of primitive amino acid sequences', *Science*, 152(3720), pp. 363–366. doi: 10.1126/science.152.3720.363.

Deamer, D. W. (1997) 'The first living systems: a bioenergetic perspective', *Microbiology and Molecular Biology Reviews*, 61(2), pp. 239–261.

Decker, K., Jungermann, K. and Thauer, R. K. (1970) 'Energy production in anaerobic organisms', *Angewandte Chemie International Edition*, 9(2), pp. 138–158.

Dobbek, H. *et al.* (2001) 'Crystal structure of a carbon monoxide dehydrogenase reveals a [Ni-4Fe-5S] cluster', *Science*, 293(5533), pp. 1281–1285. doi: 10.1126/science.1061500.

de Duve, C. (2003) 'A research proposal on the origin of life. Closing Lecture given at the ISSOL Congress in 727 Oaxaca, Mexico, on July 4, 2002.', *Origins of Life and Evolution of the Biosphere*, 33(6), pp. 559–574. doi: 10.1023/a:1025760311436.

Dyson, F. J. (1999) *Origins of Life*. Cambridge: Cambridge University Press.

Ertl, G. (1983) 'Primary steps in catalytic synthesis of ammonia', *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, 1(2), pp. 1247–1253. doi: 10.1116/1.572299.

Etiope, G. and Ionescu, A. M. (2015) 'Low-temperature catalytic CO_2 hydrogenation with geological quantities of ruthenium: a possible abiotic CH_4 source in chromitite-rich serpentinized rocks', *Geofluids*, 15, pp. 438–452. doi: 10.1111/gfl.1210.

Etiope, G. and Schoell, M. (2014) 'Abiotic gas: atypical, but not rare', *Elements*, 10, pp. 291–296. doi: 10.2113/gselements.10.4.291.

Farley, J. (1978) 'Philosophical and historical aspects of the origin of life', *Treballs de la Societat Catalana de Biologia*, 39, pp. 37–47.

Fontecilla-Camps, J. C. and Ragsdale, S. W. (1999) 'Nickel-iron-sulfur active sites: Hydrogenase and CO dehydrogenase.', *Advances in Inorganic Chemistry*, 47(2), pp. 283–333. doi: 10.1016/S0898-8838(08)60081-1.

Forster, D. (1979) 'Mechanistic pathways in the catalytic carbonylation by rhodium and iridium Complexes', *Advances in Organometallic Chemistry*, 17, pp. 255–267. doi: 10.1016/s0065-3055(08)60325-1.

Frese, K. W. (1993) 'Electrochemical reduction of CO_2 at solid electrodes', in Sullivan, B. P., Krist, K., and Guard, H. E. (eds) *Electrochemical and electrocatalytic reactions of carbon dioxide*. New York: Elsevier, pp. 145–216.

Friedrich, B., Fritsch, J. and Lenz, O. (2011) 'Oxygen-tolerant hydrogenases in hydrogen-based technologies', *Current Opinion in Biotechnology*, 22(3), pp. 358–364. doi: 10.1016/j.copbio.2011.01.006.

Frost, R. B. and Beard, J. S. (2007) 'On silica activity and serpentinization', *Journal of Petrology*, 48(7), pp. 1351–1368. doi: 10.1093/petrology/egm021.

Früh-Green, G. L. *et al.* (2017) 'Northern sites', *Proceedings of the International Ocean Discovery Program*, 357. doi: 10.14379/iodp.proc.357.106.2017.

Fuchs, G. (1986) 'CO_2 fixation in acetogenic bacteria: variations on a theme', *FEMS Microbiology Letters*, 39(3), pp. 181–213. doi: 10.1111/j.1574-6968.1986.tb01859.x.

Fuchs, G. (2011) 'Alternative pathways of carbon dioxide fixation: Insights into the early evolution of life?', *Annual Review of Microbiology*, 65(1), pp. 631–658. doi: 10.1146/annurev-micro-090110-102801.

Gilbert, W. (1986) 'The RNA world', *Nature*, 319, p. 618. doi: 10.1038/319618a0.

Guan, G. *et al.* (2003) 'Reduction of aqueous CO_2 at ambient temperature using zero-valent iron-based composites', *Green Chemistry*, 5(5), p. 630. doi: 10.1039/b304395a.

