Phylogenetische Analyse der anaeroben Corrinsynthese und Glykolyse im Kontext früher metabolischer Evolution

Inaugural-Dissertation

zur Erlangung des Doktorgrades der

Mathematisch-Naturwissenschaftlichen Fakultät der

Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Luca David Modjewski

geboren in Wuppertal

Düsseldorf, September 2025

Aus dem Institut für Molekulare Evolution
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

Berichterstatter:
1. Prof. Dr. William F. Martin
2. Prof. Dr. Martin Lercher

Tag der mündlichen Prüfung: 19. Februar 2026

Eidesstattliche Erklärung

Hiermit versichere ich an Eides statt, dass diese Dissertation von mir selbstständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf" erstellt worden ist. Die Arbeit wurde bisher keiner Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht. Ich habe bisher keinen erfolglosen Promotionsversuch unternommen.

_____

Luca David Modjewski, Düsseldorf 2025

Für

Franzisca, Silke und Frank.

Im Laufe dieser Dissertation wurden mit Zustimmung des Betreuers folgende Beiträge publiziert:

Publikationen in Fachzeitschriften thematisiert in dieser Thesis

I **Luca D. Modjewski**, Val Karavaeva, Natalia Mrnjavac, Michael Knopp, William F. Martin, Filipa L. Sousa (2024). Evidence for corrin biosynthesis in the last universal common ancestor. *FEBS J.*, 10.1111/febs.17367.

II **Luca D. Modjewski**, Ulrike Johnsen, Peter Schönheit, William F. Martin (2025). The early evolution of the glycolytic pathway from autotrophic origins to glycogen and back. *FEMS Microbiol. Rev.*, 10.1093/femsre/fuaf049.

Weitere Publikationen in Fachzeitschriften

III Natalia Mrnjavac, Falk S. P. Nagies, Jessica L. Wimmer, Nils Kapust, Michael R. Knopp, Katharina Trost, **Luca D. Modjewski**, Nico Bremer, Marek Mentel, Mauro Degli Esposti, Itzhak Mizrahi, John F. Allen, William F. Martin (2024). The radical impact of oxygen on prokaryotic evolution-enzyme inhibition first, uninhibited essential biosyntheses second, aerobic respiration third**.** *FEBS Lett.* 598:1692–1714.

IV Bing Song, **Luca D. Modjewski**, Nils Kapust, Itzhak Mizrahi, William F. Martin (2022). The origin and distribution of the main oxygen sensing mechanism across metazoans. *Front Physiol.* 13:977391.

Inhaltsverzeichnis

1 Zusammenfassung

Autotrophe Theorien zur Entstehung des Lebens besagen, dass anzestraler Stoffwechsel auf Wasserstoff (H_2) und Kohlenstoffdioxid (CO_2) basiert. Diese anorganischen Komponenten bilden den Ausgangspunkt des Acetyl-CoA-Wegs, dem wohl ältesten Stoffwechselweg zur Fixierung von CO_2, welcher sowohl von H_2-abhängigen methanogenen Archaeen als auch acetogenen Bakterien genutzt wird. In diesem Weg der CO_2-Fixierung spielen Cobamide als Teil des Corrinoid-Eisen/Schwefel-Proteins (CoFeS) eine wichtige Rolle. Sie gehören zur Gruppe der Corrine, komplexe makrozyklische Verbindungen, die ein zentral gebundenes Cobalt-Ion enthalten und deren bekanntester Vertreter Cobalamin (Vitamin B_{12}) ist. Die Synthese von Cobalamin erfolgt entweder über eine anaerobe (sauerstoffunabhängige) oder aerobe (sauerstoffabhängige) Route und umfasst insgesamt mehr als 20 enzymatische Schritte.

Im Rahmen dieser Arbeit wird zunächst der Ursprung der Cobalaminsynthese im evolutionären Kontext analysiert. Es wird ermittelt, ob die Biosynthese von Cobalamin unabhängige Ursprünge in Archaeen und Bakterien aufweist oder ob diese möglicherweise bereits im letzten gemeinsamen Vorfahren allen Lebens (LUCA) vorhanden war. Dazu werden Phylogenien und Proteinstrukturen von 26 Enzymen untersucht, die an der anaeroben Synthese von Cobalamin und den damit verbundenen unteren Liganden beteiligt sind. Die Analyse der Proteinsequenzen von methanogenen Archaeen und acetogenen Bakterien lässt darauf schließen, dass sich die anaerobe Cobalaminsynthese in LUCA entwickelte und Enzyme aus bereits vorhanden Stoffwechselwegen übernommen wurden. Cobamide ersetzten möglicherweise, als enzymatische Cofaktoren in CoFeS, katalytisch aktive Übergangsmetalle in den ersten freilebenden Zellen und verdeutlichen exemplarisch den Übergang zwischen Geochemie und Biochemie in der frühen evolutionären Erdgeschichte. So zeigen Experimente, dass eine geochemische Variante des Acetyl-CoA-Wegs, in welchem Enzyme vollständig durch Übergangsmetalle ersetzt werden, Acetat und Pyruvat produziert.

Das aus dem Acetyl-CoA-Weg hervorgehende Pyruvat bildet den Ausgangspunkt der Gluconeogenese beziehungsweise den Endpunkt der Glykolyse und stellt somit einen zentralen Aspekt im prokaryotischen Metabolismus dar. Unter der Variation an glykolytischen Routen wird der Embden-Meyerhof (EM) Weg als klassische Glykolyse bezeichnet. Dieser umfasst einige enzymatische Schritte, die reversibel sind und so ebenfalls der Gluconeogenese dienen.

Der zweite Teil dieser Arbeit befasst sich mit der Frage, ob glykolytische Stoffwechselwege einen autotrophen, gluconeogenetischen Ursprung aufweisen und welche Rolle der Energiespeicher Glykogen in diesem Zusammenhang spielt. Im Zuge dessen werden die Verteilung und Phylogenien von Enzymen des klassischen und modifizierten EM Wegs sowie Enzyme der Synthese und des Abbaus von Glykogen in einem taxonomisch ausbalancierten Datensatz prokaryotischer Genome analysiert. Eine weite Verbreitung von glykolytischen bzw. gluconeogenetischen Enzymen in H_2-abhängigen, autotrophen Organismen, darunter methanogene Archaeen, die extrazelluläre Glucose nicht als Wachstumssubstrat mobilisieren können, korreliert mit einer breiten Verteilung von Enzymen des Glykogenmetabolismus. Zudem ergeben phylogenetische Rekonstruktionen, dass insbesondere die Enzyme der Energiegewinnungsphase (engl. *trunk glycolysis*) hoch konserviert und auf LUCA zurückzuführen sind. Dies lässt darauf schließen, dass die Glykolyse einen gluconeogenetischen Ursprung besitzt und anfangs im autotrophen Metabolismus Pyruvat über Triosephosphate in Glucose umwandelte, welche weiter zu Glykogen verkettet wurde und autotrophen Organismen als intrazelluläre Kohlenstoffreserve diente. In der Folge entwickelte sich eine Vielzahl glykolytischer Stoffwechselwege, die eine Mobilisierung von Glucose über Glykogen ermöglichten.

2 Abstract

Modern autotrophic theories on the origin of life posit that ancestral metabolism was based on hydrogen (H_2) and carbon dioxide (CO_2). These inorganic compounds form the starting point of the acetyl-CoA pathway, probably the oldest metabolic pathway for CO_2 fixation, which is utilised by both H_2-dependent methanogenic archaea and acetogenic bacteria. Cobamides play an important role in this pathway of CO_2 fixation as part of the corrinoid iron-sulfur protein (CoFeS). They belong to the group of corrins, complex macrocyclic compounds that contain a centrally bound cobalt ion and whose best-known representative is cobalamin (vitamin B_{12}). Cobalamin is synthesized either via an anaerobic (oxygen-independent) or aerobic (oxygen-dependent) route involving a total of more than 20 enzymatic steps.

This work analyses the origin of cobalamin synthesis in an evolutionary context. It determines whether the biosynthesis of cobalamin has independent origins in archaea and bacteria or whether it may have already been present in the last universal common ancestor (LUCA) of all life. Phylogenies and protein structures of 26 enzymes involved in the anaerobic synthesis of cobalamin and the associated lower ligands are examined. The analysis of protein sequences of methanogens and acetogens suggests anaerobic cobalamin synthesis evolved in LUCA, often by recruiting enzymes from already existing metabolic pathways. Cobamides, as enzymatic cofactors in CoFeS, may have replaced catalytically active solid-state transition metals on the way to the first free-living cells and exemplify the transition from geochemistry to biochemistry in the early evolutionary history of Earth. Experiments show that a geochemical variant of the acetyl-CoA pathway, in which enzymes are entirely replaced by transition metals, produces acetate and pyruvate.

Pyruvate synthesized by the acetyl-CoA pathway forms the starting point for gluconeogenesis as well as the end point for glycolysis, and thus represents a central aspect of prokaryotic metabolism. Among various glycolytic pathways, the Embden-Meyerhof (EM) pathway is referred to as classical glycolysis. It comprises several reversible enzymatic steps, which are therefore also serving gluconeogenesis.

The second part of this work deals with the question of whether glycolytic metabolic pathways have an autotrophic, gluconeogenic origin and what role the energy storage compound glycogen plays in this context. The distribution and phylogenies of enzymes of the classical and modified EM pathway as well as enzymes of glycogen synthesis and degradation

are analysed in a taxonomically balanced dataset of prokaryotic genomes. The widespread distribution of glycolytic and gluconeogenic enzymes in H_2-dependent autotrophs, including methanogens that cannot mobilise extracellular glucose as a growth substrate, correlates with a broad distribution of enzymes involved in glycogen metabolism. In addition, phylogenetic reconstructions show that the enzymes of trunk glycolysis are particularly highly conserved and can be traced back to LUCA. This suggests that glycolysis has a gluconeogenic origin and initially converted pyruvate, via triose phosphates, into glucose in primordial metabolism, which was further linked to glycogen serving as an intracellular carbon reserve in autotrophic organisms. Subsequently, a variety of glycolytic metabolic pathways evolved enabling the mobilization of glucose via glycogen.

3 Einleitung

3.1 Theorien zum Ursprung des Lebens

Die wissenschaftliche Frage zum Ursprung des Lebens beschäftigt die Menschheit spätestens seit Darwins Zeit. Bereits im 19. Jahrhundert argumentierte Ernst Haeckel gegen diverse Schöpfungstheorien, indem er annahm, dass sich sämtliches Leben nicht etwa aus einer schöpferischen Kraft, sondern aus einfachen „Urwesen“, sogenannten „Protoorganismen“ oder „Moneren“, entwickelte (Haeckel, 1866). In dem von ihm als „Autogonie“ bezeichneten Prozess formten sich erste verkettete Kohlenstoffverbindungen aus anorganischem Material, woraus sich schließlich spontan erste Formen einfacher organischer Individuen bildeten (Haeckel, 1866). Im Gegensatz dazu entwickelten John B. S. Haldane und Aleksandr I. Oparin in der ersten Hälfte des 20. Jahrhunderts unabhängig voneinander Theorien zum heterotrophen Ursprung des Lebens in einer primitiven Ursuppe. Laut Haldane formten sich einfache organische Moleküle aus Wasser (H_2O), Kohlenstoffdioxid (CO_2) und Ammoniak (NH_3) in der frühesten Phase der Erdgeschichte unter dem Einfluss von UV-Strahlung, die aufgrund von fehlendem Ozon in der Atmosphäre eine hohe Wirkung auf die Erdoberfläche ausübte. Diese organischen Substanzen, unter anderem Zuckermoleküle und Vorläufer von Proteinen, sammelten sich in primitiven Ozeanen an und brachten erste anaerobe Organismen in Form von großen molekularen Strukturen hervor, die sich ausschließlich in dieser Umgebung reproduzieren konnten (Haldane 1929). Oparin vermutete, dass organische Verbindungen in der frühen Erdatmosphäre sowohl miteinander als auch mit Sauerstoff und Stickstoff reagierten. Die daraus resultierenden komplexeren Moleküle sammelten sich nach dem Abkühlen der Atmosphäre im sich auf der Erdoberfläche befindenden Wasser und reagierten zu Verbindungen mit wachsender Komplexität. Durch Polymerisierung gewannen Makromoleküle an Struktur, grenzten sich von ihrer Umgebung ab und bildeten den Ausgangspunkt für das erste organismische Leben (Oparin 1938). Im Jahre 1953 gelang es Stanley L. Miller im sogenannten Miller-Urey-Experiment, die Aminosäuren Alanin und Glycin aus Wasser, Methan (CH_4), Ammoniak und Wasserstoff (H_2) in einem System zu bilden, welches die von Harold C. Urey beschriebenen hypothetischen Bedingungen der frühen Erdatmosphäre simulierte (Urey 1952; Miller 1953).

Eine weitere ausführlich diskutierte Fragestellung bezüglich des Ursprungs des Lebens beschäftigt sich damit, ob in der frühen Evolution zuerst primitiver Stoffwechsel oder die Replikation (RNA-Welt) entstand. Befürworter der RNA-Welt-Hypothese vermuten, dass RNA nicht nur als genetischer Informationsträger fungiert, sondern ebenfalls dazu in der Lage ist, sich selbst zu replizieren (Eigen 1971). Durch das Vorhandensein katalytisch aktiver Ribozyme spielten Protein-Enzyme in der Entstehung der ersten Lebensformen in der frühen Geschichte der Erde keine Rolle, sie entwickelten sich laut dieser Annahme erst später als effizientere Varianten der RNA-Enzyme (Cech 1986; Gilbert 1986; Orgel 2004). Die Theorie der RNA-Welt trifft allerdings auf diverse Kritikpunkte. Der Einfluss von Ribozymen auf den zellulären Stoffwechsel ist stark eingeschränkt, sie sind in den wenigsten metabolischen Kernprozessen von Bedeutung (Scossa und Fernie, 2020). Des Weiteren stellt sich die Frage, wie sich die für eine ribosomale Selbstreplikation notwendigen Metabolite in ausreichender Konzentration auf der frühen Erde ansammelten (Martin und Russell, 2003).

Theorien, die eine Entstehung erster metabolischer Prozesse der RNA-Welt-Hypothese bevorzugen (Ycas 1953; Wächtershäuser 1988), gewannen mit der Entdeckung hydrothermaler Tiefseequellen, z.B. *Lost City* (Kelley *et al*., 2001; Kelley *et al*., 2005), an Relevanz. Diese Umgebung spiegelt möglicherweise die Bedingungen wider, die in den Ozeanen der frühen Erde vorherrschten (u.a. alkalischer pH-Wert von 9 bis 11, Temperaturen von 28 bis 116 °C; Martin *et al*., 2008; Schwander *et al*., 2023). In der Gesteinskruste dieser Tiefseequellen läuft der Prozess der Serpentinisierung ab, in dem neben Serpentinit und dem Eisenoxid Magnetit (Fe_3O_4) ebenfalls Wasserstoff entsteht. Diese Bedingungen formten Theorien zur Entstehung des Lebens an hydrothermalen Tiefseequellen (Martin und Russell, 2003; Martin und Russell, 2007), welche den Übergang von geochemischen zu biochemischen Prozessen und die daraus resultierende Entwicklung der ersten Lebensform in den zuvor genannten Tiefseequellen erklären. Mikroskopische, metallhaltigen Poren bilden natürliche Nischen, die eine Erhöhung der Konzentration vorhandener Metabolite begünstigen und mithilfe von Übergangsmetallen (Eisen, Nickel, Cobalt), die als Katalysatoren fungieren, erste primitive Reaktionen ermöglichen (Martin *et al*., 2008). So werden im serpentinisierenden Hydrothermalfeld *Lost City* unter anderem die organischen Verbindungen Formiat und Methan in abiotischen Reaktionen aus Wasserstoff und Kohlenstoffdioxid gebildet (Lang *et al*., 2010). Darüber hinaus bewiesen diverse experimentelle Simulationen von Bedingungen hydrothermaler Tiefseequellen, dass Übergangsmetalle Enzyme ersetzen können und an der Bildung von Acetat und Pyruvat beteiligt sind (Preiner *et al*., 2020; Beyazay *et al*., 2023). Dabei wurden bis zu 200 µM Pyruvat gebildet (Beyazay *et al*., 2023), was der physiologischen

Pyruvatkonzentration in acetogenen Bakterien entspricht (Furdui und Ragsdale, 2000). Zudem katalysieren Übergangsmetalle Reaktionen des reduktiven Krebs-Zyklus (Muchoswka *et al*., 2017), die Reduktion von NAD^+ (Henriques Pereira *et al*., 2022) sowie Reaktionen in der Synthese von Aminosäuren (Kaur *et al*., 2024; Schlikker *et al*., 2024).

Somit bieten alkalische Hydrothermalquellen Ausgangsbedingungen, welche einerseits die Entstehung des letzten universellen gemeinsamen Vorfahren aller Zellen begünstigt haben könnten (Martin und Russell, 2003; Weiss *et al*., 2016) und andererseits konkrete Verbindungen zwischen geochemischen und biochemischen Reaktionen knüpfen.

3.2 Der letzte gemeinsame Vorfahre allen Lebens und lateraler Gentransfer

Charles Darwin formulierte bereits in der zweiten Hälfte des 19. Jahrhunderts seine berühmte Evolutionstheorie, das damit einhergehende Prinzip der natürlichen Selektion und die Annahme, dass alle Lebewesen von einem gemeinsamen Vorfahren abstammen (Darwin 1860). Der russische Biologe Konstantin S. Mereschkowsky publizierte Anfang des 20. Jahrhundert eine Arbeit, in der er unter anderem die ersten Organismen auf der jungen Erde als anaerobe Zellen beschrieb, die organische Verbindungen aus anorganischen Ausgangsstoffen ohne Chlorophyll bildeten (chemolithoautotroph) und Temperaturen von etwa 100 °C (thermophil) bevorzugten (Mereschkowsky 1910). Diese Annahme wurde seither fortlaufend aufgegriffen, unterstützt, vertieft (Fuchs und Stupperich, 1985; Martin und Russell, 2003; Ragsdale und Pierce, 2008) und in Hinblick auf die Physiologie des hypothetischen letzten gemeinsamen Vorfahren allen Lebens, oder auch *last universal common ancestor* (LUCA), untersucht (Weiss *et al*., 2016). Mit der Einstufung der Archaeen als eigenständige dritte Domäne (Woese und Fox, 1977) nahm Carl R. Woese an, dass LUCA dem letzten gemeinsamen Vorfahren der Archaeen, Bakterien und Eukaryoten entsprach (Woese 1998). Da allerdings die Endosymbiontentheorie besagt, dass die Eukaryoten-spezifischen Plastiden und Mitochondrien aus Prokaryoten hervorgegangen sind (Mereschkowsky 1905; Wallin 1927; Zimorski *et al*., 2014) und phylogenetische Rekonstruktionen dies bekräftigen (Raymann *et al*., 2015), wird angenommen, dass LUCA der letzte gemeinsame Vorfahre der Archaeen und Bakterien war.

Die Physiologie von LUCA wurde von Weiss *et al*. (2016) ausführlich analysiert. Dabei wurden 355 Proteinfamilien identifiziert, die auf LUCA zurückzuführen sind und darauf schließen lassen, dass LUCA ein H_2-abhängiger, autotropher Organismus war und sich möglicherweise in einer Umgebung entwickelte, die reich an Übergangsmetallen, H_2 und CO_2 war und so den zuvor erwähnten hydrothermalen Tiefseequellen entsprach (Weiss *et al*., 2016). Der Veröffentlichung ist außerdem zu entnehmen, dass LUCA über CO_2-Fixierung per Acetyl-CoA-Weg sowie eine ursprüngliche genetische Informationsverarbeitung mit verschiedenen ribosomalen Proteinen und tRNA-Synthetasen verfügte (Weiss *et al*., 2016). Da die Synthese von Methylgruppen im Acetyl-CoA-Weg zwischen Archaeen und Bakterien keine Homologie aufweist (Sousa und Martin, 2014) ist anzunehmen, dass LUCA Methylgruppen aus der geochemischen Umgebung mobilisierte (Sousa und Martin, 2014; Weiss *et al*., 2016). Somit markiert das theoretische Konstrukt LUCA einen wichtigen Punkt im Übergang zwischen der abiotischen und biotischen Phase in der frühen evolutionären Geschichte der Erde.

Rekonstruktionen potenzieller anzestraler Proteinfamilien können allerdings durch lateralen Gentransfer (LGT) erschwert werden. Lateraler Gentransfer, auch horizontaler Gentransfer genannt, beschreibt die Übertragung von Genen zwischen Prokaryoten, sowohl zwischen Arten als auch über Artgrenzen hinaus, innerhalb derselben Domäne oder zwischen Archaeen und Bakterien (Nelson-Sathi *et al*., 2015). Demnach findet der Gentransfer innerhalb einer Generation statt, nicht von einer Generation zur nächsten, wie es auf die vertikale Vererbung von Erbgut zutrifft. Der horizontale Austausch von Genen kann über verschiedene Prozesse erfolgen. Die Aufnahme genetischen Materials aus der unmittelbaren Umwelt eines Organismus wird als Transformation bezeichnet. Konjugation beschreibt den Prozess des Genaustauschs über direkte Zell-Zell-Interaktionen und die Übertragung von DNA mittels Bakteriophagen erfolgt durch den Vorgang der Transduktion (Jones und Sneath, 1970; Arnold et. al, 2022). Darüber hinaus gewinnt lateraler Gentransfer zwischen Prokaryoten durch sogenannte Nanoröhrchen (engl. *nanotubes*), Membranstrukturen, die eine physische Schnittstelle zwischen Zellen bilden, an Bedeutung (Dubey und Ben-Yehuda, 2011; Morawska und Kuipers, 2023). Durch dessen verschiedenen Variationen beeinflusst lateraler Gentransfer die Evolution von prokaryotischen Gruppen (Ochman *et al*., 2000), hinterlässt daher ebenfalls Spuren in phylogenetischen Rekonstruktionen und erschwert unter anderem die Berechnung eines universellen archaeellen und bakteriellen Stammbaums (Doolittle 1999). So kommt es vor, dass Gene unterschiedlicher Spezies nächste Nachbarn in einem Genbaum sind, obwohl dies nicht ihrer realen Beziehung entspricht, da diese möglicherweise eine größere evolutionäre

Distanz im Stammbaum aufweisen und im äußersten Fall zwei verschiedenen Domänen angehören.

Beispielhaft anzuführen ist der Methylsynthese-Weg des Bakteriums *Methylobacterium extorquens*, der, so die Annahme, eine laterale Abstammung von archaeellen Organismen aufweist (Chistoserdova *et al*., 1998). Ein weiteres prominentes Beispiel für lateralen Gentransfer spiegeln die Haloarchaeen wider, eine taxonomische Gruppe aerober, halophiler und heterotropher Archaeen, die ihren Ursprung in anaeroben, H_2-abhängigen methanogenen Archaeen finden und deren Evolution entscheidend durch den lateralen Transfer von bakteriellen Genen geprägt wurde (Nelson-Sathi *et al*., 2012). Einige der in etwa 1.000 horizontal transferierten Gene kodieren Membrantransporter, Komponenten bakterieller Atmungsketten und Enzyme, die Funktionen im katabolen Kohlenstoffmetabolismus ausüben (Nelson-Sathi *et al*., 2012).

Die Beispiele zeigen, dass lateraler Gentransfer einen großen Einfluss auf die Entwicklung prokaryotischer Gruppen nehmen kann und dies in phylogenetischen Analysen zu berücksichtigen ist.

3.3 Der Acetyl-CoA-Weg zur Kohlenstoffdioxid-Fixierung

Wie im vorherigen Kapitel beschrieben ergaben Rekonstruktionen, dass der letzte universelle gemeinsame Vorfahre allen Lebens über die Fähigkeit verfügte, Kohlenstoffdioxid über den Acetyl-CoA-Weg, auch Wood-Ljungdahl-Weg, zu fixieren (Weiss *et al*., 2016). Er dient neben der Fixierung von CO_2 ebenfalls der Energiekonservierung (Fuchs 2011; Martin 2020) und besteht aus einem Carbonyl-Zweig sowie einem Methyl-Zweig, die auch als westlicher und östlicher Zweig (engl. *Western* und *Eastern branch*) bezeichnet werden (Ragsdale 2011). Im Carbonyl-Zweig wird Kohlenstoffmonoxid (CO) aus Kohlenstoffdioxid durch die Kohlenmonoxid-Dehydrogenase (CODH) gebildet und an die Acetyl-CoA-Synthase (ACS) gebunden, wobei CODH und ACS häufig einen Enzymkomplex bilden (Adam *et al*., 2018). Der Methyl-Zweig dient der Umwandlung von CO_2 in eine Methylgruppe (CH_3) durch eine Reihe an Reduktionen von C1-Verbindungen, die an Tetrahydrofolat (H_4F) in Bakterien beziehungsweise an Tetrahydromethanopterin (H_4MPT) in Archaeen gebunden sind. Die dabei gebildete Methylgruppe wird durch das Corrinoid-Eisen/Schwefel-Protein (CoFeS) ebenfalls auf ACS übertragen, wo die Synthese von Acetyl-CoA aus CO, CH_3 und Coenzym A erfolgt

(**Abbildung 1**). Acetyl-CoA kann anschließend durch die Pyruvatsynthase (auch Pyruvat:Ferredoxin Oxidoreduktase) in Pyruvat oder unter der Bildung von 1 ATP in Acetat umgewandelt werden (Fuchs 2011; Ragsdale 2011).

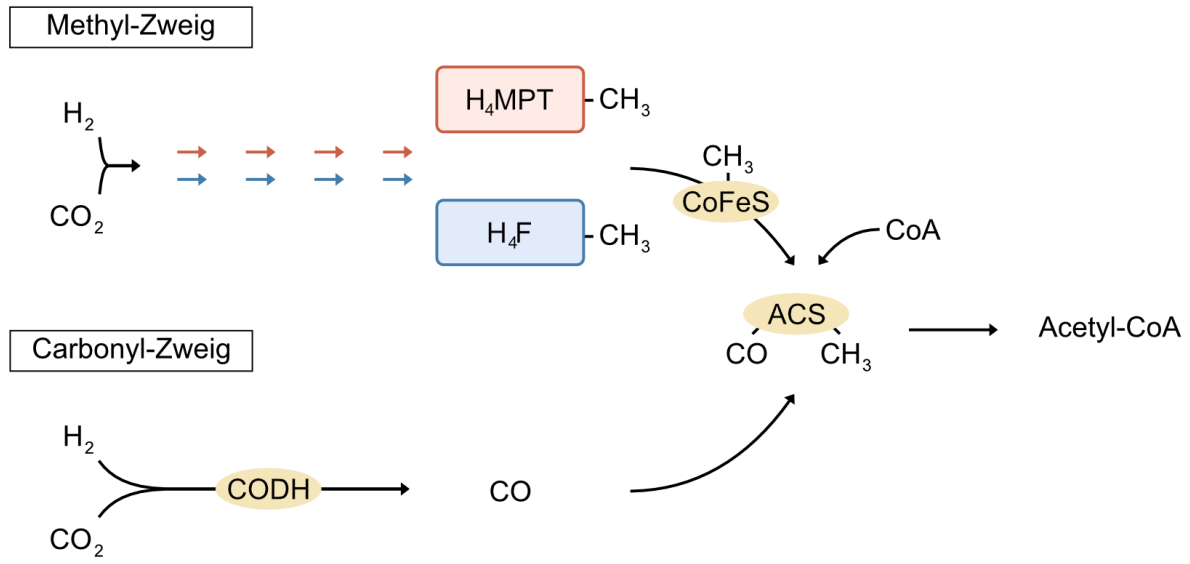


Abbildung 1. Schematische Darstellung des Acetyl-CoA-Wegs zur CO_2-Fixierung. Der Methyl-Zweig dient der schrittweisen Synthese von Methylgruppen (CH_3), welche in Archaeen (rot) und Bakterien (blau) unterschiedliche Enzyme umfasst. Das Corrinoid-Eisen/Schwefel-Protein überträgt Methylgruppen von einem Pterin-Cofaktor, Tetrahydromethanopterin (H_4MPT) in Archaeen oder Tetrahydrofolat (H_4F) in Bakterien, auf die Acetyl-CoA-Synthase (ACS). Diese katalysiert die Synthese von Acetyl-CoA aus CH_3, Coenzym A (CoA) und Kohlenmonoxid (CO), welches im Carbonyl-Zweig mithilfe der Kohlenmonoxid-Dehydrogenase (CODH) gebildet wird. Die Abbildung wurde nach Fuchs (2011), Ragsdale (2008) und Modjewski *et al*. (2024) modifiziert.

Einige Aspekte lassen darauf schließen, dass der Acetyl-CoA-Weg der älteste Prozess zur Kohlenstoffdioxidfixierung ist. Er läuft als einziger Weg zur Fixierung von CO_2 sowohl in Archaeen als auch in Bakterien ab und wird innerhalb dieser Domänen von H_2-abhängigen methanogenen Archaeen sowie acetogenen Bakterien genutzt. Vertreter dieser frühen prokaryotische Gruppen sind in H_2-produzierenden Hydrothermalquellen zu finden und spielen daher eine wichtige Rolle in der Erforschung der frühen Evolution in eben dieser Umgebung (Martin und Russell, 2007; Weiss *et al*., 2016; Colman *et al*., 2022). Darüber hinaus haben isotopische Daten von Methan ergeben, dass die Bildung von Methan (Methanogenese), und somit auch die Fixierung von CO_2 per Acetyl-CoA-Weg, bereits vor bis zu 3,8 Milliarden Jahren stattgefunden hat (Ueno *et al*., 2006). Außerdem bilden serpentinisierende Hydrothermalquellen unter abiotischen Bedingungen Formiat aus Wasserstoff und Kohlenstoffdioxid, ein Zwischenprodukt des Acetyl-CoA-Wegs (Lang *et al*., 2010). Experimente, die Eigenschaften hydrothermaler Tiefseequellen simulieren, produzieren

Formiat, Acetat und Pyruvat mithilfe von metallischen Katalysatoren, welche die Enzyme des Acetyl-CoA-Wegs ersetzen können (Varma *et al*., 2018; Preiner *et al*., 2020; Beyazay *et al*., 2023; Song *et al*., 2024).

Somit bildet der Acetyl-CoA-Weg ein Beispiel für die frühesten Phasen biochemischer Evolution, in welchen Stoffwechselprozesse durch Proteine, Cofaktoren und Metalle, die an Gesteinsoberflächen in der geochemischen Umgebung gebunden waren, vorangetrieben wurden (Mrnjavac *et al*., 2024a). Spuren dieser evolutionären Periode sind ebenfalls in der Struktur des im nachfolgenden beschriebenen Corrinoid-Eisen/Schwefel-Proteins in der Form von Cobalt und Eisen verankert (Abbanat und Ferry, 1991).

3.3.1 Das Corrinoid-Eisen/Schwefel-Protein

Im Acetyl-CoA-Weg der zuvor erwähnten anzestralen Archaeen und Bakterien spielt das Corrinoid-Eisen/Schwefel-Protein (CoFeS) eine wichtige Rolle (Svetlitchnaia *et al*., 2006). Dessen Corrinoid gehört zur Gruppe der Corrine, komplexe ringförmige, makrozyklische Cofaktoren, die unter anderem Cobamide, Coenzyme mit zentralen Cobaltatomen, umfassen. Sie unterscheiden sich in ihren unteren und oberen Liganden, beispielhaft sind 5,6-Dimethylbenzimidazol, 5-Hydroxybenzimidazol, 5-Methoxybenzimidazol oder Adenosin anzuführen (Cheong *et al*., 2001; Crofts *et al*., 2013). Corrine sind hauptsächlich an dem Transfer von Methylgruppen beteiligt, Corrin-abhängige Enzyme katalysieren allerdings zusätzlich Reaktionen in weiteren metabolischen Prozessen, beispielsweise in der Biosynthese von Aminosäuren (Gruber *et al*., 2011), Bacteriochlorophyll (Gough *et al*., 2000) und Hopanoiden (Schmerk *et al*., 2015). Darüber hinaus sind diese Enzyme ebenfalls an Fermentationen (Banerjee und Matthews, 1990; Matthews 2009) oder Entgiftungsvorgängen (Matthews 2009) beteiligt.

Im Acetyl-CoA-Weg katalysiert das Corrinoid-Eisen/Schwefel-Protein zwei Cobamid-abhängige Transferreaktion: die Methylgruppe eines Pterin-Cofaktors (H_4F in Acetogenen, H_4MPT in Methanogenen) wird (i) auf das zentrale Cobaltatom des Cobamids in CoFeS übertragen und (ii) anschließend auf ein Nickelatom im aktiven Zentrum der Acetyl-CoA-Synthase transferiert (Svetlitchnaia *et al*., 2006). Somit bildet das Corrinoid-Eisen/Schwefel-Protein als Verbindung zwischen dem Carbonyl- und Methyl-Zweig des Acetyl-CoA-Wegs einen essenziellen Bestandteil in der Biosynthese von Acetyl-CoA.

Es wird angenommen, dass Corrine bis in die frühe Evolution zurückreichen (Decker *et al*., 1970; Eschenmoser 1988), in diesem Zusammenhang wurden bereits einige Corrin-abhängige Enzyme auf LUCA zurückgeführt (Weiss *et al*., 2016). Zudem unterstützt das Auftreten von CoFeS in methanogenen Archaeen und acetogenen Bakterien die Annahme des Vorhandenseins dieses Proteins im letzten universellen gemeinsamen Vorfahren.

3.4 Cobalamin und dessen Biosynthese

Der vermutlich prominenteste Vertreter der Cobamide ist Cobalamin, ein Tetrapyrrol, welches auch als Vitamin B_{12} bekannt ist und 5,6-Dimethylbenzimidazol als unteren Liganden sowie eine Methyl- oder Adenosylgruppe als oberen Liganden in dessen aktiven Form trägt (Banerjee und Ragsdale, 2003). Die Biosynthese von Cobalamin und weiteren Cobamiden erfolgt ausschließlich in Prokaryoten (Warren *et al*., 2002). Viele Organismen, von Prokaryoten über einfache Eukaryoten bis hin zum Menschen, besitzen eine Corrin-Auxotrophie (Kolhouse und Allen, 1977; Rodionov *et al*., 2003; Zhang *et al*., 2009) und sind somit auf die Aufnahme des Cofaktors aus ihrer Umwelt angewiesen.

Prokaryoten, die in der Lage sind Cobalamin zu synthetisieren, bewerkstelligen dies entweder mithilfe eines aeroben (sauerstoffabhängigen) oder anaeroben (sauerstoffunabhängigen) Synthesewegs, beide Routen teilen sich in den ersten und finalen Reaktionsschritten diverse Enzyme (Bryant *et al*., 2020). Die zwei wesentlichen Unterschiede zwischen den Routen sind (i) die Verwendung von Sauerstoff während der aeroben Cobalaminsynthese bei der Umwandlung von Precorrin-3A in Precorrin-3B durch die Precorrin-3B-Synthase CobG, wodurch die Kontraktion des Corrin-Rings eingeleitet wird (Scott *et al*., 1993; Warren *et al*., 2002) und (ii) der Zeitpunkt der Cobalt-Insertion, welche entweder früh (anaerobe Cobalaminsynthese) oder spät (aerobe Cobalaminsynthese) erfolgt (Bryant *et al*., 2020). In der sauerstoffabhängigen Synthese von Cobalamin chelatiert der ATP-abhängige Klasse-I-Cobaltchelatasen-Komplex CobNST Hydrogenobyrinsäure a-c-Diamid, wodurch Cobyrinsäure a,c-Diamid entsteht (Debussche *et al*., 1992). Die anaerobe Route umfasst eine Klasse-II-Cobaltchelatase (CbiK/CbiX), welche die Insertion von Cobalt in Sirohydrochlorin mit der daraus resultierenden Bildung von Cobalt-Faktor II katalysiert (Raux *et al*., 1997; Brindley *et al*., 2003). Das eingefügte Cobalt-Ion wird in der Gesamtstruktur durch die vier Stickstoffatome des Tetrapyrrolrings sowie den unteren und oberen Liganden

koordinativ fixiert (Bryant *et al*., 2020). Bevor sich die Pfade der anaeroben und aeroben Cobalaminsynthese trennen, erfolgen jedoch bereits einige notwendige enzymatische Schritte.

Der Startpunkt für die O_2-abhängige und O_2-unabhängige Biosynthese von Cobalamin ist Uroporphyrinogen III (UroIII), das gemeinsame Ausgangssubstrat in der Synthese verschiedener modifizierter Tetrapyrrole, die unter anderem den Kofaktor F_{430}, Chlorophyll und Häm umfassen (Bryant *et al*., 2020). Die Produktion von UroIII startet mit 5-Aminolävulinsäure (5-ALA), die entweder über den C4-Weg aus Succinyl-CoA und Glycin (Shemin und Ritterberg, 1946) oder über den C5-Weg aus Glutamat (Huang *et al*., 1984) gebildet wird. Drei weitere Enzyme (Porphobilinogensynthase, Hydroxymethylbilansynthase und Uroporphyrinogen III-Synthase) wandeln anschließend die 5-Aminolävulinsäure schrittweise in Uroporphyrinogen III um. Abhängig von der Route, über die 5-ALA synthetisiert wird, sind somit entweder vier (C4-Weg) oder sechs (C5-Weg) Enzyme an der Produktion von UroIII beteiligt (Bryant *et al*., 2020).

Sowohl die aerobe als auch anaerobe Biosynthese von Cobalamin verläuft sukzessive über diverse Precorrin-Intermediate (Cobalt-Precorrin-Intermediate in der anaeroben Route) und Cobyrinsäure, wobei jeweils 15 Enzyme die Synthese von Adenosylcobyrinsäure (anaerobe Route) oder Adenosylcobyrinsäure-a,c-diamid (aerobe Route) aus UroIII katalysieren, bevor sich die zwei Varianten die finalen enzymatischen Schritte der Cobalaminbiosynthese teilen (Fang *et al*., 2017). In der finalen Phase wandeln drei weitere Enzyme (Adenosylcobalamin/α-Ribazol-Phosphatase, Adenosylcobinamidkinase und Adenosylcobinamid-GDP-Ribazoltransferase) die zuvor synthetisierte Adenosylcobyrinsäure in Adenosylcobalamin um (Bryant *et al*., 2020). Darüber hinaus werden für die Bildung des unteren Liganden 5,6-Dimethylbenzimidazol die Enzyme BzaA, BzaB, BzaC, BzaD und BzaE benötigt. BzaF, ein Radikal-S-Adenosylmethionin-Enzym, ist homolog zu BzaA, BzaB sowie zu der Phosphomethylpyrimidin Synthase ThiC und katalysiert in einigen Organismen die Synthese eines alternativen Liganden, 5-Hydroxybenzimidazol (Hazra *et al*., 2015). S-Adenosylmethionin (SAM) bezeichnet eine Gruppe enzymatischer Kofaktoren, die unter anderem dem Transfer von Methylgruppen und der Katalyse von Radikalreaktionen dienen (Broderick *et al*., 2014), so auch in der Biosynthese von Cobalamin, in welcher acht durch SAM vermittelte Methylgruppen in die Grundstruktur des Tetrapyrrols eingefügt werden (Frank *et al*., 2005).

Das im Acetyl-CoA-Weg aktive Corrinoid-Eisen/Schwefel-Protein von *Moorella thermoacetica* besteht unter anderem aus einem Corrin-Cofaktor, welcher 5-Methoxybenzimidazolyl als unteren Liganden trägt (Ragsdale *et al*., 1987). Durch die Variation

des unteren Liganden (Hazra *et al*., 2015) von Cobalamin, sowie zwei verschiedene Routen, die zu 5-Aminolävulinsäure führen, variiert die Anzahl an Enzymen, die an der Synthese dieses strukturell komplexen Tetrapyrrols beteiligt sind.

3.5 Glykolyse, Gluconeogenese und Glykogen-Synthese

Acetyl-CoA wird im bereits beschriebenen Acetyl-CoA-Weg in einer Reihe enzymatischer Schritte, unter anderem mithilfe von CoFeS, aus Kohlenstoffdioxid und Wasserstoff synthetisiert. Dieses kann im Anschluss in einer reversiblen Reaktion durch die Pyruvat:Ferredoxin Oxidoreduktase, auch als Pyruvat Synthase bekannt, zu Pyruvat umgewandelt werden (Shieh und Whitman, 1987), welches das Ausgangssubstrat für die Bildung von Glucose (Gluconeogenese) beziehungsweise das Endprodukt dessen Abbaus, der Glykolyse, bildet.

Für den Abbau von Glucose werden verschiedene glykolytische Wege mobilisiert. So bilden der Embden-Meyerhof (EM) Weg, der Entner-Doudoroff Weg, der Pentosephosphatweg sowie weitere metabolische Prozesse, die beispielsweise verkettete Glucosemoleküle (Glykogen) als intrazelluläre Kohlenstoff- und Energiereserven bereitstellen, einen zentralen Aspekt prokaryotischen Stoffwechsels (Bräsen *et al*., 2014; Wang *et al*., 2019; Cifuente *et al*., 2024). Im Vergleich zu Bakterien umfassen die archaeellen Routen teilweise abweichende, sequenziell und strukturell nicht-homologe Enzyme (Budgen und Danson, 1986; Hensel *et al*., 1989; Adams 1994; Schönheit und Schäfer, 1995; Selig *et al*., 1997; Ronimus und Morgan, 2003; Hansen *et al*., 2005; Siebers und Schönheit, 2005; Reher *et al*., 2007; Bräsen *et al*., 2014). Dementsprechend nutzen Archaeen hauptsächlich modifizierte Versionen des Embden-Meyerhof Wegs, Entner-Doudoroff Wegs, wie beispielsweise der von halophilen Archaeen verwendete semi-phosphorylierende ED Weg (Kuprat *et al*., 2021), oder eine Kombination aus beidem (Bräsen *et al*., 2014). Obwohl einige autotrophe Archaeen ebenfalls glykolytische Stoffwechselwege besitzen sind sie nicht dazu in der Lage, extrazelluläre Glucose als Wachstumssubstrat zu mobilisieren (Yu *et al*., 1994; Siebers und Schönheit, 2005; Sattler *et al*., 2024), was die Frage des Nutzens glykolytischer Routen in jenen Organismen aufkommen lässt. Im Gegensatz dazu können einige Bakterien ihren Metabolismus so anpassen, dass sie zwischen autotrophem und heterotrophem Wachstum variieren (Keller *et al*., 2019).

3.5.1 Der klassische und modifizierte Embden-Meyerhof Weg

Der nach Otto Meyerhof und Gustav Embden benannte Embden-Meyerhof Weg zum Abbau von Glucose wird häufig als klassische Glykolyse bezeichnet und ist in den meisten Organismen zu unterschiedlichen Vollständigkeitsgraden vorhanden (Fothergill-Gilmore und Michels, 1993). Die Route, die ein Glucosemolekül in zwei Pyruvatmoleküle umwandelt, wird üblicherweise in zwei verschiedene Phasen unterteilt: die Vorbereitungs- und die Energiegewinnungsphase.

In der Vorbereitungsphase wird Energie in Form von zwei ATP-Molekülen investiert und Glucose in Triosephosphate umgewandelt. Darauf folgt die Energiegewinnungsphase, im Englischen auch *trunk glycolysis* genannt, in der Triosephosphate weiter zu Pyruvat abgebaut werden und, wie der Name der Phase bereits vorwegnimmt, Energie erzeugt wird. Die Enzyme der Energiegewinnungsphase, Triosephosphatisomerase, Glycerinaldehyd-3-phosphat-Dehydrogenase, Phosphoglyceratkinase und -mutase, Enolase sowie Pyruvatkinase, sind deutlich konservierter als die Enzyme der Vorbereitungsphase (Fothergill-Gilmore und Michels, 1993). Fünf von ihnen finden ebenfalls Verwendung im Entner-Doudoroff Weg (Bräsen *et al*., 2014). In dieser Phase werden insgesamt vier ATP-Moleküle in den Reaktionen erzeugt, die (i) 1,3-Bisphosphoglycerat in 3-Phosphoglycerat (Phosphoglyceratkinase) und (ii) Phosphoenolpyruvat in Pyruvat (Pyruvatkinase) umwandeln. Da die Reaktionen der Glucokinase und 6-Phosphofructokinase in der Vorbereitungsphase zwei ATP verbrauchen, stellt der EM Weg einen Nettoenergiegewinn von zwei ATP-Molekülen pro Glucosemolekül zur Verfügung.

Neben dem klassischen Embden-Meyerhof Weg, der größtenteils von Bakterien genutzt wird, existiert eine weitere, modifizierte Variante dieses Stoffwechselwegs, die häufiger in Archaeen zu finden ist und abweichende Enzyme beinhaltet (Siebers und Hensel, 1993; Selig *et al*., 1997; Sakuraba und Ohshima, 2002; Sato und Atomi, 2011; Bräsen *et al*., 2014). So besitzen Archaeen häufig ADP-Glucokinasen und ADP-6-Phosphofructokinasen anstelle der ATP-abhängigen, meist von Bakterien genutzten Äquivalente (Siebers und Schönheit, 2005). Ein weiteres Beispiel bildet die Phosphoglucoseisomerase, welche die reversible Umwandlung von Glucose-6-phosphat in Fructose-6-phosphat katalysiert und dessen modifiziertes, archaeelles Pendant der Cupin-Superfamilie angehört (Verhees *et al*., 2001).

Klassischerweise läuft der Embden-Meyerhof Weg katabol (C6 → C3) ab, dennoch sind die meisten Enzyme in der Lage, die betroffenen Reaktionen reversibel (sowohl

glykolytisch als auch gluconeogenetisch) ablaufen zu lassen. Einige Ausnahmen bilden die Pyruvatkinase und 6-Phosphofructokinase, diese sind ausschließlich glykolytisch aktiv und müssen in der Gluconeogenese durch Phosphoenolpyruvatsynthase oder Pyruvat:Phosphat-Dikinase beziehungsweise Fructose-Bisphosphatase ersetzt werden (Bräsen *et al*., 2014). Darüber hinaus wandeln Archaeen im modifiziertem EM Weg Glycerinaldehyd-3-phosphat (GAP) in 3-Phosphoglycerat mithilfe von zwei weiteren, nicht-reversiblen Enzymen um (Reher *et al*., 2007). Dazu nutzen sie entweder die nicht-phosphorylierende Glycerinaldehyd-3-phosphat-Dehydrogenase GAPN, die Teil der Aldehyd-Dehydrogenase Superfamilie ist (Lorentzen *et al*., 2004; Reher und Schönheit, 2006) und auch in einigen Bakterien nachgewiesen wurde (Iddar *et al*., 2005), oder die Ferredoxin-abhängige Glycerinaldehyd-3-phosphat Oxidoreduktase GAPOR (Adams 1994). GAPN und GAPOR übernehmen in Archaeen die glykolytische Aufgabe der GAP-Dehydrogenase (GAPDH) und Phosphoglyceratkinase, da die archaeelle GAPDH im Gegensatz zur bakteriellen GAPDH, welche bidirektional agiert und per lateralem Gentransfer von Bakterien auf halophile Archaeen übertragen wurde, ausschließlich gluconeogenetisch fungiert (Tästensen und Schönheit, 2018). Die phylogenetisch anzestrale Einordnung der archaeellen GAPDH lässt vermuten, dass Archaeen den modifizierten Embden-Meyerhof Weg vorzugsweise für die Gluconeogenese nutzen (Schäfer und Schönheit, 1993; Siebers und Hensel, 1993; Bräsen *et al*., 2014; Tästensen und Schönheit, 2018), was zu der Annahme führte, dass der EM Weg ursprünglich anabol ablief und sich im Laufe der Evolution zu einem katabolen Stoffwechselweg entwickelte (Romano und Conway, 1996).