Haldane, J. B. S. (1929) 'The origin of life', *Rationalist Annual*, pp. 3–10.

Halpern, J. (1980) 'Homogeneous catalytic hydrogenation: a retrospective account', *Journal of Organometallic Chemistry*, 200, pp. 133–144. doi: 10.1016/S0022-328X(00)88642-8.

He, C. *et al.* (2010) 'A mild hydrothermal route to fix carbon dioxide to simple carboxylic acids', *Organic Letters*, 12(4), pp. 649–651. doi: 10.1021/ol9025414.

Holm, N. G. *et al.* (2015) 'Serpentinization and the formation of H_2 and CH_4 on celestial bodies (planets, moons, comets)', *Astrobiology*, 15(7), pp. 587–600. doi: 10.1089/ast.2014.1188.

Holm, N. G. and Charlou, J. L. (2001) 'Initial indications of abiotic formation of hydrocarbons in the Rainbow ultramafic hydrothermal system, Mid-Atlantic Ridge', *Earth*, 191, pp. 1–8. doi: 10.1016/s0012-821x(01)00397-1.

Horita, J. and Berndt, M. E. (1999) 'Abiogenic methane formation and isotopic fractionation under hydrothermal conditions', *Scientific Reports*, 285, pp. 1055–1057. doi: 10.1126/science.285.5430.1055.

Hou, B. *et al.* (2012) 'Study on kinetics of iron oxide reduction by hydrogen', *Chinese Journal of Chemical Engineering*, 20(1), pp. 10–17. doi: 10.1016/S1004-9541(12)60357-7.

Huber, C. and Wächtershäuser, G. (1997) 'Activated acetic acid by carbon fixation on (Fe,Ni)S under primordial conditions', *Science*, 276(5310), pp. 245–248. doi: 10.1126/science.276.5310.245.

Huxley, T. H. (1869) 'On the physical basis of life', *Fortnightly Review*, 5, pp. 129–145.

Jeoung, J.-H. and Dobbek, H. (2007) 'Carbon dioxide activation at the Ni,Fe-cluster of anaerobic carbon monoxide dehydrogenase', *Conservation & Exhibitions*, 318, pp. 1461–1464. doi: 10.1016/b978-0-408-01434-2.50020-6.

Jia, T. Z. *et al.* (2019) 'Membraneless polyester microdroplets as primordial compartments at the origins of life', *Proceedings of the National Academy of Science, U.S.A.*, 116(32), pp. 15830–15835. doi: 10.1073/pnas.1902336116.

Jordan, S. F. *et al.* (2019) 'Promotion of protocell self-assembly from mixed amphiphiles at the origin of life', *Nature Ecology and Evolution*, 3, pp. 1705–1714. doi: 10.1038/s41559-019-1015-y.

Kamminga, H. (1988) 'Historical perspective: the problem of the origin of life in the context of developments in biology', *Origins of Life and Evolution of Biospheres*, 18, pp. 1–11.

Kandemir, T. *et al.* (2013) 'The Haber-Bosch process revisited: on the real structure and stability of "ammonia iron" under working conditions', *Angewandte Chemie - International Edition*, 52(48), pp. 12723–12726. doi: 10.1002/anie.201305812.

Karl, T. R. and Trenberth, K. E. (2003) 'Modern global climate change', *Science*, 302, pp. 1719–1723. doi: 10.1126/science.1090228.

Keene, F. R. (1993) 'Thermodynamic, kinetic, and product considerations in carbon dioxide reactivity', in Sullivan, B. P., Krist, K., and Guard, H. E. (eds) *Electrochemical and electrocatalytic reactions of carbon dioxide*. New York: Elsevier, pp. 1–18.

Kelley, D. S. *et al.* (2001) 'An off-axis hydrothermal vent field near the Mid-Atlantic Ridge at 30 degrees N.', *Nature*, 412(6843), pp. 145–149. doi: 10.1038/35084000.

Kelley, D. S., Baross, J. A. and Delaney, J. R. (2002) 'Volcanoes, fluids, and life at mid-ocean ridge spreading centers', *Annual Review of Earth and Planetary Sciences.*, 30, pp. 385–491. doi: 10.1146/annurev.earth.30.091201.141331.

Kelley, D. S. and Früh-Green, G. L. (1999) 'Abiogenic methane in deep-seated mid-ocean ridge environments: Insights from stable isotope analyses', *Journal of Geophysical Research*, 104(B5), pp. 10439–10460. doi: doi.org/10.1029/1999JB900058.

Kelly, T. D. and Matos, G. R. (2015) *Historical statistics for mineral and material commodities in the united states*, *U.S. Geological Survey (USGS)*. Available at: https://www.usgs.gov/centers/nmic/historical-statistics-mineral-and-material-commodities-united-states.

Kiehl, J. T. and Trenberth, K. E. (1997) 'Earth's annual global mean energy budget', *Bulletin of the American Meteorological Society*, 78, pp. 197–208. doi: 10.1175/1520-0477(1997)078<0197:EAGMEB>2.0.CO;2.

Klein, F. and Bach, W. (2009) 'Fe-Ni-Co-O-S phase relations in peridotite-seawater interactions', *Journal of Petrology*, 50(1), pp. 37–59. doi: 10.1093/petrology/egn071.

Klein, F., Bach, W. and McCollom, T. M. (2013) 'Compositional controls on hydrogen generation during serpentinization of ultramafic rocks', *Lithos*, 178, pp. 55–69. doi: 10.1016/j.lithos.2013.03.008.

Klein, F. and McCollom, T. M. (2013) 'From serpentinization to carbonation: New insights from a CO_2 injection experiment', *Earth and Planetary Science Letters*, 379, pp. 137–145. doi: 10.1016/j.epsl.2013.08.017.

Konn, C. *et al.* (2015) 'The production of methane, hydrogen, and organic compounds in ultramafic-hosted hydrothermal vents of the Mid-Atlantic Ridge', *Astrobiology*, 15(5), pp. 381–399. doi: 10.1089/ast.2014.1198.

Kruger, K. *et al.* (1982) 'Self-Splicing RNA : Autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena', *Cell*, 31, pp. 147–157.

Kubas, G. J. *et al.* (1984) 'Characterization of the first examples of isolable molecular hydrogen complexes, $M(CO)_3(PR_3)_2(H_2)$ (M = Mo, W; R = Cy, i-Pr). Evidence for a side-on bonded H_2 ligand', *Journal of American Chemical Society*, 106, pp. 451–452.

Kubas, G. J. (2007) 'Fundamentals of H_2 binding and reactivity on transition metals underlying hydrogenase function and H_2 production and storage', *Chemical Reviews*, 107(10), pp. 4152–4205. doi: 10.1021/cr050197j.

Lang, S. Q. *et al.* (2010) 'Elevated concentrations of formate, acetate and dissolved organic carbon found at the Lost City hydrothermal field', *Geochimica et Cosmochimica Acta*, 74(3), pp. 941–952. doi: 10.1016/j.gca.2009.10.045.

Leigh, G. J. (2004) 'Haber– Bosch and other industrial processes', in Smith, B. E., Richards, R. L., and Newton, W. E. (eds) *Catalysts for nitrogen fixation: nitrogenases, relevant chemical models, and commercial processes*. Kluwer Academic Publishers, pp. 33–54. doi: 10.1007/978-1-4020-3611-8_2.

Li, P. *et al.* (2016) 'Challenges and perspectives on carbon fixation and utilization technologies: an overview', *Aerosol and Air Quality Research*, 16, pp. 1327–1344. doi: 10.4209/aaqr.2015.12.0698.

Liu, X. M. *et al.* (2003) 'Recent advances in catalysts for methanol synthesis via hydrogenation of CO and CO_2', *Industrial and Engineering Chemistry Research*, 42(25), pp. 6518–6530. doi: 10.1021/ie020979s.

Lubitz, W. *et al.* (2014) 'Hydrogenases', *Chemical Reviews*, 114, pp. 4081–4148. doi: 10.1021/cr4005814.