3.5.2 Die Frage nach dem Ursprung der Glykolyse

Auf den ersten Blick passt die Glykolyse in das Bild der frühen Evolution, weil die Erforschung dieses Stoffwechselwegs die ersten stöchiometrischen Reaktionen hervorbrachte, die Energie in der Form von ATP konservieren (Lipmann 1941), auch Substratkettenphosphorylierung genannt. Eine ursprüngliche Rolle glykolytischer Stoffwechselwege wurde bereits vor Jahrzehnten von einigen Wissenschaftlern vorgeschlagen (Wald 1964; de Duve 1991; Fothergill-Gilmore und Michels, 1993), darüber hinaus findet diese Annahme Befürworter in der jüngeren Vergangenheit (Keller *et al*., 2014). Wie kann jedoch ein heterotropher Ursprung der Glykolyse, der nicht mit H_2 und CO_2, sondern mit Glucose beginnt, in Einklang mit Theorien autotrophen Ursprungs des Metabolismus gebracht werden (Schönheit *et al*., 2016)?

Für Letzteren existieren standfeste Indizien, z.B. das Alter des Acetyl-CoA-Wegs zur Fixierung von Kohlenstoffdioxid (Fuchs und Stupperich, 1985; Fuchs 2011; Martin 2020) und dessen Ersetzbarkeit durch von Metallen katalysierten Reaktionen in Experimenten, welche die geochemischen Bedingungen hydrothermaler Tiefseequellen simulieren (Preiner *et al*., 2020; Belthle und Tüysüz, 2023; Beyazay *et al*., 2023).

Im Einklang mit Theorien zu autotrophen Stoffwechselprozessen in der frühen Evolution wird die Möglichkeit herangezogen, dass sich die Glykolyse ursprünglich gluconeogenetisch entwickelte und dem Aufbau von Glucose diente (Romano und Conway, 1996; Ronimus und Morgan, 2003). Zudem stellt freie Glucose, die von Mikroorganismen aus der Umwelt mobilisiert werden kann, wahrscheinlich eine relativ junge Entwicklung in der Geschichte der Evolution dar und geht möglicherweise auf den Ursprung der Zellulosesynthese in Landpflanzen vor etwa 500 Millionen Jahren zurück (Allen *et al*., 2019). Die Reversibilität glykolytischer Enzyme deutet darauf hin, dass diese in autotrophen Szenarien möglicherweise in erster Linie für die Gluconeogenese sowie die Synthese von intrazellulären Kohlenstoff- und Energiereserven in Form von verketteten Glucosemolekülen genutzt und im Laufe der Evolution als glykolytische Enzyme mobilisiert wurden. Dies wird ebenfalls durch die ausschließlich anabol agierende, phylogenetisch anzestrale archaeelle GAPDH (Tästensen und Schönheit, 2018) angedeutet. Zwei Enzyme der Glykolyse bzw. Gluconeogenese, Triosephosphatisomerase und Phosphoglucoseisomerase der Cupin-Superfamilie, wurden bereits dem letzten gemeinsamen Vorfahren allen Lebens zugeschrieben (Weiss *et al*., 2016).

3.5.3 Glykogen als frühe prokaryotische Energiereserve

Neben dem Nutzen in biosynthetischen Stoffwechselprozessen, z.B. in der Synthese von Zellwandbestandteilen (Albers und Mayer, 2011) kann Glucose weiter in Glykogen verarbeitet werden. Glykogen ist ein verzweigtes Polysaccharid, welches aus einer Vielzahl von Glucosemolekülen besteht, die 1,4-α-glykosidische- sowie 1,6-α-glykosidische-Verknüpfungen aufweisen und Zellen aller drei Domänen des Lebens unter anderem als Energiespeicher oder der Stressresistenz dienen (Wilkinson 1959; Wang *et al*., 2019).

Das Polymer kann über drei verschiedene enzymatische Routen produziert werden, wobei der klassische Synthese- und Abbauweg fünf Enzyme umfasst: GlgA, GlgB, GlgC, GlgP und GlgX. Im ersten Schritt katalysiert eine NDP-Glucose-Pyrophosphorylase (GlgC) die Umwandlung von Glucose in NDP-Glucose. Glykogensynthasen (GlgA) verknüpfen mehrere

NDP-Glucosemoleküle zu einem linearen Polymer, welches durch das Verzweigungsenzym (engl. *branching enzyme*) GlgB durch α-glykosidische Verbindungen weiter verzweigt und verdichtet wird (Preiss 2014; Cifuente *et al*., 2019). Der Glykogenabbau erfolgt durch die Glykogenphosphorylase GlgP in Kombination mit dem Entzweigungsenzym GlgX (engl. *debranching* enzyme) oder Glucoamylase (Chandra *et al*., 2011; Lee *et al*., 2021; Cifuente *et al*., 2024). Archaeen besitzen, wie auch im zuvor beschriebenen Embden-Meyerhof Weg, teils abweichende Enzyme. So unterscheiden sich das archaeelle Verzweigungsenzym und die archaeelle Glykogenphosphorylase der klassischen Glykogensynthese strukturell wesentlich von den bakteriellen Enzymen (Murakami *et al.*, 2006; Wang *et al*., 2019). Neben der klassischen Route existieren zwei nicht-klassische Wege zur Bildung von Glykogen, die vier (TreS, Pep2, GlgE, GlgB) und drei (GlgB, GlgC, Rv3032) Enzyme umfassen (Wang *et al*., 2019), Ersterer verknüpft den Glykogenstoffwechsel mit Trehalose (Chandra *et al*., 2011).

Obwohl Berichte über Glykogen in Archaeen bereits auf die 1980er Jahre zurückgehen (König *et al*., 1982), wurde dieser prokaryotischen Gruppe im Gegensatz zu den meisten Bakterien kein vollständiger klassischer Synthese- und Abbauweg mit allen fünf Enzymen zugeschrieben, allerdings besitzen einige Archaeen zumindest drei der involvierten Enzyme (Wang *et al*., 2019). Im Gegensatz dazu ist der nicht-klassische Weg, der sich aus vier Enzymen zusammensetzt, weiter in Archaeen verbreitet als zuvor angenommen (Chandra *et al*., 2011) und die zweite nicht-klassische Route scheint hauptsächlich in thermophilen Spezies verteilt zu sein (Wang *et al*., 2019).

Glykogen und die damit verbundene Synthese wurde ebenfalls in methanogenen Archaeen nachgewiesen (Yu *et al*., 1994; Gonzales-Ordenes *et al*., 2024). Dieses akkumuliert im Zytosol als Energiereserve (König *et al*., 1985; Murray und Zinder, 1987, Maitra *et al*., 2001) und dient somit unter anderem anzestralen prokaryotischen Gruppen als intrazelluläre Quelle für C6-Kohlenstoffe.

4 Zielsetzung

Die Frage zum Ursprung des Lebens beschäftigt die Wissenschaft und Menschheit seit jeher. Es existieren daher diverse Theorien zur Entstehung der ersten lebenden Zellen und dem damit einhergehenden frühzeitlichen Stoffwechsel. So gibt es zahlreiche Forschungen, die sich mit einem autotrophen Ursprungsszenario befassen, in dem einfache Stoffwechselprozesse auf Wasserstoff und Kohlenstoffdioxid basieren. Hydrothermale Tiefseequellen spielen in diesem Kontext eine prominente Rolle, da sie optimale Bedingungen für die Entstehung anzestraler metabolischer Stoffwechselwege, wie dem H_2- und CO_2-abhängigen Acetyl-CoA-Weg, bieten (Martin und Russell, 2007). In diesem Zusammenhang wurden unter anderem Analysen zum Umfang des autotrophen Metabolismus (Wimmer *et al*., 2021) in der frühen Erdgeschichte und zum Gengehalt des letzten gemeinsamen Vorfahren allen Lebens (Weiss *et al*., 2016), LUCA, unternommen. Die Erforschung der evolutionären Historie verschiedener Stoffwechselwege ist in dieser Hinsicht von großer Bedeutung, da sie zum Gesamtbild eines autotrophen Ursprungs des Lebens beitragen.

Diese Arbeit verfolgt das Ziel, den Ursprung der anaeroben Synthese von Cobalamin und deren Rolle im anzestralen, autotrophen Metabolismus zu ermitteln. In diesem Zuge wird die Bedeutung der Cobalaminsynthese mit Blick auf das Corrinoid-Eisen/Schwefel-Protein des Acetyl-CoA-Wegs im evolutionären Kontext betrachtet. Dazu werden initial prokaryotische Referenzgenome auf das Vorhandensein von Proteinsequenzen der betroffenen Enzyme untersucht. Vergleiche von Sequenzen und Strukturen sowie anschließende phylogenetische Rekonstruktionen bilden den Ausgangspunkt, um die evolutionäre Verwandtschaft und den Ursprung archaeeller und bakterieller Enzyme der anaeroben Cobalaminsynthese zu ergründen.

Darüber hinaus beschäftigt sich diese Arbeit mit der Frage, wie die Vielzahl glykolytischer Stoffwechselwege mit Theorien zu vereinigen ist, die einen H_2- und CO_2-abhängigen Ursprung des Lebens vorschlagen und welche Rolle Glykogen in diesem Zusammenhang einnimmt. Die Identifizierung prokaryotischer Enzyme des klassischen und modifizierten Embden-Meyerhof Wegs sowie Enzyme des Glykogenmetabolismus in Referenzgenomen, sequenzielle Vergleiche sowie umfangreiche phylogenetische Berechnungen dienen der Einordnung der Glykolyse in den Kontext eines autotrophen Stoffwechsels in der frühen evolutionären Phase der Erdgeschichte.

5 Publikationen

I Evidence for corrin biosynthesis in the last universal common ancestor

Luca D. Modjewski[$1], Val Karavaeva[$2,3], Natalia Mrnjavac[$1], Michael Knopp[1], William F. Martin[1], Filipa L. Sousa[2]

[1] Institut für Molekulare Evolution, Heinrich-Heine-Universität Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Deutschland

[2] Department für Funktionelle und Evolutionäre Ökologie, Universität Wien, Djerassipl. 1, 1030 Wien

[3] Vienna Doctoral School of Ecology and Evolution, Universität Wien, Djerassipl. 1, 1030 Wien

[$] Gleicher Beitrag

Diese Publikation wurde am 21. Dezember 2024 in *The FEBS Journal* veröffentlicht.

Beitrag von Luca D. Modjewski (Geteilte Erstautorenschaft)

Ich war an folgenden Aspekten der Publikation beteiligt: Methodologie, Recherche, Zusammenstellung und Analyse des Datensatzes, Visualisierung der zuvor gewonnen Ergebnisse (Präsenz-Abwesenheits-Matrix und Phylogenien der zugrundeliegenden Enzyme), Visualisierung des anaeroben Cobalamin-Biosynthesewegs, Anfertigung der ergänzenden Abbildungen S1, S2, S3 und S4 sowie Erstellung und Überarbeitung des Manuskripts.

The FEBS Journal

FEBS PRESS science publishing by scientists

Evidence for corrin biosynthesis in the last universal common ancestor

Luca D. Modjewski[1], Val Karavaeva[2,3], Natalia Mrnjavac[1], Michael Knopp[1], William F. Martin[1] and Filipa L. Sousa[2]

1 Institute of Molecular Evolution, Faculty of Mathematics and Natural Sciences, Heinrich Heine University Düsseldorf, Germany
2 Department of Functional and Evolutionary Ecology, University of Vienna, Austria
3 Vienna Doctoral School of Ecology and Evolution, University of Vienna, Austria

Keywords
biochemical evolution; cobalamin; cobamide; tetrapyrroles; vitamin B_{12}

Correspondence
F. L. Sousa, Department of Functional and Evolutionary Ecology, University of Vienna, Djerassipl. 1, 1030 Vienna, Austria
Tel: +43 1 4277 76510
E-mail: filipa.sousa@univie.ac.at

Luca D. Modjewski, Val Karavaeva, and Natalia Mrnjavac contributed equally to this article.

(Received 12 August 2024, revised 4 October 2024, accepted 10 December 2024)

doi:10.1111/febs.17367

Corrinoids are cobalt-containing tetrapyrroles. They include adenosylcobalamin (vitamin B_{12}) and cobamides that function as cofactors and coenzymes for methyl transfer, radical-dependent and redox reactions. Though cobamides are the most complex cofactors in nature, they are essential in the acetyl-CoA pathway, thought to be the most ancient CO_2-fixation pathway, where they perform a pterin-to-cobalt-to-nickel methyl transfer reaction catalyzed by the corrinoid iron–sulphur protein (CoFeS). CoFeS occurs in H_2-dependent archaeal methanogens, the oldest microbial lineage by measure of physiology and carbon isotope data, dating corrinoids to ca. 3.5 billion years. However, CoFeS and cobamides are also essential in the acetyl-CoA pathway of H_2-dependent bacterial acetogens. To determine whether corrin biosynthesis was established before archaea and bacteria diverged, whether the pathways arose independently or whether cobamide biosynthesis was transferred from the archaeal to the bacterial lineage (or *vice versa*) during evolution, we investigated phylogenies and structural data for 26 enzymes of corrin ring and lower ligand biosynthesis. The data trace cobamide synthesis to the common ancestor of bacteria and archaea, placing it in the last universal common ancestor of all lifeforms (LUCA),

Abbreviations
ACS, acetyl-CoA synthase; AcsE, Methyl-H_4F:CoFeS methyltransferase; BchE, anaerobic magnesium-protoporphyrin IX monomethyl ester cyclase; BchQ, chlorophyll/bacteriochlorophyll a synthase; BchR, bacteriochlorophyllide d C-12(1)-methyltransferase; BioD, dethiobiotin synthetase; BtuR/CobA/PduO, corrinoid adenosyltransferase; BzaABCDEF, 5-hydroxybenzimidazole synthase; CbiA, cobyrinic acid a, c-diamide synthetase; CbiB, adenosylcobinamide-phosphate synthase; CbiC, precorrin-8X/cobalt-precorrin-8 methylmutase; CbiD, cobalt-precorrin-5B (C1)-methyltransferase; CbiE, cobalt-precorrin-7 (C5)-methyltransferase; CbiF, precorrin-4/cobalt-precorrin-4 C11-methyltransferase; CbiG, cobalt-precorrin 5A hydrolase; CbiH, precorrin-3B C17-methyltransferase/cobalt-factor III methyltransferase; CbiJ, precorrin-6A/cobalt-precorrin-6A reductase; CbiL, precorrin-2/cobalt-factor-2 C20-methyltransferase; CbiT, cobalt-precorrin-6B (C15)-methyltransferase; $CbiX_L$/CbiK/CbiXs, sirohydrochlorin cobaltochelatase; CfbB, Ni-sirohydrochlorin a,c-diamide synthase; CobB, NAD-dependent protein deacetylase/lipoamidase; CobC/CobZ, adenosylcobalamin/alpha-ribazole phosphatase; CobQ, adenosylcobyric acid synthase; CobS, adenosylcobinamide-GDP ribazoletransferase; CobU/CobY, adenosylcobinamide-phosphate guanylyltransferase; CODH, carbon monoxide dehydrogenase; CoFeS, corrinoid iron–sulphur protein; DMB, 5,6-dimethylbenzimidazole; DsrAB, dissimilatory sulphite reductase; EutT, ethanolamine utilization cobalamin adenosyltransferase; FdhA, formate dehydrogenase; Fhs, N10-formyl-H_4F synthetase; FolD, 5,10-methenyl-H_4F cyclohydrolase/dehydrogenase; Ftr, formylmethanofuran:H_4MPT formyltransferase; FwdBD, formylmethanofuran dehydrogenase; Ga, Giga annum; H_4F, tetrahydrofolate; H_4MPT, tetrahydromethanopterin; HBI, 5-hydroxybenzimidazole; HemH/PpfC/CpfC, ferrochelatase; HGT/LGT, horizonal/lateral gene transfer; Hmd, 5,10-methenyltetrahydromethanopterin hydrogenase; ICGs, informational core genes; LUCA, last universal common ancestor; MazG, nucleoside triphosphate diphosphatase; Mch, methenyl-H_4MPT cyclohydrolase; Mer, 5,10-methylenetetrahydromethanopterin reductase; MetF, 5,10-methylene-H_4F reductase; MetF, 5,10-methylenetetrahydrofolate reductase; MeTr, methyltransferase present in some acetogens; MinD, septum site-determining protein; MoCo, molybdenum cofactor; Mtd, F_{420}-dependent methylene-H_4MPT dehydrogenase; ParA, chromosome partitioning protein; PFOR, pyruvate:ferredoxin oxidoreductase,; SAM, *S*-adenosylmethionine; SirB/ShfC, sirohydrochlorin ferrochelatase; SirC, precorrin-2 dehydrogenase; ThiC, phosphomethylpyrimidine synthase; UroIII, uroporphyrinogen III; UroM/SuMT/SirA, uroporphyrin-III C-methyltransferase.

while pterin-dependent methyl synthesis pathways likely arose independently post-LUCA in the lineages leading to bacteria and archaea. Enzymes of corrin biosynthesis were recruited from preexisting ancient pathways. Evolutionary forerunners of CoFeS function were likely Fe-, Ni- and Co-containing solid-state surfaces, which, in the laboratory, catalyze the reactions of the acetyl-CoA pathway from CO_2 to pyruvate under serpentinizing hydrothermal conditions. The data suggest that enzymatic corrin biosynthesis replaced insoluble solid-state catalysts that tethered primordial CO_2 assimilation to the Earth's crust, suggesting a role for corrin synthesis in the origin of free-living cells.

Introduction

Corrinoids are nature's most complex cofactors and coenzymes. They include cobalt-containing cobamides, which differ with respect to the lower ligand: 5,6-dimethylbenzimidazole in cobalamin, and 5-hydroxybenzimidazole, 5-methoxybenzimidazole, 5-methylbenzimidazole, benzimidazole, 5-methoxy-6-methyl benzimidazole, phenol, *p*-cresol, 2-methylsulfinyladenine, 2-methylsulfonyladenine, 2-methyladenine or adenosine in other cobamides [1–6]. These cobalt-containing corrins are mainly involved in methyl transfer [7–10] and in radical-based reactions [11], including the generation of deoxyribose nucleotides for DNA by adenosylcobalamin-dependent ribonucleotide reductase [12–14]. Corrinoid-dependent enzymes are also involved in a variety of other processes, such as the biosyntheses of amino acids (Met, [10,15,16]), queuosine [17], bacteriochlorophyll [18], hopanoids [19], polytheonamide [10,20], antibacterials and antivirals [21,22], as well as the detoxification of aromatic and aliphatic chlorinated organics [16,23], and a range of fermentations: amino acids [15,24–27], one-carbon compounds (e.g., methylamine, methanol and methylthiol [16]), ethanolamine and choline [16,28,29], nicotinic acid [16,30], 1,2-diols [16,31–33] and fatty acids [16,26,27]. Corrinoids are required by many organisms but are synthesized by few [34–37]. The biosynthetic pathway in anaerobes entails over 20 enzymatic steps if starting to count from uroporphyrinogen III, the first macrocyclic intermediate and the universal precursor of tetrapyrrole biosynthesis [38], the number of steps to the final product depending on the nature of the lower ligand [39]. Though the genes for corrin synthesis can be laterally transferred across lineages, and some prokaryotes have partial cobalamin synthesis, for example, from cobyric acid to adenosylcobalamin [40], by far the most common strategy to satisfy corrin requirements is to obtain the cofactor from corrin-producing organisms via the environment [34]. The list of corrin auxotrophs includes prokaryotes [35–37,41], algae [37,42], protists [37,43] and humans who strictly require cobalamin for only two enzymatic reactions: the isomerization of methylmalonyl-CoA to succinyl-CoA and the synthesis of methionine from homocysteine [44]. Corrins are thought to be ancient [45,46], and earlier studies traced some corrin-dependent enzymes to the last universal common ancestor, LUCA [47], but the issue of whether the enzymes of corrin biosynthesis themselves trace to LUCA is so far unresolved.

The evidence for corrin antiquity stems from its essential role in the acetyl-CoA pathway, the most ancient among CO_2-fixing pathways [48–51]. In the acetyl-CoA pathway of H_2-dependent acetogens (bacteria) and H_2-dependent methanogens (archaea), the corrinoid iron–sulphur protein (CoFeS) catalyzes a cobamide-dependent cobalt-to-nickel methyl transfer reaction that is essential for acetyl-CoA synthesis [52,53] (Fig. 1A). The geological age of the corrins is given by methanogens, which are the most ancient lineage of microbes by the measure of physiology [46] and carbon isotope data. Methane isotopes [65] provide the earliest evidence for life on Earth, tracing methanogenesis, hence the acetyl-CoA pathway, hence an ancestral CoFeS (judging by the homology between bacterial and archaeal CoFeS), hence corrins, into rocks 3.8 billion years of age. Isotopically ultralight carbon is also reported in 3.8 Ga sediment carbon [66,67], which suggests the presence of the acetyl-CoA pathway [68] without discriminating an acetogen or methanogen source. Phylogenetic studies implicate acetogens and methanogens as the most ancient lineages among modern bacteria and archaea, respectively [47]. Both acetogens and methanogens inhabit H_2-producing hydrothermal systems today [69,70] and

might present windows into the physiology and habitat of the first cells [47,69–72].

In addition to physiology [48–51], isotopes [65–67] and phylogeny [47,69,70], metal catalysts provide an independent line of evidence in favour of the antiquity of the acetyl-CoA pathway. Serpentinizing (H_2-producing) hydrothermal systems synthesize formate [73], an intermediate of the acetyl-CoA pathway [60], and methane, the end product of methanogenesis, from abiotic reactions of H_2 and CO_2 [74]. Though the catalysts that facilitate formate and methane synthesis in vents have not yet been identified *in situ*, native (zero-valent) transition metals are thought to be the catalysts [60], for two reasons. First, Fe^0, Ni^0, Co^0 and their alloys are naturally deposited in serpentinizing hydrothermal vents because of their highly reducing conditions [75]. Second, Fe^0, Ni^0, Ni_3Fe and silica-supported Co-Fe alloys can replace the proteins that synthesize formate, acetate, pyruvate from H_2 and CO_2 via the acetyl-CoA pathway [60–64] (Fig. 1C). In water and in the absence of enzymes, the solid-state metals generate methyl groups from H_2 and CO_2, producing formate, acetate, pyruvate and methane, replacing the function of over 100 enzymes [76]. Furthermore, Ni^0 readily forms amino acids from 2-oxoacids, NH_3 and H_2 [77], and Fe^0 can reduce ferredoxin, functionally substituting for flavin-based electron bifurcation, a complex process that requires multi-enzyme systems capable of coupling exergonic and endergonic reactions [78] and reducing the low-potential physiological donor of electrons in the acetyl-CoA pathway, ferredoxin, with electrons from H_2 [79].