Maden, B. E. H. (2000) 'Tetrahydrofolate and tetrahydromethanopterin compared: functionally distinct carriers in C_1 metabolism', *Biochemical Journal*, 350(3), pp. 609–629. doi: 10.1042/0264-6021:3520935.

Maia, L. B. *et al.* (2016) 'Reduction of carbon dioxide by a molybdenum-containing formate dehydrogenase: a kinetic and mechanistic study', *Journal of the American Chemical Society*, 138(28), pp. 8834–8846. doi: 10.1021/jacs.6b03941.

Martin, W. F. (2011) 'Early evolution without a tree of life', *Biology Direct*, 6(1), 36. doi: 10.1186/1745-6150-6-36.

Martin, W. F. (2019) 'Carbon–metal bonds: rare and primordial in metabolism', *Trends in Biochemical Sciences*, 44, pp. 807–818. doi: 10.1016/j.tibs.2019.04.010.

Martin, W. F., Bryant, D. A. and Beatty, J. T. (2018) 'A physiological perspective on the origin and evolution of photosynthesis', *FEMS Microbiology Reviews*, 42, pp. 205–231. doi: 10.1093/femsre/fux056.

Martin, W. F. and Müller, M. (1998) 'The hydrogen hypothesis for the first eukaryote', *Nature*, 392, pp. 37–41. doi: 10.1088/1555-6611/aad92d.

Martin, W. F. and Preiner, M. (2017) 'Origin of Life, Theories of', *Reference Module in Life Sciences*. Elsevier Inc., Oxford. doi: 10.1016/B978-0-12-809633-8.02403-1.

Martin, W. F. and Russell, M. J. (2003) 'On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 358(1429), pp. 59–85. doi: 10.1098/rstb.2002.1183.

Martin, W. F. and Russell, M. J. (2007) 'On the origin of biochemistry at an alkaline hydrothermal vent', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1486), pp. 1887–1925. doi: 10.1098/rstb.2006.1881.

McCollom, T. M. (2013a) 'Laboratory simulations of abiotic hydrocarbon formation in Earth's deep subsurface', *Reviews in Mineralogy and Geochemistry*, 75(1), pp. 467–494. doi: 10.2138/rmg.2013.75.15.

McCollom, T. M. (2013b) 'Miller-Urey and beyond: What have we learned about prebiotic organic synthesis reactions in the past 60 years?', *Annual Review of Earth and Planetary Sciences*, 41, pp. 207–29. doi: 10.1146/annurev-earth-040610-133457.

McCollom, T. M. (2016) 'Abiotic methane formation during experimental serpentinization of olivine.', *Proceedings of the National Academy of Sciences, U.S.A.*, 113(49), pp. 13965–13970. doi: 10.1073/pnas.1611843113.

McCollom, T. M. *et al.* (2016) 'Temperature trends for reaction rates, hydrogen generation, and partitioning of iron during experimental serpentinization of olivine', *Geochimica et Cosmochimica Acta*, 181, pp. 175–200. doi: 10.1016/j.gca.2016.03.002.

McDermott, J. M. *et al.* (2015) 'Pathways for abiotic organic synthesis at submarine hydrothermal fields', *Proceedings of the National Academy of Sciences, U.S.A.*, 112(25), pp. 7668–7672. doi: 10.1073/pnas.1506295112.

Ménez, B. *et al.* (2018) 'Abiotic synthesis of amino acids in the recesses of the oceanic lithosphere', *Nature*, 564(7734), pp. 59–63. doi: 10.1038/s41586-018-0684-z.

Miller, H. M. *et al.* (2017) 'Low temperature hydrogen production during experimental hydration of partially-serpentinized dunite', *Geochimica et Cosmochimica Acta*, 209, pp. 161–183. doi: 10.1016/j.gca.2017.04.022.

Miller, S. L. (1953) 'A production of amino acids under possible primitive earth conditions', *Science*, 117(3046), pp. 528–529.

Miller, S. L. (1955) 'Production of some organic compounds under possible primitive earth conditions', *Journal of the American Chemical Society*, 77(9), pp. 2351–2361.

Milshteyn, D. *et al.* (2018) 'Amphiphilic compounds assemble into membranous vesicles in hydrothermal hot spring water but not in seawater', *Life*, 8(2), 11. doi: 10.3390/life8020011.

Moore, B. (1913) *The origin and nature of life*. Edited by H. Holt. New York.

Morowitz, H. and Smith, E. (2006) 'Energy flow and the organisation of life', *Complexity*, 13, pp. 51–59. doi: 10.1002/cplx.20191.

Müller, V., Chowdhury, N. P. and Basen, M. (2018) 'Electron bifurcation: A long-hidden energy-coupling mechanism', *Annual Review of Microbiology*, 72, pp. 331–353. doi: 10.1146/annurev-micro-090816- 093440.

Oivanen, M., Kuusela, S. and Lönnberg, H. (1998) 'Kinetics and mechanisms for the cleavage and isomerization of the phosphodiester bonds of RNA by Brønsted acids and bases', *Chemical Reviews*, 98, pp. 961–990. doi: 10.1021/cr960425x.

Oparin, A. (1938) *The origin of life*. New York: Dover Publications, Inc.

Oparin, A. (1957) *Die Entstehung des Lebens auf der Erde*. Edited by T. Creutziger. Berlin: Deutscher Verlag der Wissenschaften.

Oparin, A. (1967) 'The origin of life, transl. A. Synge', in Bernal, J. D. (ed.) *The Origin of Life*. London: Oxford University Press, pp. 197–234.

Orgel, L. E. (2004) 'Prebiotic Chemistry and the Origin of the RNA World Prebiotic Chemistry and the Origin of the RNA World', *Critical Reviews in Biochemistry and Molecular Biology*, 39(2), pp. 99–123. doi: 10.1080/10409230490460765.

Orgel, L. E. (2008) 'The implausibility of metabolic cycles on the prebiotic Earth', *PLoS Biology*, 6(1), e18-9. doi: 10.1371/journal.pbio.0060018.

Ostwald, W. and Köhler, R. (1927) 'Ueber die flüssig-flüssige Entmischung von Gelatine durch Sulfosalizylsäure und über die Beziehungen dieses Systems zur Phasenregel', *Kolloid-Zeitschrift*, 43, pp. 131–150.

Pachauri, R. K. and Meyer, L. A. (eds) (2014) *Climate change 2014: Synthesis report. Contribution of working groups I, II and III to the fifth assessment report of the intergovernmental panel on climate change*, *IPCC, 2014*. Geneva, Switzerland. doi: 10.1016/S0022-0248(00)00575-3.

Pisarev, A. A. (2012) 'Hydrogen adsorption on the surface of metals', in Gangloff, R. P. and Somerday, B. P. (eds) *Gaseous Hydrogen Embrittlement of Materials in Energy Technologies Part 1*. Woodhead Publishing Limited, pp. 3–26. doi: 10.1533/9780857095374.1.3.

Porosoff, M. D. *et al.* (2015) 'Identifying trends and descriptors for selective CO_2 conversion to CO over transition metal carbides', *Chem. Commun.*, 51, pp. 6988–6991. doi: 10.1039/C5CC01545F.

Porosoff, M. D., Yan, B. and Chen, J. G. (2016) 'Catalytic reduction of CO_2 by H_2 for synthesis of CO, methanol and hydrocarbons: Challenges and opportunities', *Energy and Environmental Science*, 9(1), pp. 62–73. doi: 10.1039/c5ee02657a.

Powner, M. W., Gerland, B. and Sutherland, J. D. (2009) 'Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions', *Nature*, 459(7244), pp. 239–242. doi: 10.1038/nature08013.

Preiner, M. *et al.* (2018) 'Serpentinization: connecting geochemistry, ancient metabolism and industrial hydrogenation', *Life*, 8(4), 41. doi: 10.3390/life8040041.

Preiner, M., Igarashi, K., *et al.* (2020) 'A hydrogen-dependent geochemical analogue of primordial carbon and energy metabolism', *Nature Ecology and Evolution*. doi: 10.1038/s41559-020-1125-6.

Preiner, M., Xavier, J. C., *et al.* (2019) 'Catalysts, autocatalysis and the origin of metabolism', *Interface Focus*, 9(6), 20190072. doi: 10.1098/rsfs.2019.0072.