Countering the case for corrin antiquity, its biosynthetic pathway is the longest, and the product is structurally the most complex of any known tetrapyrrole [38,80,81], raising the question of its possible functional precursor during biochemical evolution (Fig. 1). Corrin synthesis itself holds no clear evidence for antiquity, other than its basal position in tetrapyrrole biogenesis [46,82]. While some cofactors such as thiamine or pyridoxal phosphate participate in their own biosynthesis, forming small autocatalytic cycles [71], that is not the case for corrins. The pathway involves the introduction of eight methyl groups to the tetrapyrrole backbone, one of which aids methylene elimination and contraction of the porphyrin to the corrin ring [83,84]. The seven methyl groups present in the corrinoid structure [9] (Fig. 1B) are donated by *S*-adenosylmethionine (SAM), a cofactor of methyl transfers and radical reactions [85], yet SAM does not substitute for cobamide in either the CoFeS reaction [52] or in the pumping reaction at MtrA-H in methanogens [86].

The age of methanogens and the identification of functional abiotic homologues of the acetyl-CoA pathway point to simple environmental precursors of corrin-dependent methyl transfer (Fig. 1C) but do not directly address corrin biosynthesis antiquity. To determine whether corrin synthesis required for the acetyl-CoA pathway traces to the common ancestor of bacteria and archaea – the last universal common ancestor (LUCA) – or was the result of later lateral gene transfers [47], we have investigated the evolutionary history of the enzymes of corrin biosynthesis in acetogens and methanogens [38], including the synthesis of the lower ligand (Fig. 1B) [39] and CoFeS [87,88].

Results and Discussion

There are two pathways of corrin synthesis, an O_2-dependent (aerobic) pathway and an oxygen-sensitive (anaerobic) pathway, where the aerobic pathway requires oxygen for ring contraction and cobalt insertion [38,81]. Only the enzymes of the assumed-to-be-more-ancient anaerobic pathway are of interest in early biochemical evolution for three reasons. First, O_2 is a product of cyanobacterial photosynthesis, which was lacking [89] in the environment at the origin of the acetyl-CoA pathway and methanogens > 3.5 billion years ago [65,82]. Second, both abiotic [90] and enzymatic reactions to pyruvate synthesis along the acetyl-CoA pathway are extremely O_2-sensitive due to the involvement of FeS clusters in electron transfer and other radical-generating reactions [91,92], such that the corrins required for the pathway arose in the absence of O_2 (as with chlorophyll synthesis, the O_2-dependent corrin synthesis reactions arose in response to O_2-dependent inhibition of radical-dependent reactions [93]). Third, acetogens and methanogens themselves, which depend upon CoFeS and the acetyl-CoA pathway for growth [48,49,78], are extremely O_2-sensitive and use the anaerobic corrin synthesis pathway. Because physiology [46], phylogeny [47,69,70], and carbon isotopes [65,66,68] point to acetogens and methanogens as the most ancient bacterial and archaeal lineages, respectively, and because both require cobamide [7,8], the corresponding genes of anaerobic corrin biosynthesis were extracted from the genome of the model acetogen *Moorella thermoacetica* [94] and from the genome of the methanogen *Methanococcus maripaludis* [95].

Corrinoids are not generally viewed as likely products of prebiotic synthesis because of their complex structure and biosynthesis [38,39,96], which is outlined in Fig. 2 (and in Fig. S1; see Supporting Information

for a description of the initial biosynthetic steps). Of note, the initial steps, from Uro-III to the formation of sirohydrochlorin, are common to siroheme biosynthesis [99], and homologues of CbiA and CbiX are involved in F_{430} biosynthesis [100,101], highlighting the duplications that occurred during tetrapyrrole evolution. At a late stage of the pathway (Fig. 2), adenosylation (catalyzed by BtuR/CobA) introduces an adenosyl moiety as the upper ligand at this stage of the adenosylcobalamine synthesis pathway [38]. In the cobamides that are employed in the CoFeS reaction, the adenosyl upper ligand is missing, the corresponding Co coordination site being occupied by water or by the methyl group during the methyl transfer

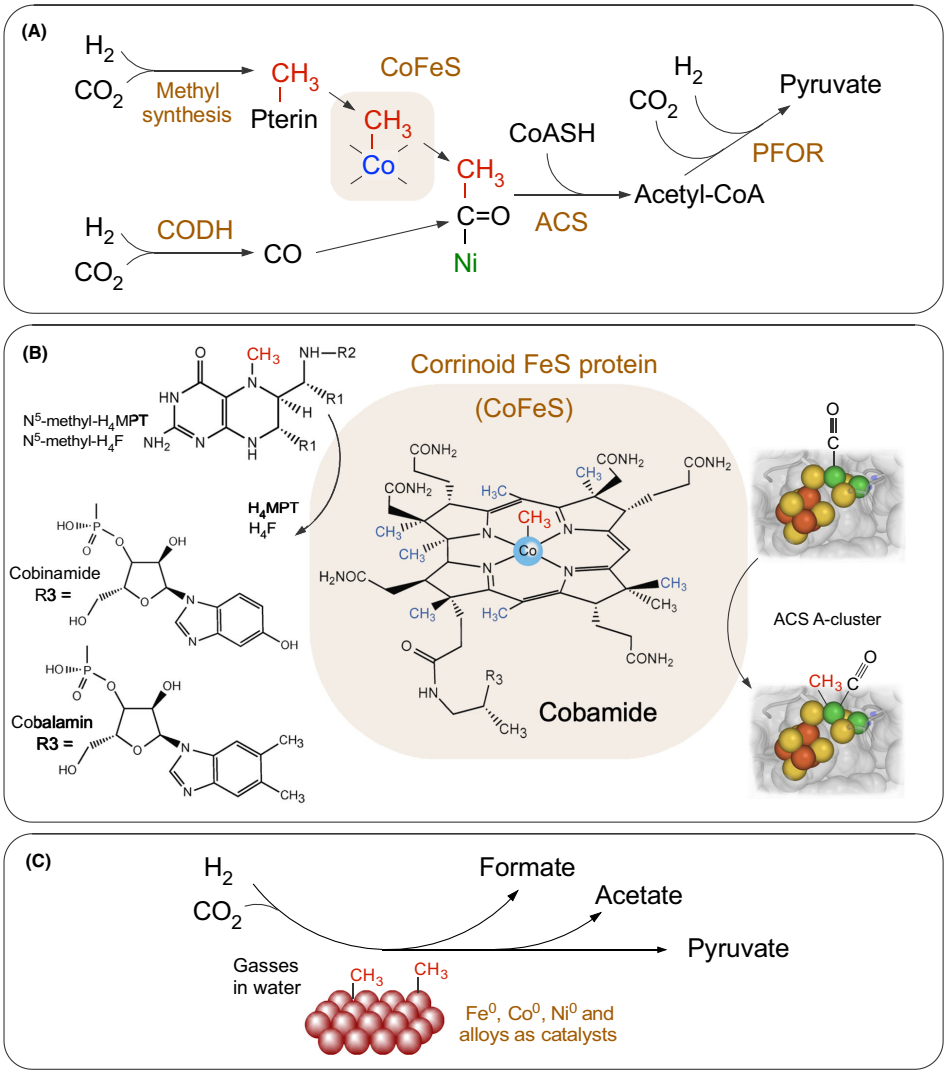


Fig. 1. The role of corrins in the acetyl-CoA pathway. (A) The acetyl-CoA pathway, redrawn from references [49, 54]. Methyl synthesis: pterin-dependent methyl synthesis in the acetyl-CoA pathway of bacteria and archaea from H_2 and CO_2 [49]. ACS, acetyl-CoA synthase [53]; CODH, carbon monoxide dehydrogenase [55]; CoFeS, corrinoid iron–sulphur protein [52]; PFOR, pyruvate:ferredoxin oxidoreductase [54]. The methyl group is highlighted in red, the corrinoid cobalt atom in blue, the proximal nickel atom of the ACS active site [53] in green and enzyme names or steps are given in sepia letters. (B) Corrinoid-dependent methyl transfer from a pterin cofactor to the active site metal cluster (A-cluster) of acetyl-CoA synthase (ACS) (PDB ID: 7ZKJ [56]) in the acetyl-CoA pathway of acetogens and methanogens. The corrinoid iron–sulphur protein CoFeS catalyzes the methyl-transfer reaction, in some cases aided by a separate methyltransferase (MeTr) [52,57]. H_4F and H_4MPT: tetrahydrofolate and tetrahydromethanopterin, respectively. R1 = –H in H_4F and $–CH_3$ in H_4MPT.For the chemical structure of R2,in H4F and H4MPT, , see[58]. Methylcobalamin is shown in the base-off configuration [59], with –R3 replacing the nucleotide loop. Methyl groups derived from *S*-adenosylmethionine (SAM) are shown in blue. Fe atoms are shown in orange, S atoms in yellow and Ni atoms in green. Lower left: Examples of two cobamide lower ligands (–R3 in A); 5-hydroxybenzimidazole (HBI) is found in the cobinamides of some methanogens [39], while 5,6-dimethylbenzimidazole (DMB, the lower ligand base of cobalamin) is found in some bacteria [52]. (C) Native metals catalyze formate, acetate and pyruvate synthesis from H_2 and CO_2 over Ni^0, Fe^0, Co^0, Ni_3Fe and other alloys [60–64], identifying a nonenzymatic, geochemical precursor of corrin-dependent methyl transfer shown in panels A and B. More is known about the enzymatic reaction mechanisms [53] than about the metal-catalyzed forerunners [60–64]. The catalyst-bound methyl groups in panel C are probable intermediates in the reaction mechanism, as methanol and methane are sometimes observed as products [60,61]. In some experiments, Fe_3S_4 and Fe_3O_4 are effective as catalysts, but it cannot be excluded that these catalysts are reduced to Fe^0 at some sites during the reaction with H_2.

reaction [52]. Accordingly, some methanogens (and other archaea) lack BtuR/CobA homologues [36,102]. The biosynthetic route depicted in Figs 2 and 3 follows that described for adenosylated intermediates [38,96], the corresponding genes, except for BtuR/CobA, being present in methanogens and acetogens that are included in this study.

The anaerobic synthesis of the cobalamin lower ligand dimethylbenzimidazole (DMB) is performed by the *bzaABCDE* gene products [39], starting from 5-aminoimidazole ribotide (AIR), an intermediate of thiamine and purine biosynthesis. In some organisms, only the product of *bzaF*, a radical SAM enzyme homologous to *bzaA*, *bzaB* and *thiC* products, is present [39,109], generating 5-hydroxybenzimidazole (HBI) as the cobamide lower ligand (Fig. 1C). Other organisms employ different intermediates of lower ligand synthesis [39], and the lower ligands of cobamides isolated from acetogens [7] and methanogens vary across species [8]. Here we analysed only the biosynthesis of the lower ligand DMB since it is found in both archaeal and bacterial organisms and has a large impact in marine settings [110]. Moreover, it was shown that at least in *Salmonella enterica*, the incorporation of DMB or of adenine as lower ligands is performed by the same enzymes, and, in DMB-limiting conditions, adenine can replace DMB [111].

Homologies trace corrin synthesis to the acetogen-methanogen common ancestor

Alignments of structures generated from amino acid sequences with AlphaFold [103,105] show that the structures of enzymes for methanogen and acetogen cobamide synthesis superimpose well in all cases except one, indicating homology of the corrin synthesis pathways in methanogens and acetogens (Fig. 3, left panel). The same is true for the two enzymes of the carbonyl branch of the acetyl-CoA pathway, acetyl-CoA synthase (ACS) and carbon monoxide dehydrogenase (CODH) (Fig. 3, lower right). In addition, the subunits of the corrinoid iron–sulphur protein (CoFeS) that binds cobamide also appear homologous in the methanogen–acetogen comparison.

CoFeS catalyzes two cobamide-dependent methyltransfer reactions, one from the pterin cofactor to the cobalt atom of the cobamide cofactor and a second from Co to the active-site Ni atom of ACS, in some cases aided by an additional, separate methyltransferase in acetogens [52,57,112,113]. The two CoFeS subunits of *Methanococcus maripaludis* [88] and *M. thermoacetica* [87] were scored as present in a target genome if the genome had best hits for both subunits, a condition fulfilled in many cases. CoFeS showed a similar phylogenetic distribution as ACS (EC 2.3.1.169) and CODH (EC 1.2.7.4) (Fig. 4), which were previously known to be homologous across the bacterial–archaeal divide based on sequence similarity [47,114]. In methanogens, CoFeS is often part of a larger CODH/ACS complex [107]. The CoFeS subunits of *Methanococcus maripaludis* and *M. thermoacetica* show 31.4% and 37.8% amino acid sequence identity in global pairwise comparisons (methanogenic CODH/ACS complex subunits delta and gamma, respectively). The sequence homology of acetogenic and methanogenic CoFeS was previously reported [14,42]. Their structural alignments using AlphaFold-modelled structures [103,105] (Fig. S2) showed TM-scores of 0.931 (methanogen CODH/ACS complex

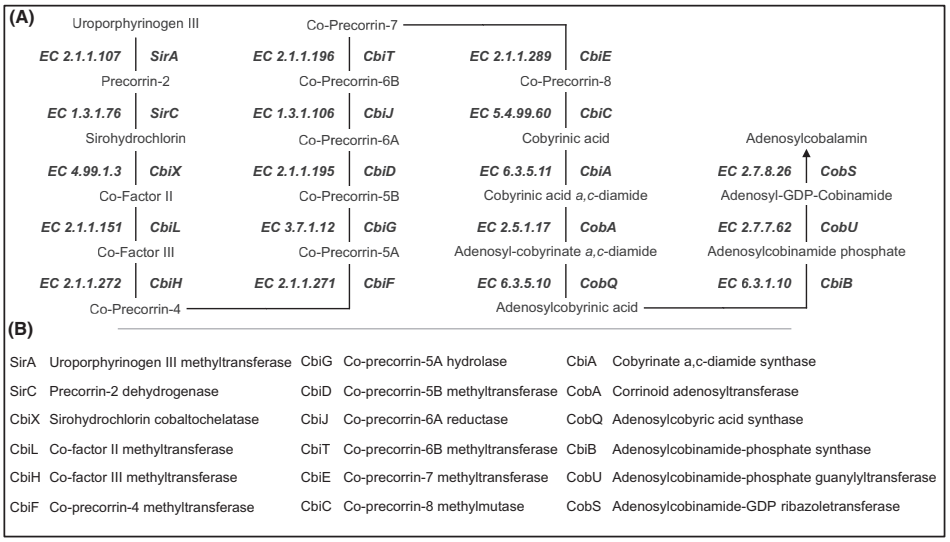
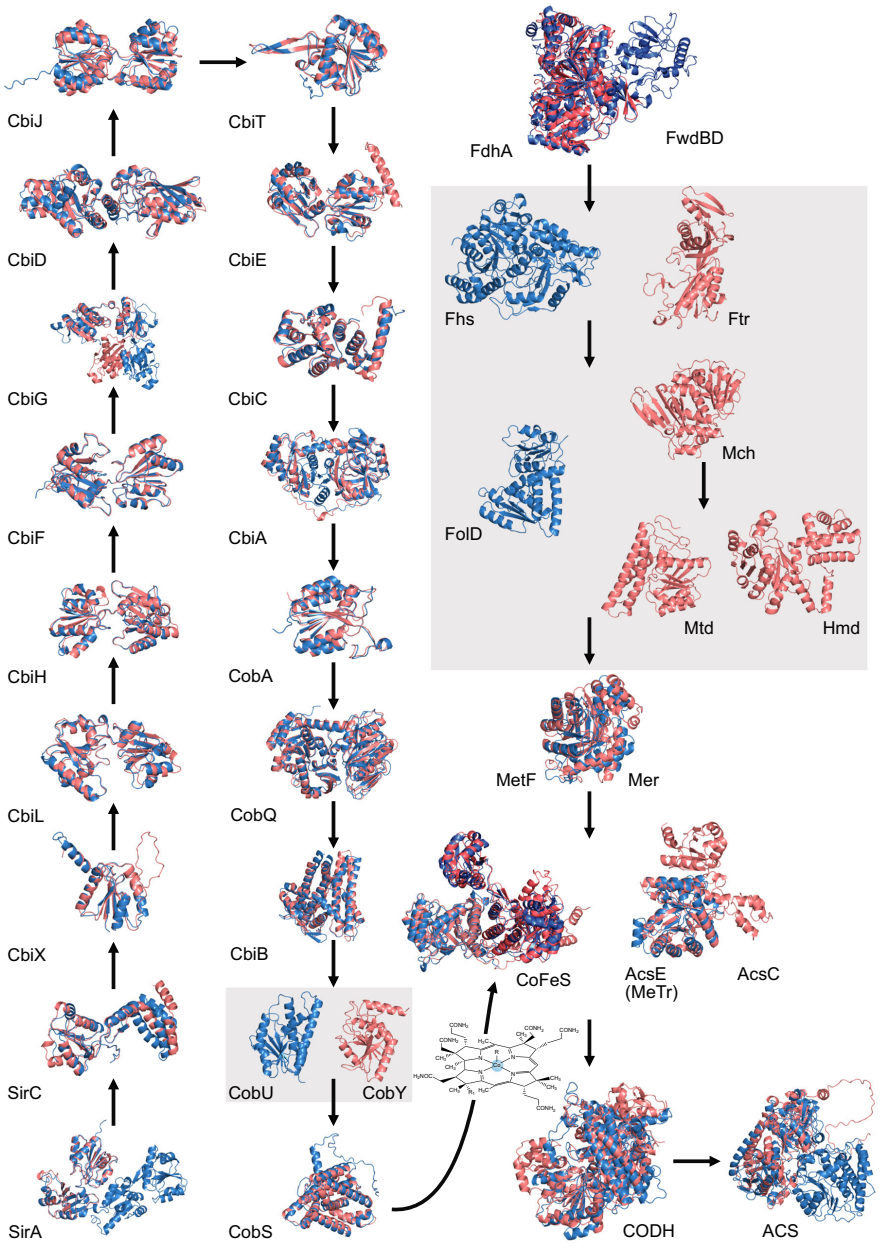


Fig. 2. The anaerobic cobalamin biosynthesis pathway. (A) The schematic shows the abbreviations and EC numbers of enzymes involved in the biosynthesis of adenosylcobalamin as found in the KEGG database, starting with the conversion of uroporphyrinogen III to precorrin-2 and the early insertion of cobalt. Lower ligand synthesis is shown in Fig. S1. (B) In previously published literature, other abbreviations like UroM [97] or SuMT [98] have been used to refer to uroporphyrinogen III methyltransferase (SirA). Abbreviations and full names of enzymes found in this article are displayed in the schematic. A short description of the steps and citations for the individual reactions is given in the Suppporting Information.

subunit delta, acetogen CoFeS small subunit) and 0.836 (methanogen CODH/ACS complex subunit gamma, acetogen CoFeS large subunit), normalized by the length of the respective acetogen protein. These scores indicate that the structures share a common fold, confirming their homology [115]. The methanogen corrinoid-iron sulphur protein CoFeS (AcsC) also shares a common fold with the methyl-H_4F:CoFeS methyltransferase (AcsE), which aids acetogen CoFeS in the methyl transfer step from the pterin to the corrin cobalt (Fig. 3), as previously reported [52].

The active corrins in the acetyl-CoA pathway of acetogens and methanogens accept methyl groups at the open Co coordination site. In adenosylcobalamin,

Fig. 3. Structurally homologous and divergent enzymes in anaerobic cobalamin synthesis and the acetyl-CoA pathway. Left panel: cobalamin synthesis pathway. Right panel: the acetyl-CoA pathway. Structures from acetogens (blue) and methanogens (salmon) are shown superimposed [103], while non-homologous structures are shown adjacent. Adjacent structures of the same colour point to non-homologous enzymatic variants in the same prokaryotic domain, such as in the case of Hmd. Hmd is the Fe-only hydrogenase of archaea that directly reduces methenyl-H_4MPT to methylene-H_4MPT, is expressed under nickel limitation, uses a unique iron-guanylylpyridinol cofactor and is not homologous to any enzymes of the pathways shown here [104]. Grey shading highlights non-homologous enzymes in methanogen-acetogen comparisons. Structures were modelled with AlphaFold [105], except for FwdBD (PDB ID: 5T5I [106]) and CODH/ACS (PDB IDs: 3CF4 and 1MJG [107,108]). Methanogen ACS and methyl synthesis were modelled based on sequences from *Methanothermobacter marburgensis*. The acetogen methyl branch is from *Moorella thermoacetica*. In cobalamin biosynthesis, methanogen sequences are from *Methanococcus maripaludis* (and *Methanosarcina barkeri* in the case of CobA, which is specific to adenosylcobalamin synthesis and missing in many methanogens; see text) and acetogen sequences from *M. thermoacetica*. In multimers, only the structurally homologous subunit was shown for clarity. TM-scores of structural alignments can be found in Table S1. For the nomenclature of enzymes involved in anaerobic cobalamin synthesis, see the legend and pannel (B) of Fig. 2. See also Svetlitchnaia *et al.* [52] for a structural comparison of CoFeS with structurally similar proteins.



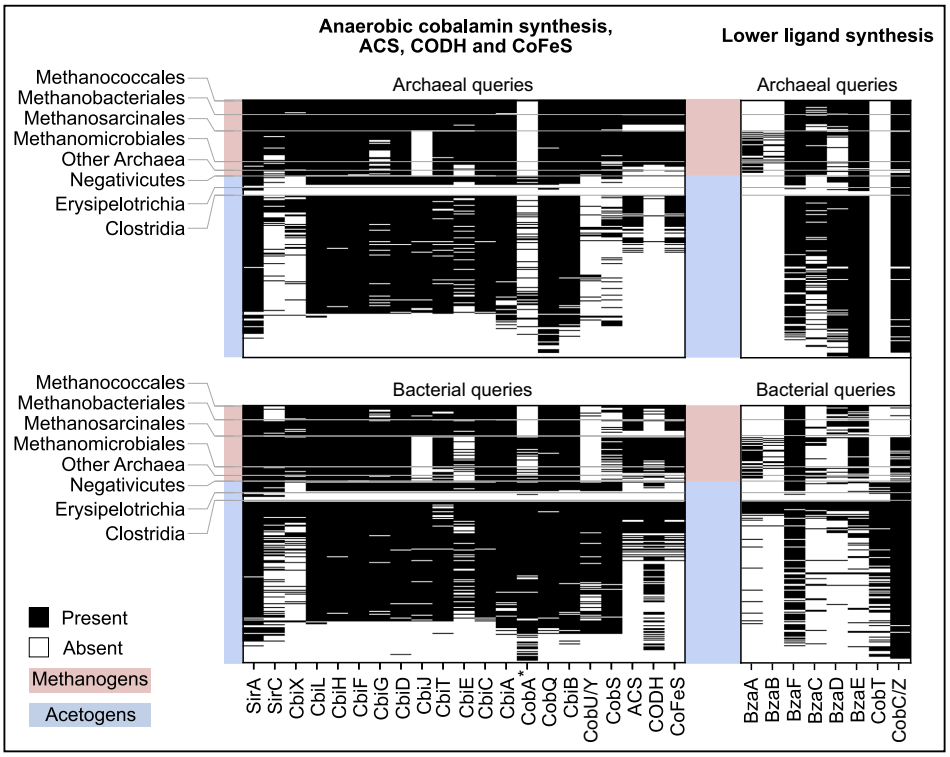


Fig. 4. Presence-absence pattern of enzymes of the anaerobic biosynthesis of cobalamin, acetyl-CoA synthase, carbon monoxide dehydrogenase, CoFeS and lower ligand synthesis enzymes in methanogens and acetogens. Best BLAST hits for all archaeal and bacterial query sequences (see Materials and methods) were scored in data sets of 80 methanogens and 189 acetogens. Analysed enzymes are displayed on the *X*-axis, and groups of methanogens (red) and acetogens (blue) are shown on the *Y*-axis. Black indicates presence and white indicates absence of the enzyme. The column (*) labelled "CobA" covers homologs of CobA, PduO and EutT. Nonorthologous alternatives of CobU (CobY) and CobC (CobZ) [36] were also included.

the central cobalt is coordinated by an upper ligand adenosyl moiety that is introduced by the corrinoid adenosyltransferase BtuR/CobA [116], which can also be replaced by two alternative enzymes, PduO [117] or EutT [118]. BtuR/CobA was generally missing in methanogen genomes from the *Methanococcales* (including our query species *Methanococcus maripaludis*, Fig. 4) and *Methanobacteriales*, which contain hydrogenotrophic species. This is consistent with a lack of adenosylcobalamin in several studied species [7]. CobU (adenosylcobinamide-phosphate guanylyltransferase) and its non-orthologous alternative CobY [36] lack structural homology across the methanogen-acetogen divide (Fig. 3) and, like CobS (adenosylcobinamide-GDP ribazoletransferase), are not widely distributed in reciprocal comparisons among acetogens and methanogens represented here (Fig. 4). Among the enzymes involved in the synthesis of the cobalamin lower ligand, only BzaF, which catalyzes the synthesis of 5-hydroxybenzimidazole (HBI) [39], is well conserved across the bacterial–archaeal divide, reflecting natural variation in lower ligand synthesis observed in methanogens and acetogens [7,8] relative to *Eubacterium limosum*, the species in which lower ligand synthesis was identified [39]. Archaeal and bacterial CobT proteins are not homologous, as previously reported [119]. The distribution of the genes for enzymes of corrin biosynthesis across

the bacterial–archaeal divide at the 25% amino acid sequence identity threshold is shown in Fig. 4.

Independent origins of enzymatic methyl synthesis in acetogens and methanogens

The methyl synthesis branch of the acetyl-CoA pathway provides a continuous supply of methyl groups for the CoFeS reaction in acetogen and methanogen carbon assimilation. The enzymes of methyl synthesis and those underpinning the cofactors synthesis, tetrahydrofolate (H_4F) in bacteria and tetrahydromethanopterin (H_4MPT) in archaea, were previously found to be non-homologous across the bacterial–archaeal divide [50] by the measure of amino acid sequence conservation. Because the methanogen enzyme formylmethanofuran dehydrogenase reveals structural homology to formate dehydrogenase of the acetogen pathway [116], which was not detected at the amino acid sequence level, we reinvestigated the potential common ancestry of acetogen and methanogen methyl synthesis pathways on the basis of structural homology. Identifiers for genes of the methyl branch of the acetyl-CoA pathway for the methanogen *Methanothermobacter marburgensis* and the acetogen *M. thermoacetica* were taken from previous studies [49,94]. Archaeal formylmethanofuran dehydrogenase (FwdBD) and bacterial formate dehydrogenase (FdhA) were not reinvestigated as structural similarity has already been shown [104]. Alignments of structures modelled with AlphaFold [103,105] showed that the reduction of the pterin-bound methylene to a methyl group is also catalyzed by enzymes that share a common fold, with a TM-score > 0.67 (Fig. 3). The electron donor for methylene-H_4MPT reductase in methanogens (Mer) is the cofactor F_{420}; the donor for the acetogen 5,10-methylene-H_4F reductase (MetF) is unclear [120].

The remaining two steps of methyl synthesis in acetogens, N^{10}-formyl-H_4F synthetase (Fhs) and 5,10-methenyl-H_4F cyclohydrolase/dehydrogenase (FolD), correspond to three steps in methanogens, catalyzed by formylmethanofuran:H_4MPT formyltransferase (Ftr), methenyl-H_4MPT cyclohydrolase (Mch) and F_{420}-dependent methylene-H_4MPT dehydrogenase (Mtd). These conversions are catalyzed by structurally non-homologous enzymes in acetogens and methanogens (TM-score < 0.35 in all cases) (Fig. 3). The non-homology of these methyl synthesis enzymes in acetogens and methanogens, their use of different redox cofactors at individual steps, in addition to the divergent syntheses of their corresponding pterin cofactors, indicates that the pathways arose independently in the acetogen and methanogen lineages, respectively [50]. This in turn suggests that the common ancestor of acetogens and methanogens was dependent upon a supply of chemically reactive methyl groups provided by the environment [47,50] and that the use of environmental methyl groups in acetyl and methane synthesis is more ancient than the pterin-dependent biochemical pathways used to synthesize the methyl groups. Such geochemically supplied reactive methyl groups could include methanol, a soluble substrate for the acetyl-CoA pathway in methanogens and acetogens [49] and a synthesis product from H_2 and CO_2 in geochemical analogues of the acetyl-CoA pathway [60], hydrothermal methyl sulphide [71,121], methyl thioacetate [122] or methyl groups bound to solid-state metal catalysts that synthesize methyl groups from H_2 and CO_2 [60,62–64].

The corrin synthesis pathways in acetogens and methanogens consist of homologous enzymes (Figs 3 and 4) that provide the cofactor for the corrin-binding methyltransferase, CoFeS, which is also homologous (Fig. S2), as are the enzymes of the acetyl synthesis branch of the acetyl-CoA pathway (CODH/ACS) (Fig. 3). In the methyl synthesis branch, the reduction of CO_2 to a formyl group (in methanogens) or formate (in acetogens) and the reduction of pterin-bound methylene to a methyl group are both steps catalyzed by structurally homologous enzymes across the acetogen-methanogen divide (Fig. 3, Fig. S2). Two steps of methyl synthesis are not homologous in the acetogen/methanogen comparison, indicating that these two steps "resulted from convergent developments" [48], after the divergence of the acetogen and methanogen lineages.

One might challenge the view that the acetyl-CoA pathway is the oldest pathway of CO_2 fixation [48–51]. However, the evidence in favour of its antiquity is unambiguous. There are seven known pathways of CO_2 fixation (the Calvin cycle, the reverse citric acid cycle, the acetyl-CoA pathway, the dicarboxylate/4-hydroxybutyrate cycle, the 3-hydroxypropionate/4-hydroxybutyrate cycle, the 3-hydroxypropionate bicycle and the glycine reductase pathway [123]), but only one of them, the acetyl-CoA pathway, occurs in both bacteria and archaea [49,124]. Moreover, only one of those seven, again only the acetyl CoA pathway, operates *in vitro* to form its main products – formate, acetate and pyruvate [49,90] – from H_2 and CO_2 without enzymes or cofactors, using only metals or minerals that occur in hydrothermal vents as catalysts [60–64,125]. None of the CO_2-fixing reactions of the other six pathways have been shown to operate *in vitro* without enzymes. In addition, the amount of pyruvate

formed via metal-catalyzed, non-enzymatic versions of the acetyl-CoA pathway can reach up to 200 μM [62], which is the physiological concentration of pyruvate in the cytosol of clostridial acetogens growing from H_2 and CO_2 via the acetyl-CoA pathway [126]. Finally, on the point of the antiquity of the acetyl CoA pathway, the same metal catalysts that allow it to unfold without ATP, enzymes, or cofactors convert 2-oxoacids to amino acids at room temperature [77], reduce NAD^+ [127] and reduce ferredoxin with H_2 [79]. The other six pathways consume ATP during CO_2 fixation [49]; the acetyl-CoA pathway is integral to ATP synthesis from H_2 and CO_2 in acetogens and methanogens [78,128]. These findings offer clear experimental evidence that the acetyl-CoA pathway is the most ancient among known CO_2-fixing pathways, and no comparable findings for any of the other six pathways have been reported.

Phylogenies trace corrin synthesis to the acetogen-methanogen common ancestor

To determine whether the corrin pathway of methanogens and acetogens traces to the last universal common ancestor, we constructed alignments and phylogenies based on genes for enzymes of anaerobic cobalamin biosynthesis and lower ligand biosynthesis (see Materials and methods) (Fig. 5). We recovered three classes of phylogenies: (a) bacterial–archaeal monophyly (nine cases: SirA, SirC, CbiX, CbiE, CbiB, CobS, BzaA, BzaB and BzaF), (b) bacterial–archaeal monophyly violated by a few recent transfers of bacterial genes into the same methanogen lineages or by transfers of archaeal genes into the acetogen lineage (seven cases: CbiJ, CbiT, CbiC, CbiA, CobA, CobQ and CobU/Y) and (c) highly interleaved phylogenies suggesting multiple LGT events across the bacterial–archaeal divide (nine cases: CbiL, CbiH, CbiF, CbiG, CbiD, BzaC, BzaD, BzaE and CobC).

The nine cases of domain monophyly trace anaerobic corrin and lower ligand synthesis to the common ancestor of bacteria and archaea. Among those, SirA and SirC are involved in siroheme biosynthesis as well, a cofactor also identified in Weiss *et al.* [47] as being present in LUCA. The seven cases of rarely violated monophyly also trace anaerobic corrin and lower ligand synthesis to the common ancestor of bacteria and archaea, yet with phylogenies punctuated by late lateral gene transfer events, indicated with asterisks in Fig. 5. In some cases, these transfers involved the import of multiple genes from acetogens into the same methanogen species. For instance, the genes of CbiJ, CbiT and CbiC were transferred from acetogens into *Methanomassiliicoccus intestinalis* (GCF_000404225.1), *Methanoplasma termitum* (GCF_000800805.1), *Methanomethylophilus alvus* (GCF_000300255.2) and methanogenic archaeon ISO4-H5 (GCF_001560915.1). A common ancestor of these four archaea acquired the bacterial corrin synthesis genes *cbiJ*, *cbiT* and *cbiC*, apparently in a single event, because the four archaea in question share a common ancestor in a species tree constructed from a concatenated alignment of 23 prokaryotic informational core genes (ICGs) (Fig. S3), constructed mainly from ribosomal proteins and aminoacyl-tRNA synthetases (Table S2). In addition, a deep separation between archaeal and bacterial CbiA/CobB proteins was also recently observed in a phylogeny comprising homologous amidases of the different tetrapyrroles present in 2070 assemblies [129], where, besides CbiA/CobB, the CfbB protein, involved in the amidation of the F_{430} cofactor of methyl-coenzyme M reductase [100,130], and the DsrN protein [131], responsible for the amidation of the siroheme cofactor present in DsrAB proteins [132], are present. This phylogeny clearly shows that the methanogenic CfbB proteins and the DsrN proteins derived from archaeal CbiA/CobB proteins and that archaeal and bacterial CbiA/CobB proteins form two large clades separated by the root. This indicates not only the presence of this protein in LUCA but also the ancestry of CbiA/CobA in relation to F_{430}. A closer inspection of the phylogeny also indicated events of interdomain CbiA/CobA transfers with 11 archaeal sequences found in four small clades inside a clade of 1879 bacterial CbiA/CobB proteins. If indeed the proteins trace back to LUCA, it is expected that, during the almost 4 billion years of evolution, several lateral gene transfers would have occurred, leading to possible synonymous replacements of the original gene by the newly acquired one, as observed in other cases [129,133,134]. Moreover, studies have indicated that an average of 42.5% of genes per species is affected by lateral gene transfer when considering recent transfers [135], and this number raises up to 81% when the oldest transfers are included [136]. Thus, cases in which only a few HGTs are identified can be considered as present in LUCA. Putative LGT events are highlighted in Fig. S3.

The remaining nine trees showed abundant interleaving of bacterial and archaeal corrin synthesis genes, indicating that the homologous genes are, in terms of structure and function, freely interchangeable across the bacterial–archaeal divide. This indicates in turn (a) conservation of function since their divergence from a common ancestral gene and (b) the presence and function of a member of that gene family in the

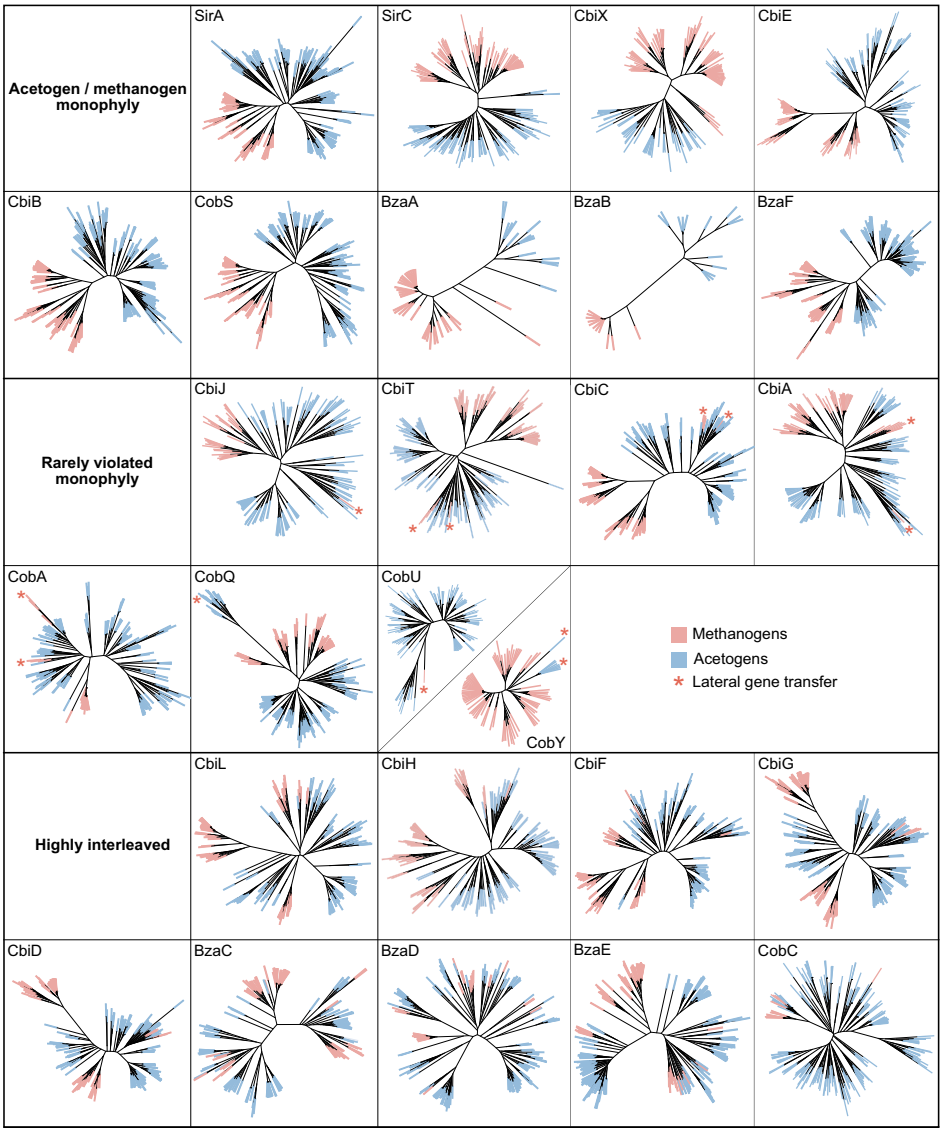


Fig. 5. Unrooted gene trees of enzymes involved in the anaerobic biosynthesis of cobalamin and lower ligand synthesis. The overview shows single-gene trees containing sequences of methanogens (red) and acetogens (blue). Calculations of trees are based on the best BLAST hits of acetogen and methanogen query sequences against genome data sets of methanogens and acetogens. The phylogenies were computed with RAXML version 8.2.12 and are grouped according to their category (monophyly, rarely violated monophyly, highly interleaved) and arranged corresponding to the order displayed in Fig. 2.

bacterial–archaeal common ancestor. Despite the interleaving of methanogens and acetogens in individual trees, the concatenated tree of corrin biosynthesis enzymes recovers domain monophyly (Fig. S4). Given the possible non-specificity of reactions of primordial enzymes, it can be envisaged that CbiE (tracing to LUCA) could have replaced (in LUCA) some of the homologous proteins that show intercalated archaeal-bacterial phylogenies, such as CbiF, CbiH and CbiL. Phylogeny and functional equivalence indicate the presence of an almost complete corrin pathway in the bacterial–archaeal common ancestor with a partial lower ligand biosynthesis.

Cobamide synthesis, a hub in biochemical pathway evolution

None of the intermediates in the 26-enzyme cobamide synthesis pathway are known to serve a biochemical function beyond their role as a pathway intermediate. This would present an evolutionary paradox, because all individual enzymes of the entire pathway would have to be functional before selection could set in – unless, of course, the domains of cobamide synthesis enzymes were recruited from simpler preexisting pathways. To see which pathways might have contributed genes to corrin evolution, we took three approaches.

First, a DIAMOND BLASTP search of cobalamin biosynthesis proteins against a dataset of 404 autotrophic core reactions [137] uncovered homologues of CbiA in F_{430} biosynthesis, homologues of CbiA and CobQ in biotin biosynthesis, and homologues of CbiG and CobQ in histidine metabolism (Table S3). Second, we mapped the sequences of the 355 protein clusters from Weiss *et al.* [47] to InterPro and compared the SUPERFAMILY, TIGR and PFAM assignments with those obtained for cobalamin biosynthesis proteins. We found that the ancient 'P-loop containing nucleoside triphosphate hydrolase' domain, which is the only domain of CobU and BtuR/CobA proteins, is present in 37 different protein families from Weiss *et al.* [47] belonging to seven different functional categories (Fig. 6). CbiT belongs to the large superfamily of *S*-adenosyl-L-methionine-dependent (SAM-dependent) methyltransferases (SSF53335) and is homologous with five protein families of Weiss *et al.* [47], including three involved in RNA modifications. CbiE shares homology with the remaining cobalamin methylases (SirA, CbiF, CbiH and CbiL), indicating ancient diversification of the family. CobQ is a member of the Class I glutamine amidotransferase-like superfamily, which includes a domain of CbiA; the second, nucleotide-binding domain of CbiA is also present in the data of Weiss *et al.* [47] in a cluster annotated as MinD. In total, 12 out of 18 cobalamin biosynthesis proteins belong to protein families that trace to LUCA [47], indicating that their ancestral module was already present in LUCA (Table S4).

Next, we looked for domains shared between cobalamin biosynthesis and domains present in 3939 sequenced genomes of cultured organisms using InterPro domain scans (Fig. 6). This revealed homologues of genes for anaerobic cobalamin biosynthesis in additional ancient pathways. Homologues of cobalamin biosynthesis are involved in amino acid synthesis (tyrosine and histidine) and in cofactor biosynthesis pathways for molybdopterin (MoCo), pyridoxal phosphate and biotin. The class III tetrapyrrole methylases CbiE, CbiF, CbiH, CbiL and SirA are homologous to diphthine synthase, 16S rRNA (cytidine(1402)-2′-O)-methyltransferase and MazG proteins. The class I methyltransferase CbiT shares homology with a large number of SAM-dependent RNA-modifying methyltransferases, which is in accordance with the literature [138]. SirC contains a very common NAD(P)-binding Rossmann-fold domain (Table S5). CobQ is related to BioD (biotin pathway), an enzyme responsible for ring formation [139] and homologous to MinD, CobB and ParA proteins [140]. No homologs were identified for CbiG, CbiJ, CbiD and CbiB. In the pathway for synthesis of the lower ligand, BzaA, BzaB and BzaF are homologous to ThiC, which catalyzes the first committed step of thiamine biosynthesis [141]. Other homologues included enzymes involved in the synthesis of F_{430}, a Ni-containing tetrapyrrole specific to the methane-forming step of methanogenesis [142] whose synthesis shares common steps with cobamide biosynthesis [38].

The shared homologies of *cob* enzymes reflect links to cofactor synthesis, amino acid synthesis and RNA modification (Fig. 6), functions that were clearly present in LUCA [47,137], indicating that the 26-step cobamide synthesis pathway was assembled from protein domains involved in ancient preexisting pathways. Cobalamin synthesis genes were, however, also recruited later in evolution, for example, in chlorophyll biosynthesis (BchE, BchR and BchQ) of photosynthetic bacteria [38]. In addition, all tetrapyrrole class II chelatases derive from a common ancestor [143,144], that evolved through duplication to give rise to $CbiX_L$, CbiK, HemH (PpfC/CpfC) and SirB (ShfC). Both *Methanococcus maripaludis* and *M. thermoacetica* use the short version of the enzyme ($CbiX_s$), having retained the primordial class II tetrapyrrole chelatase and further underscoring the common ancestry of the pathways in these organisms. Cobamide biosynthesis

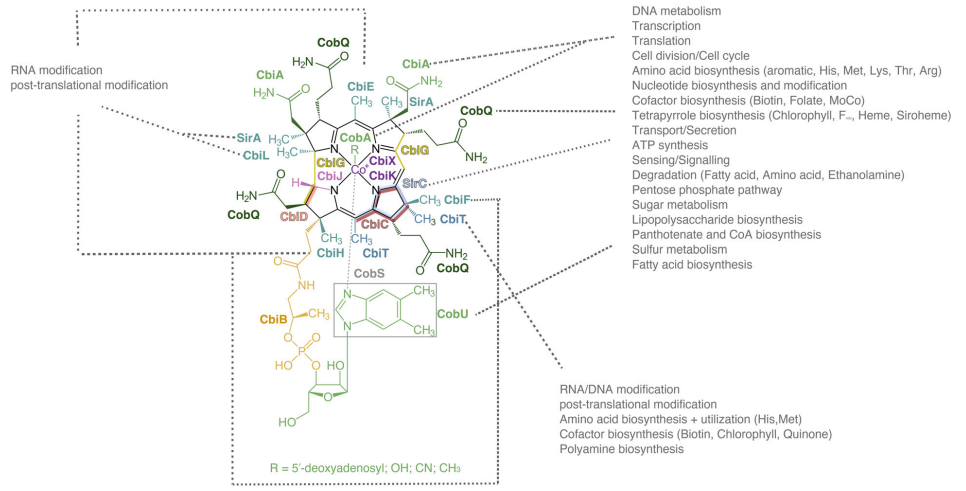


Fig. 6. Homologs of cobalamin biosynthesis enzymes in other pathways. The chemical structure of adenosylcobalamin is coloured according to the enzymes that are involved in the synthesis of the respective region of the cobalamin structure. Dashed arrows point to KEGG pathway categories linked to identified homologs of the corresponding enzymes.

represents a hub of functional domains that was assembled early in the evolution of biochemical pathways and tapped for useful activities during lineage divergence.

The acetogen–methanogen common ancestor and LUCA

The data indicate that corrin biosynthesis was present in the acetogen-methanogen common ancestor. This traces it to LUCA, because there are only three possibilities concerning the relationship between the common ancestor of bacteria and archaea and LUCA. (a) Bacteria arose from a specific lineage of archaea, making LUCA an archaeon. (b) Archaea arose from a specific lineage of bacteria, making LUCA a bacterium. (c) Neither domain arose from the other, such that LUCA was the common ancestor of bacteria and archaea sampled here. No current molecular phylogenetic studies support an origin of one prokaryotic domain from the other or one member of one prokaryotic domain from the other. Though debates exist about LUCA's organizational state [145,146], very few studies directly address the physiology of LUCA [47] and numerous studies support possibility 3 [47,147,148], which, given the homologies (Fig. 3) and phylogenies (Fig. 5) presented here, would place corrin biosynthesis in the common ancestor of bacteria and archaea, hence in LUCA.