Preiner, M., Asche, S., *et al.* (2020) 'The future of origin of life research: bridging decades-old divisions', *Life*, 10(3), 20. 10.3390/life10030020.

Qiao, Z. A. *et al.* (2014) 'Lab-in-a-shell: encapsulating metal clusters for size sieving catalysis', *Journal of the American Chemical Society*, 136(32), pp. 11260–11263. doi: 10.1021/ja505903r.

Ragsdale, S. W. and Kumar, M. (1996) 'Nickel-containing carbon monoxide dehydrogenase/acetyl-CoA synthase', *Chemical Reviews*, 96(7), pp. 2515–2540. doi: 10.1021/cr950058+.

Ragsdale, S. W. and Pierce, E. (2008) 'Acetogenesis and the Wood-Ljungdahl pathway of CO_2 fixation', *Biochimica et Biophysica Acta*, 1784, pp. 1873–1898. doi: 10.1016/j.bbapap.2008.08.012.

Riedel, T. *et al.* (2001) 'Kinetics of CO_2 hydrogenation on a K-promoted Fe catalyst', *Industrial & Engineering Chemistry Research*, 40, pp. 1355–1363. doi: 10.1021/ie000084k.

Riedel, T. *et al.* (2003) 'Fischer–Tropsch on iron with H_2/CO and H_2/CO_2 as synthesis gases: The episodes of formation of the Fischer-Tropsch regime and construction of the catalyst', *Topics in Catalysis*, 26(1–4), pp. 41–54. doi: 10.1023/B:TOCA.0000012986.46680.28.

Rodriguez, J. A. *et al.* (2015) 'Hydrogenation of CO_2 to methanol: Importance of metal-oxide and metal-carbide interfaces in the activation of CO_2', *ACS Catalysis*, 5, pp. 6696–6706. doi: 10.1021/acscatal.5b01755.

Roldan, A. *et al.* (2015) 'Bio-inspired CO_2 conversion by iron sulfide catalysts under sustainable conditions', *Chem. Commun.*, 51, pp. 7501–7504.

Ruiz-Mirazo, K., Briones, C. and De La Escosura, A. (2014) 'Prebiotic systems chemistry: New perspectives for the origins of life', *Chemical Reviews*, 114(1), pp. 285–366. doi: 10.1021/cr2004844.

Russell, M. J. *et al.* (1993) 'On the emergence of life via catalytic iron-sulphide membranes',

Terra Nova, 5, pp. 343–347. doi: 10.1111/j.1365-3121.1993.tb00267.x.

Russell, M. J. (2007) 'The alkaline solution to the emergence of life: Energy, entropy and early evolution', *Acta Biotheoretica*, 55(2), pp. 133–179. doi: 10.1007/s10441-007-9018-5.

Russell, M. J. (2019) 'Prospecting for life', *Interface Focus*, 9(6), 20190050. doi: 10.1098/rsfs.2019.0050.

Russell, M. J. and Hall, A. J. (1997) 'The emergence of life from iron monosulphide bubbles at a submarine hydrothermal redox and pH front', *Journal of the Geological Society*, 154, pp. 377–402. doi: 10.1.1.520.7821.

Russell, M. J., Hall, A. J. and Martin, W. F. (2010) 'Serpentinization as a source of energy at the origin of life', *Geobiology*, 8(5), pp. 355–371. doi: 10.1111/j.1472-4669.2010.00249.x.

Russell, M. J., Hall, A. J. and Turner, D. (1989) 'In vitro growth of iron sulphide chimneys: possible culture chambers for origin-of-life experiments', *Terra Nova*, 1(3), pp. 238–241. doi: 10.1111/j.1365-3121.1989.tb00364.x.

Russell, M. J. and Martin, W. F. (2004) 'The rocky roots of the acetyl-CoA pathway', *Trends in Biochemical Sciences*, 29(7), pp. 358–363. doi: 10.1016/j.tibs.2004.05.007.

Russell, M. J., Nitschke, W. and Branscomb, E. (2013) 'The inevitable journey to being', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1622), 20120254. doi: 10.1098/rstb.2012.0254.