The presence of corrin synthesis in LUCA contrasts with the lack of homologous methyl synthesis branches across the bacterial and archaeal versions of the acetyl-CoA pathway that delivers methyl groups for CoFeS (Fig. 3), suggesting that LUCA utilized geochemically supplied methyl groups [47,50]. A recent study reported that LUCA possessed a nearly complete acetyl-CoA pathway [149], although the genes they identified (in their supplemental information) were specifically for the archaeal acetyl-CoA pathway and were often present in fewer than 10% of the bacterial plus bacterial-metagenomic lineages sampled, indicating transfer of the archaeal genes to bacteria, rather than presence of the archaeal pathway in LUCA, which their method inferred. Many such archaea-to-bacteria transfers are known for the archaeal methyl synthesis pathway, starting with the O_2-dependent bacterial methylotroph *Methylobacterium extorquens*, where genes for the entire methyl synthesis pathway and methanopterin cofactor synthesis were acquired by transfer from archaea [150]. Several other studies uncovered further examples in which individual genes of the archaeal methyl synthesis and methanopterin pathways had been transferred from archaea to bacterial recipients [50,151–154], which are

often O_2-dependent methylotrophs and methanotrophs. Lateral gene transfer is a general hurdle to identifying genes that were present in LUCA [47]. The reactions of methyl synthesis in Fig. 3 are reversible. The presence of homologous enzymes generating formate (FdhA, FwdBD) could reflect formate utilization (generation of CO_2 and reductant) as reductant in early serpentinizing environments [70]; the presence of homologous enzymes for synthesis of methylene-pterins (MetF, Mer) (Fig. 3) could reflect methyl-dependent generation of methylene-H_4F and methylene-H_4MPT as C1 units for biosyntheses.

Implications for early biochemical evolution

Earlier gene-based investigations of corrin biosynthesis evolution [80,155–157] have missed the role of corrinoids in CoFeS and the acetyl-CoA pathway, even though their function in ancient pathways has long been evident [7,8,48,49]. The present findings trace corrinoids to the earliest phases of biochemical evolution, to a time when some essential biochemical functions were performed by proteins, some were performed by cofactors alone and some were still performed by solid-state metals [62,64,77,90]. Isotope data trace the ultralight isotope signature of the acetyl-CoA pathway [68] back to > 3.5 Ga [65,66], consistent with the structural (Fig. 3) and phylogenetic (Fig. 5) evidence for the age of corrins.

The present data suggest a sequence of events as outlined in Fig. 7. Under the conditions of serpentinizing hydrothermal vents, transition metals alone convert H_2 and CO_2 into formate, acetate and pyruvate in water [60,62,63,158]. Methyl synthesis and transfer can occur on solid-state catalyst surfaces (Fig. 7A). The exergonic nature of carbon and energy metabolism from H_2-dependent CO_2 fixation can, in principle, energetically support the origin of genes and proteins with the help of substrate-level phosphorylation [71,159]. Proteins, once present, can organize Fe and Ni into active sites through ligation to cysteine sulfhydryls and perform the functions of ferredoxin [79], CODH and ACS [60,61], such that acetyl and C≡O synthesis become independent of the environmental surface but remain tied to solid-state catalysts for methyl group supply from H_2 and CO_2 (Fig. 7B). In this stable intermediate configuration, enzymatic carbon metabolism remains tethered to catalysts on the Earth's crust. More ancient than enzymatic methyl synthesis, corrin synthesis could supply soluble and diffusible Co-bound methyl groups, freeing ACS and acetyl synthesis from contact with the solid state and allowing acetyl synthesis to migrate to the aqueous phase (Fig. 7C). Pterin-dependent enzymatic methyl synthesis from H_2 and CO_2 freed methyl synthesis from solid-state catalysts, involving folate in the lineage leading to acetogens and methanopterin in the lineage leading to methanogens [50] (Fig. 7D). This required the origin of additional complexes essential for methanogens and acetogens, namely, hydrogenases and flavin-based electron bifurcation [78], the precursor of which was Fe^0-dependent ferredoxin reduction with H_2 under hydrothermal conditions [79].

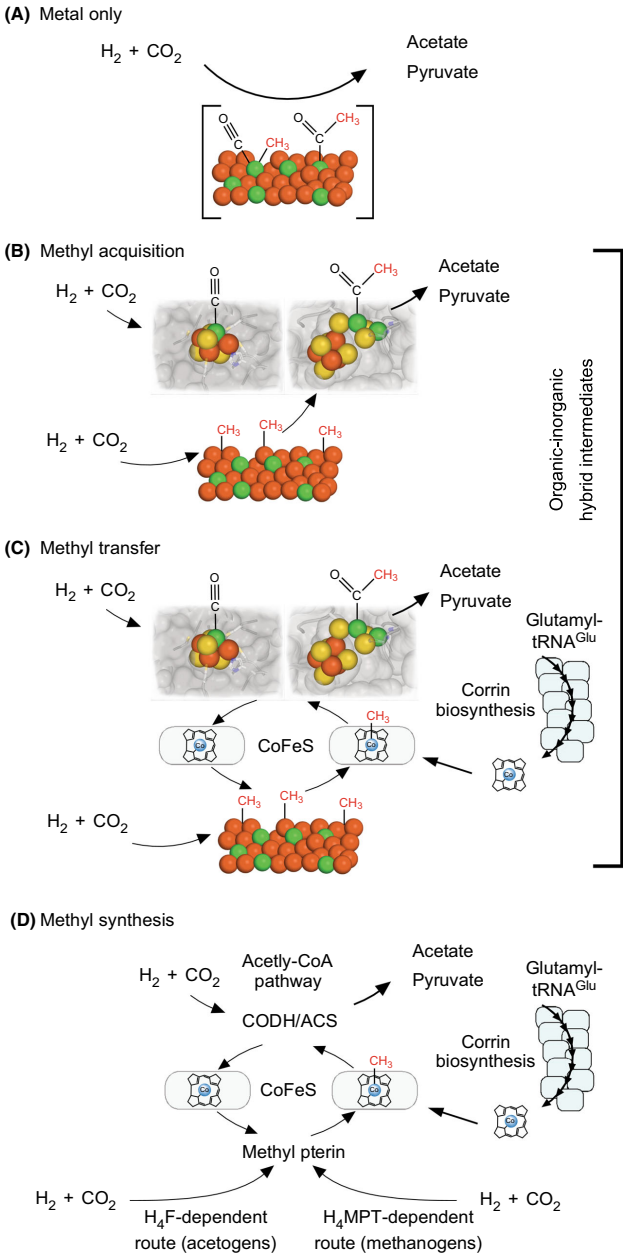
A critical observer might maintain that even enzymatic corrin synthesis is 'just too complicated' to be ancient [80], but the argument fails. Why? Corrins are far less complicated than ribosomes, with 1/2000th the molecular mass [160], and ribosome synthesis had to be in place before any of the homologous protein-coding genes for corrin biosynthesis arose. Complexity is not an argument against antiquity *per se*, as the ribosome attests. At biochemical origins, complex enzymatic pathways require functional catalytic precursors. The present findings establish a catalytic continuity between natural geochemical environments [75] and the requirement for Fe, Ni and Co in enzymes and corrin cofactors of the acetyl-CoA pathway, which might foster the identification of additional intermediate states in early biochemical evolution.

Materials and methods

Data

A dataset of 80 genomes of methanogens and 189 genomes of acetogens from the Reference Sequence database

Fig. 7. Early phases of biochemical evolution in serpentinizing hydrothermal vents. (A) Transition metals alone convert H_2 and CO_2 into acetate and pyruvate; methyl synthesis and transfer occur on the solid-state catalyst surface. (B) Still tied to solid-state catalysts for methyl supply, proteins incorporate Fe- and Ni-atoms into their active sites, allowing the synthesis of carbon monoxide and acetyl to become soluble and independent of the environmental surface. (C) The origin of corrin synthesis generates a supply of mobile, cobalt-bound methyl groups, releasing ACS from the solid state and enabling acetyl synthesis to migrate to the aqueous phase. (D) H_2-dependent methyl formation from CO_2 also migrated to the aqueous phase with the origin of enzymatic methyl synthesis after the archaeal–bacterial divergence.



(September 2016) [161] was analyzed. The methanogens belong to five archaeal orders (including "Other Archaea"), and three classes of acetogens are represented. A taxonomy report can be found in Table S6. Sequences of enzymes involved in the anaerobic biosynthesis of cobalamin (Fig. 2; based on [38]), as well as acetyl-CoA synthase (ACS), carbon monoxide dehydrogenase (CODH) and CoFeS of *M. thermoacetica* (acetogen representative) and *Methanococcus maripaludis* (methanogen representative), were acquired from the KEGG (Kyoto Encyclopedia of Genes and Genomes) [162] database in June 2023 (KEGG organisms *mtho*/*mta* and *mmp*, respectively). *M. thermoacetica* sirohydrochlorin cobaltochelatase CbiX was downloaded from GenBank (TYL17187.1). Acetogen sequences of enzymes involved in lower ligand synthesis were also acquired from the Reference Sequence database and GenBank. Methanogen query sequences were selected from BLAST results against the acetogen set of sequences. Sequences of enzymes involved in the methyl synthesis in *M. thermoacetica* [94] were acquired from the IMG/M database [163] and UniProt [164]. For *Methanothermobacter marburgensis*, sequences (as described in Ref. [49]) were downloaded from GenBank. A dataset of 3939 genomes downloaded from NCBI (November 2019) was analyzed for the presence of homologues of the biosynthesis of cobalamin but excluding the aerobic cobalamin biosynthesis pathway. All analyzed sequences, including the CobA alternatives PduO/EutT and the corresponding organisms, are listed in Table S7.

Sequence homology search in methanogens and acetogens

The representative archaeal and bacterial sequences were scored in methanogens and acetogens with DIAMOND version 2.0.1.139 [165,166]. The homology search was performed using an expectation value threshold of 1×10^{-10} and an identity threshold of 25%. For each enzyme, the best sequence hit was selected per genome for further analyses. INTERPROSCAN (version 5.66-9.8) [167] was run on the retrieved hits to obtain further functional annotations. Global sequence alignments of methanogen and acetogen CoFeS subunits were calculated based on the Needleman-Wunsch algorithm [168].

Synteny analysis

Genomes containing best-hit sequences were used for the synteny analysis, which was used in phylogenetic analysis. The best hits were mapped to their gene location files (feature tables), and a syntenic neighbourhood was defined using a window of two upstream and two downstream proteins in the neighbourhood of the proteins of interest (same chromosome). The identified syntenic neighbourhood for each best hit was extracted and mapped onto the phylogenies with the corresponding accessions, protein types, PFAM annotations and gene descriptions for a detailed phylogenetic analysis and better distinction of paralogous sequences.

Expanded sequence homology searches

DIAMOND (version 2.18) was also run against the dataset from Wimmer *et al.* [137] with an identity cutoff of 25% and *E*-value cutoffs of 10^{-10}, 10^{-8} and 10^{-5}, and the best hits were further analyzed. In addition, INTERPROSCAN (version 5.66-9.8) [167] was run on the dataset of 3939 genomes and on the sequences from the 355 protein families mapped to LUCA from Weiss *et al.* [47]. All hits that shared the same SUPERFAMILY annotations as cobalamin biosynthesis enzymes were further analysed.

Computation of multiple sequence alignments and phylogenetic trees

Gene trees containing sequences of methanogens and acetogens for 26 analysed enzymes involved in the anaerobic biosynthesis of cobalamin and lower ligand synthesis were computed based on the best hit selection of the homology search or from the synteny analysis. For that, multiple sequence alignments were made using MAFFT L-INS-I version 7.471 [169] and unrooted gene trees were reconstructed with RAXML version 8.2.12 [170] using the PROTCATWAG model and a random starting seed (*12345*). The same pipeline was used to compute a concatenated tree of methanogens and acetogens with best hits for ≥ 17 enzymes. Single sequence alignments of the analysed enzymes were combined to obtain a large alignment for tree reconstruction. MAFFT and RAXML were used to calculate a species tree of all 80 methanogens present in the dataset based on a concatenated alignment combining sequences of prokaryotic informational core gene families (ICGs; [93]), which were determined by using the verticality value [171], the number of genomes and the taxonomic distribution of a protein family. All 80 methanogens are represented by at least one sequence in 23 different prokaryotic informational core gene families (Table S2). The trees were visually modified with the use of the INTERACTIVE TREE OF LIFE [172].

Structural alignments of enzymes

In order to compare the structures of methanogenic and acetogenic enzymes involved in anaerobic cobalamin synthesis and methyl synthesis in the acetyl-CoA pathway, structures were either obtained from the Protein Data Bank (PDB, [173]) or modelled using AlphaFold version 2.2.0 [105]. Structural alignments and TM-scores were determined using US-ALIGN version 20230609 [103]. Alignments were visually inspected, and figures were prepared using PYMOL version 2.5.7 (The PyMOL Molecular Graphics

System, Version 2.5.7; Schrödinger, LLC (http://www.pymol.org/pymol)]). In those cases when the structures overlapped over one or more domains, but not the full-length protein (usually reflected in significant differences in the TM scores normalized by the lengths of the two proteins), the overlapping domains were separately aligned to obtain a more accurate TM-score. This allowed us to determine structural homology for the superimposed domains, in line with the modular nature of protein evolution.

Acknowledgements

We thank Katharina Trost for calculating and providing informational core genes. Computational infrastructure and support were provided by the Centre for Information and Media Technology at Heinrich Heine University Düsseldorf and by the Life Science Computer Cluster (LiSC) at the University of Vienna. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 101018894). For funding, WFM thanks the ERC (101018894), the Deutsche Forschungsgemeinschaft (MA 1426/21-1) and the Volkswagen Foundation (Grant 96_742), For funding, FLS thanks the Wiener Wissenschafts-, Forschungs- und Technologiefonds (grant agreement VRG15-007) and the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation program (grant agreement 803768).

Conflict of interest

The authors declare no conflict of interest.

Author contributions

WFM designed research; LDM, VK, MK, NM, WFM and FLS performed research; LDM, VK, MK, NM, WFM and FLS analyzed data; and LDM, VK, MK, NM, WFM and FLS wrote the paper.

Data availability statement

Alignments, phylogenies, commands, synteny script and the dataset of 3939 genomes are available at Figshare: https://doi.org/10.6084/m9.figshare.25765518.

References

1 Matthews RG, Koutmos M & Datta S (2008) Cobalamin-dependent and cobamide-dependent methyltransferases. *Curr Opin Struct Biol* **18**, 658–666.

2 Crofts TS, Seth EC, Hazra AB & Taga ME (2013) Cobamide structure depends on both lower ligand availability and CobT substrate specificity. *Chem Biol* **20**, 1265–1274.

3 Cheong CG, Escalante-Semerena JC & Rayment I (2001) Structural investigation of the biosynthesis of alternative lower ligands for cobamides by nicotinate mononucleotide: 5,6-dimethylbenzimidazole phosphoribosyltransferase from *Salmonella enterica*. *J Biol Chem* **276**, 37612–37620.

4 Rondon MR, Trzebiatowski JR & Escalante-Semerena JC (1997) Biochemistry and molecular genetics of cobalamin biosynthesis. *Prog Nucleic Acid Res Mol Biol* **56**, 347–384.

5 Renz P, Blickle S & Friedrich W (1987) Two new vitamin B-12 factors from sewage sludge containing 2-methylsulfinyladenine or 2-methylsulfonyladenine as base component. *Eur J Biochem* **163**, 175–179.

6 Kräutler B, Kohler HPE & Stupperich E (1988) 5′-Methylbenzimidazolyl-cobamides are the corrinoids from some sulfate-reducing and sulfur-metabolizing bacteria. *Eur J Biochem* **176**, 461–469.

7 Stupperich E & Kräutler B (1988) Pseudo vitamin B_{12} or 5-hydroxybenzimidazolyl-cobamide are the corrinoids found in methanogenic bacteria. *Arch Microbiol* **149**, 268–271.

8 Stupperich E, Eisinger HJ & Kräutler B (1988) Diversity of corrinoids in acetogenic bacteria. P-cresolylcobamide from *Sporomusa ovata*, 5-methoxy-6-methylbenzimidazolylcobamide from *Clostridium formicoaceticum* and vitamin B_{12} from *Acetobacterium woodii*. *Eur J Biochem* **172**, 459–464.

9 Mathur Y & Hazra AB (2022) Methylations in vitamin B_{12} biosynthesis and catalysis. *Curr Opin Struct Biol* **77**, 102490.

10 Gruber K, Puffer B & Kräutler B (2011) Vitamin B_{12}-derivatives—enzyme cofactors and ligands of proteins and nucleic acids. *Chem Soc Rev* **40**, 4346–4363.

11 Buckel W & Golding B (1996) Glutamate and 2-methyleneglutarate mutase: from microbial curiosities to paradigms for coenzyme B_{12}-dependent enzymes. *Chem Soc Rev* **25**, 329–337.

12 Jordan A, Torrents E, Jeanson C, Eliasson R, Hellman U, Wernstedt C, Barbé J, Gibert I & Reichard P (1997) B_{12}-dependent ribonucleotide reductases from deeply rooted eubacteria are structurally related to the aerobic enzyme from *Escherichia coli*. *Proc Natl Acad Sci USA* **94**, 13487–13492.

13 Harder J (1993) Ribonucleotide reductases and their occurrence in microorganisms: a link to the RNA/DNA transition. *FEMS Microbiol Rev* **12**, 273–292.

14 Tauer A & Benner SA (1997) The B12-dependent ribonucleotide reductase from the archaebacterium *Thermoplasma acidophila*: an evolutionary solution to

the ribonucleotide reductase conundrum. *Proc Natl Acad Sci USA* **94**, 53–58.
15 Banerjee RV & Matthews RG (1990) Cobalamin-dependent methionine synthase. *FASEB J* **4**, 1450–1459.
16 Matthews RG (2009) Cobalamin- and corrinoid-dependent enzymes. In Metal-Carbon Bonds in Enzymes and Cofactors (Sigel A, Sigel H & Sigel RKO, eds), pp. 53–114. The Royal Society of Chemistry, London.
17 Dowling DP, Miles ZD, Köhrer C, Maiocco SJ, Elliott SJ, Bandarian V & Drennan CL (2016) Molecular basis of cobalamin-dependent RNA modification. *Nucleic Acids Res* **44**, 9965–9976.
18 Gough SP, Petersen BO & Duus JO (2000) Anaerobic chlorophyll isocyclic ring formation in *Rhodobacter capsulatus* requires a cobalamin cofactor. *Proc Natl Acad Sci USA* **97**, 6908–6913.
19 Schmerk CL, Welander PV, Hamad MA, Bain KL, Bernards MA, Summons RE & Valvano MA (2015) Elucidation of the *Burkholderia cenocepacia* hopanoid biosynthesis pathway uncovers functions for conserved proteins in hopanoid-producing bacteria. *Environ Microbiol* **17**, 735–750.
20 Parent A, Guillot A, Benjdia A, Chartier G, Leprince J & Berteau O (2016) The B_{12}-radical SAM enzyme PoyC catalyzes valine C_{β}-methylation during polytheonamide biosynthesis. *J Am Chem Soc* **138**, 15515–15518.
21 Bridwell-Rabb J, Li B & Drennan CL (2022) Cobalamin-dependent radical *S*-adenosylmethionine enzymes: capitalizing on old motifs for new functions. *ACS Bio Med Chem Au* **2**, 173–186.
22 Wang Y & Begley TP (2020) Mechanistic studies on CysS – a vitamin B_{12}-dependent radical SAM methyltransferase involved in the biosynthesis of the tert-butyl group of cystobactamid. *J Am Chem Soc* **142**, 9944–9954.
23 Banerjee R & Ragsdale SW (2003) The many faces of vitamin B12: catalysis by cobalamin-dependent enzymes. *Annu Rev Biochem* **72**, 209–247.
24 Zelder O, Beatrix B, Leutbecher U & Buckel W (1994) Characterization of the coenzyme-B12-dependent glutamate mutase from *Clostridium cochlearium* produced in *Escherichia coli*. *Eur J Biochem* **226**, 577–585.
25 Kreimeyer A, Perret A, Lechaplais C, Vallenet D, Medigue C, Salanoubat M & Weissenbach J (2007) Identification of the last unknown genes in the fermentation pathway of lysine. *J Biol Chem* **282**, 7191–7197.
26 Mancia F, Keep NH, Nakagawa A, Leadlay PF, McSweeney S, Rasmussen B, Bö Secke P, Diat O & Evans PR (1996) How coenzyme B12 radicals are generated: the crystal structure of methylmalonyl-coenzyme A mutase at 2 A resolution. *Structure* **4**, 339–350.
27 Takahashi-Iñiguez T, García-Hernandez E, Arreguín-Espinosa R & Flores ME (2012) Role of vitamin B12 on methylmalonyl-CoA mutase activity. *J Zhejiang Univ Sci B* **13**, 423–437.
28 Kaplan BH & Stadtman ER (1968) Ethanolamine deaminase, a cobamide coenzyme-dependent enzyme. I. Purification, assay, and properties of the enzyme. *J Biol Chem* **243**, 1787–1793.
29 Kofoid E, Rappleye C, Stojiljkovic I & Roth J (1999) The 17-gene ethanolamine (eut) operon of *Salmonella typhimurium* encodes five homologues of carboxysome shell proteins. *J Bacteriol* **181**, 5317–5329.
30 Beatrix B, Zelder O, Linder D & Buckel W (1994) Cloning, sequencing and expression of the gene encoding the coenzyme B12-dependent 2-methyleneglutarate mutase from *Clostridium barkeri* in *Escherichia coli*. *Eur J Biochem* **221**, 101–109.
31 Forage RG & Foster MA (1982) Glycerol fermentation in *Klebsiella pneumoniae*: functions of the coenzyme B12-dependent glycerol and diol dehydratases. *J Bacteriol* **149**, 413–419.
32 Toraya T (2002) Enzymatic radical catalysis: coenzyme B_{12}-dependent diol dehydratase. *Chem Rec* **2**, 352–366.
33 Seyfried M, Daniel R & Gottschalk G (1996) Cloning, sequencing, and overexpression of the genes encoding coenzyme B_{12}-dependent glycerol dehydratase of *Citrobacter freundii*. *J Bacteriol* **178**, 5793–5796.
34 Sokolovskaya OM, Shelton AN & Taga ME (2020) Sharing vitamins: cobamides unveil microbial interactions. *Science* **369**, eaba0165.
35 Shelton AN, Seth EC, Mok KC, Han AW, Jackson SN, Haft DR & Taga ME (2019) Uneven distribution of cobamide biosynthesis and dependence in bacteria predicted by comparative genomics. *ISME J* **13**, 789–804.
36 Rodionov DA, Vitreshak AG, Mironov AA & Gelfand MS (2003) Comparative genomics of the vitamin B12 metabolism and regulation in prokaryotes. *J Biol Chem* **278**, 41148–41159.
37 Zhang Y, Rodionov DA, Gelfand MS & Gladyshev VN (2009) Comparative genomic analyses of nickel, cobalt and vitamin B12 utilization. *BMC Genomics* **10**, 78.
38 Bryant DA, Hunter CN & Warren MJ (2020) Biosynthesis of the modified tetrapyrroles – the pigments of life. *J Biol Chem* **295**, 6888–6925.
39 Hazra AB, Han AW, Mehta AP, Mok KC, Osadchiy W, Begley TP & Taga ME (2015) Anaerobic biosynthesis of the lower ligand of vitamin B_{12}. *Proc Natl Acad Sci USA* **112**, 10792–10797.
40 Lawrence JG & Roth JR (1996) Evolution of coenzyme B_{12} synthesis among enteric bacteria: evidence for loss and reacquisition of a multigene complex. *Genetics* **142**, 11–24.

41 Gómez-Consarnau L, Sadcheva R, Gifford SM, Cutter LS, Fuhrman JA, Sanudo-Wilhelmy SA & Moran MA (2018) Mosaic patterns of B-vitamin synthesis and utilization in a natural marine microbial community. *Environ Microbiol* **20**, 2809–2823.

42 Bunbury F, Helliwell KE, Mehrshani P, Davey MP, Salmon DL, Holzer A, Smirnoff N & Smith AG (2020) Responses of a newly evolved auxotroph of *Chlamydomonas* to B_{12} deprivation. *Plant Physiol* **183**, 167–178.

43 Torrents E, Trevisiol C, Rotte C & Hellman U (2006) *Euglena gracilis* ribonucleotide reductase: the eukaryote class II enzyme and the possible antiquity of eukaryote B_{12} dependence. *J Biol Chem* **281**, 5604–5611.

44 Kolhouse JF & Allen RH (1977) Recognition of two intracellular cobalamin binding proteins and their identification as methylmalonyl-CoA mutase and methionine synthetase. *Proc Natl Acad Sci USA* **74**, 921–925.

45 Eschenmoser A (1988) Vitamin B_{12}: experiments concerning the origin of its molecular structure. *Angew Chem Int Ed* **27**, 5–39.

46 Decker K, Jungermann K & Thauer RK (1970) Energy production in anaerobic organisms. *Angew Chem Int Ed* **9**, 138–158.

47 Weiss M, 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.

48 Fuchs G & Stupperich E (1985) Evolution of autotrophic CO_2 fixation. In Evolution of Prokaryotes (Schleifer KH & Stackebrandt E, eds), pp. 235–251. Academic Press, London.

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

50 Sousa FL & Martin WF (2014) Biochemical fossils of the ancient transition from geoenergetics to bioenergetics in prokaryotic one carbon compound metabolism. *Biochim Biophys Acta* **1837**, 964–981.

51 Martin WF (2020) Older than genes: the acetyl-CoA pathway and origins. *Front Microbiol* **11**, 817.

52 Svetlitchnaia T, Svetlichnyi V, Meyer O & Dobbek H (2006) Structural insights into methyltransfer reactions of a corrinoid iron-sulfur protein involved in acetyl-CoA synthesis. *Proc Natl Acad Sci USA* **103**, 14331–14336.

53 Can M, Abernathy MJ, Wiley S, Griffith C, James CD, Xiong J, Guo Y, Hoffman BM, Ragsdale SW & Sarangi R (2023) Characterization of methyl- and acetyl-Ni intermediates in acetyl CoA synthase formed during anaerobic CO_2 and CO fixation. *J Am Chem Soc* **145**, 13696–13708.

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

55 Biester A, Grahame DA & Drennan CL (2024) Capturing a methanogenic carbon monoxide dehydrogenase/acetyl-CoA synthase complex via cryogenic electron microscopy. *Proc Natl Acad Sci USA* **121**, e2410995121.

56 Ruickoldt J, Basak Y, Domnik L, Jeoung JH & Dobbek H (2022) On the kinetics of CO_2 reduction by Ni, Fe-CO dehydrogenases. *ACS Catal* **12**, 13131–13142.

57 Goetzl S, Jeoung JH, Hennig SE & Dobbek H (2011) Structural basis for electron and methyl-group transfer in a methyltransferase system operating in the reductive acetyl-CoA pathway. *J Mol Biol* **411**, 96–109.

58 Maden BE (2000) Tetrahydrofolate and tetrahydromethanopterin compared: functionally distinct carriers in C1 metabolism. *Biochem J* **350**, 609–629.

59 Ragsdale SW, Lindahl PA & Münck E (1987) Mössbauer, EPR, and optical studies of the corrinoid/iron-sulfur protein involved in the synthesis of acetyl coenzyme A by *Clostridium thermoaceticum*. *J Biol Chem* **262**, 14289–14297.

60 Preiner M, Igarashi K, Muchowska KB, Yu M, Varma SJ, Kleinermanns K, Nobu MK, Kamagata Y, Tüysüz H, Moran J *et al.* (2020) A hydrogen-dependent geochemical analogue of primordial carbon and energy metabolism. *Nat Ecol Evol* **4**, 534–542.

61 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.

62 Beyazay T, Ochoa-Hernandez C, Song Y, Belthle KS, Martin WF & Tüysüz H (2023) Influence of composition of nickel-iron nanoparticles for abiotic CO_2 conversion to early prebiotic organics. *Angew Chem Int Ed Engl* **62**, e202218189.

63 Song Y, Beyazay T & Tüysüz H (2023) Effect of alkali- and alkaline-earth-metal promoters on silica-supported Co-Fe alloy for autocatalytic CO_2 fixation. *Angew Chem Int Ed Engl* **63**, e202316110.

64 Belthle KS, Martin WF & Tüysüz H (2024) Synergistic effects of silica-supported iron-cobalt catalysts for CO_2 reduction to prebiotic organics. *ChemCatChem* **16**, e202301218.

65 Ueno Y, Yamada K, Yoshida N, Maruyama S & Isozaki Y (2006) Evidence from fluid inclusions for microbial methanogenesis in the early Archaean era. *Nature* **440**, 516–519.

66 Tashiro T, Ishida A, Hori M, Igisu M, Koike M, Mejean 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.

67 Mojzsis SJ, Arrhenius G, McKeegan KD, Harrison TM, Nutman AP & Friend CRL (1996) Evidence for life on Earth before 3,800 million years ago. *Nature* **384**, 55–59.

68 Blaser MB, Dreisbach LK & Conrad R (2013) Carbon isotope fractionation of 11 acetogenic strains grown on H_2 and CO_2. *Appl Environ Microbiol* **79**, 1787–1794.

69 Mei R, Kaneko M, Imachi H & Nobu MK (2023) The origin and evolution of methanogenesis and Archaea are intertwined. *PNAS Nexus* **2**, pgad023.

70 Colman DR, Kraus EA, Thieringer PH, Rempfert K, Templeton AS, Spear JR & Boyd ES (2022) Deep-branching acetogens in serpentinized subsurface fluids of Oman. *Proc Natl Acad Sci USA* **119**, e2206845119.

71 Martin WF & Russell MJ (2007) On the origin of biochemistry at an alkaline hydrothermal vent. *Philos Trans R Soc Lond B Biol Sci* **362**, 1887–1925.

72 Schöne C, Poehlein A, Jehmlich N, Adlung N, Daniel R, von Bergen M, Scheller S & Rother M (2022) Deconstructing *Methanosarcina acetivorans* into an acetogenic archaeon. *Proc Natl Acad Sci USA* **119**, e2113853119.

73 Lang SQ, Butterfield DA, Schulte M, Kelley DS & Lilley MD (2010) Elevated concentrations of formate, acetate and dissolved organic carbon found at the Lost City hydrothermal field. *Geochim Cosmochim Acta* **74**, 941–952.

74 Etiope G (2017) Abiotic methane in continental serpentinization sites: an overview. *Procedia Earth Planet Sci* **17**, 9–12.

75 Schwander L, Brabender M, Mrnjavac N, Wimmer JLE, Preiner M & Martin WF (2023) Serpentinization as the source of energy, electrons, organics, catalysts, nutrients and pH gradients for the origin of LUCA and life. *Front Microbiol* **14**, 1257597.

76 Mrnjavac N, Wimmer JLE, Brabender M, Schwander L & Martin WF (2023) The moon-forming impact and the autotrophic origin of life. *ChemPlusChem* **88**, e202300270.

77 Kaur H, Rauscher SA, Werner E, Song Y, Yi J, Kazöne W, Martin WF, Tüysüz H & Moran J (2024) A prebiotic Krebs cycle analog generates amino acids with H_2 and NH_3 over nickel. *Chem* **10**, 1528–1540.

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

79 Brabender M, Henriques-Pereira DP, Mrnjavac N, Schlikker ML, Kimura ZI, Sucharitakul J, Kleinermanns K, Tüysüz H, Buckel W, Preiner M *et al.* (2024) Ferredoxin reduction by hydrogen with iron functions as an evolutionary precursor of flavin-based electron bifurcation. *Proc Natl Acad Sci USA* **121**, e2318969121.

80 Holliday GL, Thornton JM, Marquet A, Smith AG, Rebeille F, Mendel R, Schubert HL, Lawrence AD & Warren MJ (2007) Evolution of enzymes and pathways for the biosynthesis of cofactors. *Nat Prod Rep* **24**, 972–987.

81 Warren MJ, Raux E, Schubert HL & Escalante-Semerena JC (2002) The biosynthesis of adenosylcobalamin (vitamin B_{12}). *Nat Prod Rep* **19**, 390–412.

82 Martin WF & Sousa FL (2015) Early microbial evolution: the age of anaerobes. *Cold Spring Harb Perspect Biol* **8**, a018127.

83 Wada K, Harada J, Yaeda Y, Tamiaki H, Oh-Oka H & Fukuyama K (2007) Crystal structures of CbiL, a methyltransferase involved in anaerobic vitamin B biosynthesis, and CbiL in complex with S-adenosylhomocysteine – implications for the reaction mechanism. *FEBS J* **274**, 563–573.

84 Frank S, Brindley AA, Deery E, Heathcote P, Lawrence AD, Leech HK, Pickersgill RW & Warren MJ (2005) Anaerobic synthesis of vitamin B_{12}: characterization of the early steps in the pathway. *Biochem Soc Trans* **33**, 811–814.

85 Broderick JB, Duffus DR, Duschene KS & Shepard EM (2014) Radical S-adenosylmethionine enzymes. *Chem Rev* **114**, 4229–4317.

86 Aziz I, Kayastha K, Kaltwasser S, Vonck J, Welsch S, Murphy BJ, Kahnt J, Wu D, Wagner T, Shima S *et al.* (2024) Structural and mechanistic basis of the central energy-converting methyltransferase complex of methanogenesis. *Proc Natl Acad Sci USA* **121**, e2315568121.

87 Lu WP, Schiau I, Cunningham JR & Ragsdale SW (1993) Sequence and expression of the gene encoding the corrinoid/iron-sulfur protein from *Clostridium thermoaceticum* and reconstitution of the recombinant protein to full activity. *J Biol Chem* **268**, 5605–5614.

88 Maupin-Furlow J & Ferry JG (1996) Characterization of the cdhD and cdhE genes encoding subunits of the corrinoid/iron-sulfur enzyme of the CO dehydrogenase complex from *Methanosarcina thermophila*. *J Bacteriol* **178**, 340–346.

89 Slotznick SP, Johnson JE, Rasmussen B, Raub TD, Webb SM, Zi JW, Kirschvink JL & Fischer WW (2022) Reexamination of 2.5-Ga "whiff" of oxygen interval points to anoxic ocean before GOE. *Sci Adv* **8**, eabj7190.

90 Mrnjavac N, Schwander L, Brabender M & Martin WF (2024) Chemical antiquity in metabolism. *Acc Chem Res* **57**, 2267–2278.

91 Ragsdale SW (2003) Pyruvate ferredoxin oxidoreductase and its radical intermediate. *Chem Rev* **103**, 2333–2346.

92 Lu Z & Imlay JA (2021) When anaerobes encounter oxygen: mechanisms of oxygen toxicity, tolerance and defense. *Nat Rev Microbiol* **19**, 774–785.

93 Mrnjavac N, Schwander L, Brabender M & Martin WF (2024) The radical impact of oxygen on prokaryotic evolution—enzyme inhibition first, uninhibited essential biosyntheses second, aerobic respiration third. *FEBS Lett* **598**, 1692–1714.

94 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.

95 Hendrickson EL, Kaul R, Zhou Y, Bovee D, Chapman P, Chung J, Conway de Macario E, Dodsworth JA, Gillett W, Graham DE *et al*. (2004) Complete genome sequence of the genetically tractable hydrogenotrophic methanogen *Methanococcus maripaludis*. *J Bacteriol* **186**, 6956–6969.

96 Dailey HA, Dailey TA, Gerdes S, Jahn D, Jahn M, O'Brian MR & Warren MJ (2017) Prokaryotic heme biosynthesis: multiple pathways to a common essential product. *Microbiol Mol Biol Rev* **81**, e00048-16.

97 Videira MAM, Lobo SAL, Sousa FL & Saraiva LM (2020) Identification of the sirohaem biosynthesis pathway in *Staphylococcus aureus*. *FEBS J* **287**, 1537–1553.

98 Dailey HA & Medlock AE (2022) A primer on heme biosynthesis. *Biol Chem* **403**, 985–1003.

99 Stroupe ME, Leech HK, Daniels DS, Warren MJ & Getzoff ED (2003) CysG structure reveals tetrapyrrole-binding features and novel regulation of siroheme biosynthesis. *Nat Struct Biol* **10**, 1064–1073.

100 Zheng K, Ngo PD, Owens VL, Yang XP & Mansoorabadi SO (2016) The biosynthetic pathway of coenzyme F430 in methanogenic and methanotrophic archaea. *Science* **354**, 339–342 (published correction appears in *Nature* 2017, 545, 40).

101 Moore SJ, Sowa ST, Schuchardt C, Deery E, Lawrence AD, Ramos JV, Billig S, Birkenmeyer C, Chivers PT, Howard MJ *et al*. (2017) Corrigendum: elucidation of the biosynthesis of the methane catalyst coenzyme F_{430}. *Nature* **545**, 116.

102 Buan NR, Rehfeld K & Escalante-Semerena JC (2006) Studies of the CobA-type ATP:Co(I)rrinoid adenosyltransferase enzyme of *Methanosarcina mazei* strain Gö1. *J Bacteriol* **188**, 3543–3550.

103 Zhang C, Shine M, Pyle AM & Zhang Y (2022) US-align: universal structure alignments of proteins, nucleic acids, and macromolecular complexes. *Nat Methods* **19**, 1109–1115.

104 Shima S, Huang G, Wagner T & Ermler U (2020) Structural basis of hydrogenotrophic methanogenesis. *Annu Rev Microbiol* **74**, 713–733.

105 Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A *et al*. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589.

106 Wagner T, Ermler U & Shima S (2016) The methanogenic CO_2 reducing-and-fixing enzyme is bifunctional and contains 46 [4Fe-4S] clusters. *Science* **354**, 114–117.

107 Gong W, Hao B, Wei Z, Ferguson DJ, Tallant T, Krzycki JA & Chan MK (2008) Structure of the $\alpha_2\varepsilon_2$ Ni-dependent CO dehydrogenase component of the *Methanosarcina barkeri* acetyl-CoA decarbonylase/synthase complex. *Proc Natl Acad Sci USA* **105**, 9558–9563.

108 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.

109 Gagnon DM, Stich TA, Mehta AP, Abdelwahed SH, Begley TP & Britt RD (2018) An aminoimidazole radical intermediate in the anaerobic biosynthesis of the 5,6-dimethylbenzimidazole ligand to vitamin B_{12}. *J Am Chem Soc* **140**, 12798–12807.

110 Wienhausen G, Dlugosch L, Jarling R, Wilkes H, Giebel HA & Simon M (2022) Availability of vitamin B_{12} and its lower ligand intermediate α-ribazole impact prokaryotic and protist communities in oceanic systems. *ISME J* **16**, 2002–2014.

111 Anderson PJ, Lango J, Carkeet C, Britten A, Kräutler B, Hammock BD & Roth JR (2008) One pathway can incorporate either adenine or dimethylbenzimidazole as an alpha-axial ligand of B_{12} cofactors in *Salmonella enterica*. *J Bacteriol* **190**, 1160–1171.

112 Doukov TI, Seravalli J, Stezowski JJ & Ragsdale SW (2000) Crystal structure of a methyltetrahydrofolate- and corrinoid-dependent methyltransferase. *Structure* **8**, 817–830.

113 Hu SI, Pezacka E & Wood HG (1984) Acetate synthesis from carbon monoxide by *Clostridium thermoaceticum*. Purification of the corrinoid protein. *J Biol Chem* **259**, 8892–8897.

114 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.

115 Xu J & Zhang Y (2010) How significant is a protein structure similarity with TM-score = 0.5? *Bioinformatics* **26**, 889–895.

116 Vitols E, Walker GA & Huennekens FM (1966) Enzymatic conversion of vitamin B_{12s} to a cobamide coenzyme, α-(5,6-dimethylbenzimidazolyl) deoxyadenosylcobamide (adenosyl-B_{12}). *J Biol Chem* **241**, 1455–1461.

117 Johnson CLV, Buszko ML & Bobik TA (2004) Purification and initial characterization of the *Salmonella enterica* PduO ATP:Cob(I)alamin adenosyltransferase. *J Bacteriol* **186**, 7881–7887.

118 Buan NR, Suh S & Escalante-Semerena JC (2004) The eutT gene of *Salmonella enterica* encodes an oxygen-labile, metal-containing ATP:corrinoid adenosyltransferase enzyme. *J Bacteriol* **186**, 5708–5714.

119 Jeter VL, Mattes TA, Beattie NR & Escalante-Semerena JC (2019) A new class of

phosphoribosyltransferases involved in cobamide biosynthesis is found in methanogenic archaea and cyanobacteria. *Biochemistry* **58**, 951–964.
120 Öppinger C, Kremp F & Müller V (2022) Is reduced ferredoxin the physiological electron donor for MetVF-type methylenetetrahydrofolate reductases in acetogenesis? A hypothesis. *Int Microbiol* **25**, 75–88.
121 Reeves EP, McDermott JM & Seewald JS (2014) The origin of methanethiol in midocean ridge hydrothermal fluids. *Proc Natl Acad Sci USA* **111**, 5474–5479.
122 Kitadai N, Nakamura R, Yamamoto M, Okada S, Takahagi W, Nakano Y, Takahashi Y, Takai K & Oono Y (2021) Thioester synthesis through geoelectrochemical CO_2 fixation on Ni sulfides. *Commun Chem* **4**, 37.
123 Sánchez-Andrea I, Guedes IA, Hornung B, Boeren S, Lawson CE, Sousa DZ, Arren BE, Claassens NJ & Stams AJM (2020) The reductive glycine pathway allows autotrophic growth of *Desulfovibrio desulfuricans*. *Nat Commun* **11**, 5090.
124 Berg IA (2011) Ecological aspects of the distribution of different autotrophic CO_2 fixation pathways. *Appl Environ Microbiol* **77**, 1925–1936.
125 Beyazay T, Belthle KS, Farès C, Preiner M, Moran J, Martin WF & Tüysüz H (2023) Ambient temperature CO_2 fixation to pyruvate and subsequently to citramalate over iron and nickel nanoparticles. *Nat Commun* **14**, 570.
126 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.
127 Henriques Pereira DP, Leethaus J, Beyazay T, do Nascimento Vieira A, Kleinermanns K, Tüysüz H, Martin WF & Preiner M (2022) Role of geochemical protoenzymes (geozymes) in primordial metabolism: specific abiotic hydride transfer by metals to the biological redox cofactor NAD^+. *FEBS J* **289**, 3148–3162.
128 Müller V, Chowdhury NP & Basen M (2018) Electron bifurcation: a long-hidden energy-coupling mechanism. *Annu Rev Microbiol* **72**, 331–353.
129 Neukirchen S, Pereira IAC & Sousa FL (2023) Stepwise pathway for early evolutionary assembly of dissimilatory sulfite and sulfate reduction. *ISME J* **17**, 1680–1692.
130 Thauer RK (2019) Methyl (alkyl)-coenzyme M reductases: nickel F-430-containing enzymes involved in anaerobic methane formation and in anaerobic oxidation of methane or of short chain alkanes. *Biochemistry* **58**, 5198–5220.
131 Debussche L, Thibaut D, Cameron B, Crouzet J & Blanche F (1990) Purification and characterization of cobyrinic acid a,c-diamide synthase from *Pseudomonas denitrificans*. *J Bacteriol* **172**, 6239–6244.
132 Lübbe YJ, Youn HS, Timkovich R & Dahl C (2006) Siro(haem)amide in *Allochromatium vinosum* and relevance of DsrL and DsrN, a homolog of cobyrinic acid a,c-diamide synthase, for sulphur oxidation. *FEMS Microbiol Lett* **261**, 194–202.
133 Loy A, Duller S, Baranyi C, Mußmann M, Ott J, Sharon I, Béjà O, Le Paslier D, Dahl C & Wagner M (2009) Reverse dissimilatory sulfite reductase as phylogenetic marker for a subgroup of sulfur-oxidizing prokaryotes. *Environ Microbiol* **11**, 289–299.
134 Koonin EV, Makarova KS & Aravind L (2001) Horizontal gene transfer in prokaryotes: quantification and classification. *Annu Rev Microbiol* **55**, 709–742.
135 Dmitrijeva M, Tackmann J, Matias Rodrigues JF, Huerta-Cepas J, Coelho LP & von Mering C (2024) A global survey of prokaryotic genomes reveals the eco-evolutionary pressures driving horizontal gene transfer. *Nat Ecol Evol* **8**, 986–998.
136 Dagan T, Artzy-Randrup Y & Martin W (2008) Modular networks and cumulative impact of lateral transfer in prokaryote genome evolution. *Proc Natl Acad Sci USA* **105**, 10039–10044.
137 Wimmer JLE, Vieira AN, Xavier JC, Kleinermanns K, Martin WF & Preiner M (2021) The autotrophic core: an ancient network of 404 reactions converts H_2, CO_2, and NH_3 into amino acids, bases, and cofactors. *Microorganisms* **9**, 458.
138 Schubert HL, Blumenthal RM & Cheng X (2003) Many paths to methyltransfer: a chronicle of convergence. *Trends Biochem Sci* **28**, 329–335.
139 Sirithanakorn C & Cronan JE (2021) Biotin, a universal and essential cofactor: synthesis, ligation and regulation. *FEMS Microbiol Rev* **45**, fuab003.
140 Huang L, Qin Y, Yan Q, Lin G, Huang L, Huang B & Huang W (2015) MinD plays an important role in *Aeromonas hydrophila* adherence to *Anguilla japonica* mucus. *Gene* **565**, 275–281.
141 Settembre E, Begley TP & Ealick SE (2003) Structural biology of enzymes of the thiamin biosynthesis pathway. *Curr Opin Struct Biol* **13**, 739–747.
142 Wongnate T, Sliwa D, Ginovska B, Smith D, Wolf MW, Lehnert N, Raugei S & Ragsdale SW (2016) The radical mechanism of biological methane synthesis by methyl-coenzyme M reductase. *Science* **352**, 953–958.
143 Romão CV, Ladakis D, Lobo SAL, Carrondo MA, Brindley AA, Deery E, Matias PM, Pickersgill RW, Saraiva LM & Warren MJ (2011) Evolution in a family of chelatases facilitated by the introduction of active site asymmetry and protein oligomerization. *Proc Natl Acad Sci USA* **108**, 97–102.
144 Brindley AA, Raux E, Leech H, Schubert HL & Warren MJ (2003) Identification and characterization