Santos-Carballal, D. *et al.* (2018) 'Reactivity of CO_2 on the surfaces of magnetite (Fe_3O_4), greigite (Fe_3S_4) and mackinawite (FeS)', *Philosophical Transactions of the Royal Society A*, 376(2110), 20170065. doi: 10.1098/rsta.2017.0065.

Satterfield, C. N. *et al.* (1986) 'Effect of water on the iron-catalyzed Fischer–Tropsch synthesis', *Industrial and Engineering Chemistry Product Research and Development*, 25(3), pp. 407–414. doi: 10.1021/i300023a007.

Schönheit, P., Buckel, W. and Martin, W. F. (2016) 'On the origin of heterotrophy', *Trends in Microbiology*, 24(1), pp. 12–25. doi: 10.1016/j.tim.2015.10.003.

Schrenk, M. O., Brazelton, W. J. and Lang, S. Q. (2013) 'Serpentinization, carbon, and deep life', *Reviews in Mineralogy and Geochemistry*, 75, pp. 575–606. doi: 10.2138/rmg.2013.75.18.

Shapiro, R. (2000) 'A replicator was not involved in the origin of life', *IUBMB Life*, 49(3),

pp. 173–176. doi: 10.1080/713803621.

Shock, E. L. (1990) 'Geochemical contraints on the origin of organic compounds in hydrothermal systems', *Origins of Life and Evolution of the Biosphere*, 20, pp. 331–367. doi: 10.1007/BF01808115.

Siegbahn, P. E. M., Tye, J. W. and Hall, M. B. (2007) 'Computational studies of [NiFe] and [FeFe] hydrogenases', *Chemical Reviews*, 107(10), pp. 4414–4435. doi: 10.1021/cr050185y.

Sleep, N. H. *et al.* (2004) 'H_2-rich fluids from serpentinization: geochemical and biotic implications', *Proceedings of the National Academy of Sciences, U.S.A.*, 101(35), pp. 12818–12823. doi: 10.1073/pnas.0405289101.

Smith, E. and Morowitz, H. J. (2004) 'Universality in intermediary metabolism', *Proceedings of the National Academy of Science, U.S.A.*, 101(36), pp. 13168–13173. doi: 10.1073/pnas.0404922101.

Solomon, S. *et al.* (2007) *Climate change 2007: the physical Science basis*.

Sousa, F. L. and Martin, W. F. (2014) 'Biochemical fossils of the ancient transition from geoenergetics to bioenergetics in prokaryotic one carbon compound metabolism', *Biochimica et Biophysica Acta - Bioenergetics*, 1837(7), pp. 964–981. doi: 10.1016/j.bbabio.2014.02.001.

Sousa, F. L., Preiner, M. and Martin, W. F. (2018) 'Native metals, electron bifurcation, and CO_2 reduction in early biochemical evolution', *Current Opinion in Microbiology*, 43, pp. 77–83. doi: 10.1016/j.mib.2017.12.010.

Spreitzer, D. and Schenk, J. (2019) 'Reduction of iron oxides with hydrogen—a review', *Steel Research International*, 90(10), 1900108. doi: 10.1002/srin.201900108.

Strecker, A. (1850) 'Ueber die künstliche Bildung der Milchsäure und einen neuen, dem Glycocoll homologen Körper', *Justus Liebigs Annalen der Chemie*, 75, pp. 27–45.

Sunley, G. J. and Watson, D. J. (2000) 'High productivity methanol carbonylation catalysis using iridium. The CativaTM process for the manufacture of acetic acid', *Catalysis Today*, 58(4), pp. 293–307. doi: 10.1016/s0920-5861(00)00263-7.

Tashiro, T. *et al.* (2017) 'Early trace of life from 3.95 Ga sedimentary rocks in Labrador, Canada', *Nature*, 549, pp. 516–518. doi: 10.1038/nature24019.

Thauer, R. K. *et al.* (2008) 'Methanogenic archaea: ecologically relevant differences in energy conservation', *Nature Reviews Microbiology*, 6(8), pp. 579–591. doi:

10.1038/nrmicro1931.

Thauer, R. K. *et al.* (2010) 'Hydrogenases from methanogenic archaea, nickel, a novel cofactor, and H_2 storage', *Annual Review of Biochemistry*, 79(1), pp. 507–536. doi: 10.1146/annurev.biochem.030508.152103.

Thauer, R. K. (2011) 'Hydrogenases and the global H_2 cycle', *European Journal of Inorganic Chemistry*, 7, pp. 919–921. doi: 10.1002/ejic.201001255.

Thauer, R. K., Jungermann, K. and Decker, K. (1977) 'Energy conservation in chemotrophic anaerobic bacteria', *Bacteriological reviews*, 41(1), pp. 100–180. doi: 10.1073/pnas.0803850105.

Trends in atmospheric carbon dioxide (2019) *Global monthly mean CO_2*. Available at: https://www.esrl.noaa.gov/gmd/ccgg/trends/global.html#global.

Varma, S. J. *et al.* (2018) 'Native iron reduces CO_2 to intermediates and endproducts of the acetyl-CoA pathway', *Nature Ecology and Evolution*, 2(6), pp. 1019–1024. doi: 10.1038/s41559-018-0542-2.

Vincent, K. A. *et al.* (2005) 'Electrocatalytic hydrogen oxidation by an enzyme at high carbon monoxide or oxygen levels', *Proceedings of the National Academy of Science, U.S.A.*, 102(47), pp. 16951–16954. doi: 10.1073/pnas.0504499102.

Volbeda, A. *et al.* (1995) 'Crystal structure of the nickel-iron hydrogenase from *Desulfovibrio gigas*', *Nature*, 373(6515), pp. 580–587. doi: 10.1038/373580a0.

Wächtershäuser, G. (1988) 'Before enzymes and templates: Theory of surface metabolism', *Microbiological Reviews*, 52(4), pp. 452–484.

Wächtershäuser, G. (1990) 'Evolution of the first metabolic cycles.', *Proceedings of the National Academy of Sciences, U.S.A.*, 87(1), pp. 200–204. doi: 10.1073/pnas.87.1.200.

Wächtershäuser, G. (1992) 'Groundworks for an evolutionary biochemistry: The iron-sulphur world', *Progress in Biophysics and Molecular Biology*, 58(2), pp. 85–201. doi: 10.1016/0079-6107(92)90022-X.

Wächtershäuser, G. (1997) 'The origin of life and its methodological challenge', *Journal of Theoretical Biology*, 187, pp. 483–494. doi: 10.1006/jtbi.1996.0383.

Wang, W.-H., Feng, X. and Bao, M. (2018) 'Transformation of CO_2 to formic acid or formate with homogeneous catalysts', in *Transformation of carbon dioxide to formic acid and methanol. SpringerBriefs in molecular Science.* Singapore: Springer, pp. 7–42. doi: 10.1007/978-981-10-3250-9_2.

Weiss, M. C. *et al.* (2016) 'The physiology and habitat of the last universal common ancestor', *Nature Microbiology*, 1(9), 16116. doi: 10.1038/nmicrobiol.2016.116.

Weiss, M. C. *et al.* (2018) 'The last universal common ancestor between ancient Earth chemistry and the onset of genetics', *PloS Genetics*, 14(8), e1007518. doi: 10.1371/journal.pgen.1007518.

Welch, G. R. (1995) 'T. H. Huxley and the 'Protoplasmic theory of life ': 100 years later', *Trends in Biochemical Sciences*, 20(11), pp. 481–485. doi: 10.1016/S0968-0004(00)89106-9.

Wintterlin, J. *et al.* (1997) 'Atomic and macroscopic reaction rates of a surface-catalyzed reaction', *Science*, 278(5345), pp. 1931–1934. doi: 10.1126/science.278.5345.1931.

'Workshop OQOL'09: Open questions on the origins of life 2009' (2010) *Origins of Life and Evolution of Biospheres*, 40. doi: 10.1007/s11084-010-9213-2.

Xavier, J. C., Preiner, M. and Martin, W. F. (2018) 'Something special about CO-dependent CO_2 fixation', *FEBS Letters*, 285(22), pp. 4181–4195. doi: 10.1111/febs.14664.

Zambelli, T. *et al.* (1996) 'Identification of the "active sites" of a surface-catalyzed reaction', *Science*, 273, pp. 1688–1690. doi: 10.1126/science.273.5282.1688.