of a small 13-15 kDa "ancestral" cobaltochelatase (CbiX^S) in the archaea. *J Biol Chem* **278**, 22388–22395.
145 Forterre P (2001) Genomics and early cellular evolution. The origin of the DNA world. *C R Acad Sci III* **324**, 1067–1076.
146 Di Giulio M (2011) The last universal common ancestor (LUCA) and the ancestors of archaea and bacteria were progenotes. *J Mol Evol* **72**, 119–126.
147 Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, Butterfield CN, Hernsdorf AW, Amano Y, Ise K *et al.* (2016) A new view of the tree of life. *Nat Microbiol* **1**, 16048.
148 Mahendrarajah TA, Moody ERR, Schrempf D, Szantho LL, Dombrowski N, Davin AA, Pisani D, Donoghue PCJ, Szöllősi GJ, Williams TA *et al.* (2023) ATP synthase evolution on a cross-braced dated tree of life. *Nat Commun* **14**, 7456.
149 Moody ERR, Álvarez-Carretero S, Mahendrarajah TA, Clark JW, Betts HC, Dombrowski N, Szánthó LL, Boyle RA, Daines S, Chen X *et al.* (2024) The nature of the last universal common ancestor and its impact on the early Earth system. *Nat Ecol Evol* **8**, 1654–1666.
150 Chistoserdova L, Vorholt JA, Thauer RK & Lidstrom ME (1998) C1 transfer enzymes and coenzymes linking methylotrophic bacteria and methanogenic Archaea. *Science* **281**, 99–102.
151 Chistoserdova L, Jenkins C, Kalyuzhnaya MG, Marx CJ, Lapidus A, Vorholt JA, Staley JT & Lidstrom ME (2004) The enigmatic planctomycetes may hold a key to the origins of methanogenesis and methylotrophy. *Mol Biol Evol* **21**, 1234–1241.
152 Kalyuzhnaya MG, Korotkova N, Crowther G, Marx CJ, Lidstrom ME & Chistoserdova L (2005) Analysis of gene islands involved in methanopterin-linked C1 transfer reactions reveals new functions and provides evolutionary insights. *J Bacteriol* **187**, 4607–4614.
153 Chistoserdova L & Kalyuzhnaya MG (2018) Current trends in methylotrophy. *Trends Microbiol* **26**, 703–714.
154 Adam PS, Borrel G & Gribaldo S (2019) An archaeal origin of the Wood-Ljungdahl H4MPT branch and the emergence of bacterial methylotrophy. *Nat Microbiol* **4**, 2155–2163.
155 Jin X, Yang Y, Cao H, Gao B & Zhao Z (2022) Eco-phylogenetic analyses reveal divergent evolution of vitamin B_{12} metabolism in the marine bacterial family '*Psychromonadaceae*'. *Environ Microbiol Rep* **14**, 147–163.
156 Kirschning A (2022) On the evolution of coenzyme biosynthesis. *Nat Prod Rep* **39**, 2175–2199.
157 Kuznetsov S, Milenkin A & Antonov I (2022) Translational frameshifting in the chlD gene gives a clue to the coevolution of the chlorophyll and cobalamin biosyntheses. *Microorganisms* **10**, 1200.
158 Belthle KS, Beyazay T, Ochoa-Hernandez C, Miyazaki R, Foppa L, Martin WF & Tüysüz H (2022) Effects of silica modification (Mg, Al, Ca, Ti, and Zr) on supported cobalt catalysts for H_2-dependent CO_2 reduction to metabolic intermediates. *J Am Chem Soc* **144**, 21232–21243.
159 Sousa FL, Thiergart T, Landan G, Nelson-Sathi S, Pereira IAC, Allen JF, Lane N & Martin WF (2013) Early bioenergetic evolution. *Philos Trans R Soc Lond B Biol Sci* **368**, 20130088.
160 Bowman JC, Petrov AS, Frenkel-Pinter M, Penev PI & Williams MD (2020) Root of the tree: the significance, evolution, and origins of the ribosome. *Chem Rev* **120**, 4848–4878.
161 Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, Connor R, Funk K, Kelly C, Kim S *et al.* (2022) Database resources of the national center for biotechnology information. *Nucleic Acids Res* **50**, D20–D26.
162 Kanehisa M & Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27–30.
163 Chen IMA, Chu K, Palaniappan K, Ratner A, Huang J, Huntemann M, Hajek P, Ritter SJ, Webb C, Wu D *et al.* (2023) The IMG/M data management and analysis system v.7: content updates and new features. *Nucleic Acids Res* **51**, D723–D732.
164 The UniProt Consortium (2023) UniProt: the universal protein knowledgebase in 2023. *Nucleic Acids Res* **51**, D523–D531.
165 Buchfink B, Xie C & Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. *Nat Methods* **12**, 59–60.
166 Buchfink B, Reuter K & Drost HG (2021) Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat Methods* **18**, 366–368.
167 Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G *et al.* (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240.
168 Needleman SB & Wunsch CD (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol* **48**, 443–453.
169 Katoh K & Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**, 772–780.
170 Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313.
171 Nagies FSP, Brueckner J, Tria FDK & Martin WF (2020) A spectrum of verticality across genes. *PLoS Genet* **16**, e1009200.

172 Letunic I & Bork P (2021) Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* **49**, W293–W296.
173 Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN & Bourne PE (2000) The Protein Data Bank. *Nucleic Acids Res* **28**, 235–242.

Supporting information

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

Fig. S1. The anaerobic cobalamin biosynthesis pathway.

Fig. S2. Structural alignments (25) of the two CoFeS subunits of *M. maripaludis* (mmp0980 and mmp0981, red) and *M. thermoacetica* (moth_1198 and moth_1201, blue).

Fig. S3. Species tree of 80 analyzed methanogens.

Fig. S4. Concatenated tree of the anaerobic route to cobalamin in methanogens and acetogens.

Table S1. TM-scores for structural alignments of archaeal and bacterial enzymes of the methyl branch of the acetyl-CoA pathway and the anaerobic cobalamin synthesis pathway.

Table S2. Prokaryotic informational core genes (ICG) and their functions. The genes were used to calculate a species tree comprising all analysed methanogens (Fig. S3).

Table S3. Diamond blast hits of cobalamin biosynthesis proteins vs the dataset of 404 autocatalytic core reactions.

Table S4. Homologues of cobalamin biosynthesis proteins found in 355 protein families traced to LUCA (Weiss *et al*. 2016). Sorted by functional category.

Table S5. Homologues of cobalamin synthesis proteins that share SUPERFAMILY annotation found in the dataset of 3939 genomes.

Table S6. Accession numbers and taxonomy of acetogen and methanogen genomes analysed in this work. The genomes are a subset of a Reference Sequence (RefSeq) data set of 2016.

Table S7. Query sequences used for homology searches and sources for structural comparisons.

II The early evolution of the glycolytic pathway from autotrophic origins to glycogen and back

Luca D. Modjewski[*1], Ulrike Johnsen[2], Peter Schönheit[2], William F. Martin[1]

[1] Institut für Molekulare Evolution, Heinrich-Heine-Universität Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Deutschland

[2] Institut für Allgemeine Mikrobiologie, Christian-Albrechts-Universität zu Kiel, 24118 Kiel, Deutschland

[*]Korrespondenz

Diese Publikation wurde am 30. September 2025 in *FEMS Microbiology Reviews* veröffentlicht.

Beitrag von Luca D. Modjewski

Ich war an folgenden Aspekten der Publikation beteiligt: Methodologie, Recherche, Zusammenstellung und Analyse des Datensatzes, Visualisierung der zuvor gewonnen Ergebnisse (Abbildungen 2 und 3; Abbildungen 5 bis 8, Tabellen 1 und 2), Visualisierung des zugrundeliegenden Stoffwechselwegs, Anfertigung des gesamten ergänzenden Materials sowie Erstellung und Überarbeitung des Manuskripts.

FEMS Microbiology Reviews, 2025, 49, fuaf049
DOI: 10.1093/femsre/fuaf049
Advance access publication date: 30 September 2025
Review Article

The early evolution of the glycolytic pathway from autotrophic origins to glycogen and back

Luca D. Modjewski[1,*], Ulrike Johnsen[2], Peter Schönheit[2], William F. Martin[1]

[1]Institute of Molecular Evolution, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany
[2]Institut für Allgemeine Mikrobiologie, Christian-Albrechts-Universität zu Kiel, 24118 Kiel, Germany
*Corresponding author. Institute of Molecular Evolution, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany. E-mail: luca.modjewski@hhu.de
This paper is dedicated to Professor Georg Fuchs on the occasion of his 80th birthday.
Editor: Bernhard Schink

Abstract

Glycolysis stops where gluconeogenesis starts—at pyruvate, the central metabolite of biosynthesis. The early history of carbon metabolism is preserved in archaeal and bacterial enzymes for glucose synthesis and breakdown. Here, we summarize the distribution and phylogeny of enzymes involved in glycolysis, gluconeogenesis, and glycogen metabolism from genomes of cultured prokaryotes. The presence of glycolytic pathways in H_2-dependent chemolithoautotrophs, including methanogens, which cannot grow on exogenous glucose, correlates with their use of glycogen for intracellular carbon storage. Glycogen synthesis and gluconeogenesis are universal among prokaryotes, but glycolysis is not, indicating that the enzymatic conversions of glycolysis arose in the gluconeogenic direction encompassing three phases: (1) an autotrophic origin from H_2 and CO_2 to pyruvate and triosephosphate (trunk glycolysis) fulfilling basic amino acid and cofactor synthesis in the last universal common ancestor, (2) from triosephosphate to glucose supplying cell wall (murein and pseudomurein) and nucleic acid biosynthetic requirements in the first free-living autotrophs, also giving rise to intracellular carbon reserves (glycogen), followed by (3) diversification and transfer of enzymes for glycogen-mobilizing glycolytic routes. An autotrophic origin of trunk glycolysis followed by glycogen-dependent origin of glucose utilization account for conservation, distribution, and diversity of enzymes observed in microbial sugar phosphate pathways.

Keywords: gluconeogenesis; glycolysis; glycogen; autotrophic origin; evolution; carbohydrate metabolism

Abbreviations of enzymes

PK:	Pyruvate kinase
PEPS:	Phosphoenolpyruvate synthase
PPDK:	Pyruvate:phosphate dikinase
PCK:	Phosphoenolpyruvate carboxykinase
ENO:	Enolase
dPGM/iPGM:	Phosphoglycerate mutase (cofactor-dependent/-independent)
PGK:	Phosphoglycerate kinase
GAPDH (NAD):	Glyceraldehyde-3-phosphate dehydrogenase
aGAPDH (NADP):	Archaeal GAPDH, anabolic
GAPN (NADP):	Nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase
GAPOR (Fd):	Glyceraldehyde-3-phosphate:ferredoxin oxidoreductase
TIM:	Triose-phosphate isomerase
FBPA I/II:	Fructose-1,6-bisphosphate aldolase class I/II
PFK (ATP/ADP):	6-phosphofructokinase
FBPase:	Fructose-bisphosphatase
Bif. FBPAase:	Bifunctional fructose-1,6-bisphosphate aldolase/phosphatase
PGI:	Phosphoglucose isomerase
cPGI:	Cupin-type phosphoglucose isomerase
PGI/PMI:	Bifunctional phosphoglucose/phosphomannose isomerase
GK (ATP/ADP):	Glucokinase
PGLM:	Phosphoglucomutase
NGPase:	NDP-glucose pyrophosphorylase
AGPase:	ADP-glucose pyrophosphorylase
GGPase:	GDP-glucose pyrophosphorylase
UGPase:	UDP-glucose pyrophosphorylase
GlgA/GS (UDP):	Glycogen synthase
GlgB:	1,4-α-glucan branching protein
GlgP:	Glycogen phosphorylase
GlgX/TreX:	Glycogen debranching enzyme
Gaa:	Glucoamylase
GdE:	Glycogen degrading enzyme

Introduction

Glycolysis is the enzymatic conversion of glucose to pyruvate. This first biochemical pathway ever characterized, bringing a Nobel prize to Meyerhof in 1922, is not universal across prokaryotes, either in its distribution, its chemical intermediates, or its enzymes. This is in contrast to the ribosome (Yonath 2010) or the adenosine triphosphate (ATP) synthase (Walker 1998, Mayer and Müller 2014), which are highly conserved structures that were present in the last universal common ancestor (LUCA) of all cells (Weiss et al. 2016). Glycolysis traditionally figures into the topic of early

Received 11 July 2025; **revised** 17 September 2025; **accepted** 25 September 2025

biochemical evolution because it uncovered the first stoichiometric reactions that conserve energy as ATP (Lipmann 1941), reactions that are now called substrate level phosphorylation (SLP). Yet, a truly primordial role for glycolysis in biochemical evolution, as once argued by Wald (1964), de Duve (1991), and others (Fothergill-Gillmore and Michels 1993, Keller et al. 2014) now seems very unlikely. This is mainly because independent lines of evidence favor an autotrophic origin of metabolism, based upon the antiquity of the linear acetyl-CoA pathway of CO_2 fixation (Fuchs and Stupperich 1985, Fuchs 2011, Martin 2020) and its *homology* (similarity based on common ancestry) to metal-catalysed reactions of H_2 and CO_2 at hydrothermal vents (Martin et al. 2008, Sleep 2018), both in terms of nickel-dependent catalysis (Beyazay et al. 2023a,b,c, Belthle and Tüysüz 2023) and products (Preiner et al. 2020). In addition, the existence of environmentally available glucose is probably a very recent development in Earth history, possibly tracing back only to the origin of cellulose synthesis in the land plant lineage less than 500 million years ago (Mrnjavac et al. 2024a, Allen et al. 2019). From the standpoint of thermodynamics, an autotrophic origin of metabolism (Ronimus and Morgan 2003, Say and Fuchs 2010) is highly compatible with hydrothermal theories for the origin of metabolism from H_2/CO_2-dependent metabolic reactions (Mrnjavac et al. 2024b, Martin and Russell 2007, Schwander et al. 2023), but raises a number of questions about the origin and early evolution of glycolysis (Schönheit et al. 2016), which starts from glucose, not CO_2.

It is well known that common glycolytic pathways like the Embden–Meyerhof pathway (EM pathway), the Entner–Doudoroff (ED) pathway, the pentose phosphate pathway, and the metabolic pathways of energy storage compounds like glycogen are a central aspect of prokaryotic energy metabolism (Bräsen et al. 2014, Wang et al. 2019, Cifuente et al. 2024). But are they truly ancient? The earliest studies of glycolytic enzymes in Archaea uncovered a recurrent pattern that has persisted to the present. The archaeal enzymes differ in sequence and structure from their bacterial counterparts, whereby the archaeal pathways often entail different, unrelated enzymes (not in EM pathways) (Budgen and Danson 1986, Hensel et al. 1989, Adams 1994, Schönheit and Schäfer 1995, Selig et al. 1997, Ronimus and Morgan 2003, Hansen et al. 2005, Siebers and Schönheit 2005, Reher et al. 2007, Bräsen et al. 2014), with the eukaryotic enzymes being mainly bacterial in origin (Martin and Schnarrenberger 1997). Somewhat puzzling is the observation that many, if not all, strictly H_2/CO_2-dependent archaeal autotrophs harbor some form of a glycolytic pathway, even though they cannot utilize glucose as a growth substrate (Yu et al. 1994, Siebers and Schönheit 2005, Sattler et al. 2024). By contrast, many acetogenic bacteria can readily switch between H_2-/CO_2-dependent growth or glucose-dependent growth, in both cases generating acetate, depending on H_2 partial pressures (Keller et al. 2019).

The EM pathway (classic glycolysis) is present in most organisms with varying degrees of completeness (Fothergill-Gilmore and Michels 1993), whereby trunk glycolysis (the conversion of triosephosphates to pyruvate) is far more widely conserved than the conversion of glucose to triose phosphate. The EM pathway exists in different versions, the classical EM pathway, which occurs mostly in bacteria, and several modified versions of EM pathways, more common among archaeal species (Siebers and Hensel 1993, Selig et al. 1997, Sakuraba and Ohshima 2002, Siebers and Schönheit 2005, Sato and Atomi 2011, Bräsen et al. 2014). Both classical and modified EM pathways are responsible for transforming one molecule of glucose into two molecules of pyruvate using different enzymes in the first half of the pathway.

Table 1. Change of free energy for reactions ($\Delta G_r'$) involved in glycolysis, gluconeogenesis, and glycogen metabolism, estimated using eQuilibrator (Beber et al. 2023) and based on equilibrium constants (K) or apparent equilibrium constants (K') for physiological conditions.

		$\Delta G_r'$ (kJ/mol)	
Enzyme	EC number	Estimated	From K or K'
AGPase	2.7.7.27	8.3	–0.3
GGPase	2.7.7.34	2.3	3.5
UGPase	2.7.7.9	1.6	2.9
PGLM	5.4.2.2	–7.4	–7.2
GK (ATP)*	2.7.1.2	–17.9	–15.0
GK (ADP)*	2.7.1.147	–17.7	N/A
PGI	5.3.1.9	–1.1	1.0
PFK (ATP)*	2.7.1.11	–18.4	–19.5
PFK (ADP)*	2.7.1.146	–14.4	N/A
FBP*	3.1.3.11	–29.9	–13.7
FBPA	4.1.2.13	5.3	15.3
TIM	5.3.1.1	–5.6	–7.7
GAPN*	1.2.1.90	–42.2	N/A
GAPOR*	1.2.7.6	–26.1	N/A
GAPDH	1.2.1.12	21.7	13.8
PGK	2.7.2.3	–19.3	–20.0
iPGM	5.4.2.12	4.5	4.6
ENO	4.2.1.11	3.8	–0.9
PCK (GTP)	4.1.1.32	4.0	–0.4
PCK (ATP)	4.1.1.49	6.8	N/A
PEPS*	2.7.9.2	–16.7	–10.6
PPDK	2.7.9.1	–17.4	–17.5
PK*	2.7.1.40	–27.6	–17.9

Notes: eQuilibrator free energy changes were estimated at pH 7.0, pMg^{2+} 3.0, and ionic strength 0.25 M; no values were returned for the following reactions because the KEGG reaction was not balanced: phosphoglycerate mutase, cofactor-dependent (5.4.2.11); glycogen synthase (2.4.1.21, 2.4.1.11); 1,4-α-glucan branching protein (2.4.1.18); glycogen phosphorylase (2.4.1.1); glycogen debranching enzyme (3.2.1.196); and glucoamylase (3.2.1.3). The reactions indicated as irreversible in Fig. 1 (unidirectional arrows) are highlighted with an asterisk. For reversible reactions in Fig. 1, values of $\Delta G_r'$ are given for the catabolic reaction. The K/K' values used and the corresponding references are listed in Table S1. If multiple values of K or K' were listed for a given reaction, the mean $\Delta G_r'$ was calculated. Full enzyme names are listed under "Abbreviations of enzymes". N/A—no (apparent) equilibrium constant was identified.

While the classical EM pathway proceeds under physiological conditions in the catabolic direction (C6 → C3), most of the reactions are reversible, such that the enzymes can catalyse the corresponding reactions in both the glycolytic and the gluconeogenic direction. But several glycolytic reactions are not reversible under physiological conditions, which is also underscored by thermodynamics (changes of free energy $\Delta_rG'^m$ are listed in Table 1), such that alternative reactions are required for these steps (Bräsen et al. 2014, Zhang et al. 2017). For example, in order to convert pyruvate to C6 sugars via the reversed EM pathway, the nonreversible, catabolic reactions catalysed by pyruvate kinase and 6-phosphofructokinase have to be bypassed by phosphoenolpyruvate (PEP) synthase or pyruvate:phosphate dikinase and fructose bisphosphatase, respectively (Bräsen et al. 2014). In all archaea studied to date, the (classical) EM pathway seems to be used for gluconeogenesis rather than sugar degradation, based on its physiological reversibility through gene expression and enzyme activity regulation (Schäfer and Schönheit 1993, Siebers and Hensel 1993, Bräsen et al. 2014). Reversibility led to the suggestion that the EM pathway originally was anabolic (gluconeogenic) and became catabolic (glycolytic) later in evolution (Romano and Conway 1996).

Glucose that is formed in gluconeogenesis is consumed in numerous biosynthetic pathways, including cell wall synthesis (Albers and Mayer 2011, Egan et al. 2017) and supplying ribose for nucleic acids. One of the more common products of gluconeogenesis is glycogen, a branched polymer of 1,4-α-linked and 1,6-α-linked glucose molecules and one of the main energy storage compounds of cells (Wilkinson 1959). Glycogen is used for energy storage and stress resistance, and occurs in all three domains of life (Wang et al. 2019). Glycogen synthesis (GS) and degradation proceeds via three different pathways, its presence in archaeal species was reported in the early 1980s (König et al. 1982). While the classical pathway includes five enzymes (GlgC, GlgA, GlgB, GlgX/TreX, and GlgP; Fig. 1), the two nonclassical pathways include four (TreS, Pep2, GlgE, and GlgB) and three (GlgC, Rv3032, and GlgB) enzymes. The 1,4-α-glucan branching enzyme (GlgB), glycogen phosphorylase (GlgP), and glycogen debranching enzyme (GlgX/TreX) exist in two different versions each, one that is more common in bacteria and one in archaea. The archaeal 1,4-α-glucan-branching enzyme GlgB and glycogen phosphorylase GlgP are distinct from their bacterial counterparts, the archaeal glycogen debranching enzymes TreX and GlgX share high sequence similarity, rendering them indistinguishable in sequence analyses (Wang et al. 2019).

It was reported that the classical GS pathway occurs mainly in bacteria and that no archaeal species appear to possess all five enzymes, although some archaea, mainly Thermococci, possess at least three enzymes of GS (Wang et al. 2019). The nonclassical pathway utilizing four enzymes was originally thought to be more common in the bacterial domain, since *Picrophilus torridus* was the only archaeon known at the time to possess the pathway (Chandra et al. 2011). However, further studies found that the four-enzyme pathway is more widespread within the archaeal domain than originally thought (Chandra et al. 2011). The second nonclassical pathway seems to be most common among thermophiles. Homologs of the glycogen synthase Rv3032 are found in many archaea, possibly playing a central role in GS (Wang et al. 2019). Glycogen and its synthesis are known in many methanogenic species (Pellerin et al. 1987, Yu et al. 1994, Maitra et al. 2001), as recently reviewed (Gonzalez-Ordenes et al. 2024).

The degradation of glucose to pyruvate as single product in bacteria proceeds either via the EM pathway (classical glycolysis), or, via the classical ED pathway (Flamholz et al. 2013, Chen et al. 2016). The ED pathway is a catabolic pathway and requires five enzymes of the EM pathway, involved in GAP conversion to pyruvate. Gluconeogenesis via the ED pathway has not been reported so far. In Archaea, carbon metabolism is more diverse, with modified versions of both the EM pathway and the ED pathway (Bräsen et al. 2014). Hyperthermophilic archaea can generate pyruvate from glucose either via modifications of the EM pathway, the nonphosphorylative ED pathway, or a combination of both (Bräsen et al. 2014) as reported for *Thermoproteus tenax* (Selig et al. 1997, Reher et al. 2010). In halophilic archaea, a semiphosphorylative version of the ED pathway is operative (Kuprat et al. 2021). In the present paper, we will focus on the EM pathway from glucose to pyruvate and on enzymes that function in the gluconeogenic direction (Fig. 1).

Most enzymes of the classical EM pathway in bacteria are able to operate in the direction of both glycolysis and gluconeogenesis, but the catabolic reactions catalysed by phosphofructokinase and pyruvate kinase are irreversible under physiological conditions and need to be bypassed in gluconeogenesis (Bräsen et al. 2014). In contrast to the classical EM pathway in bacteria, the modified EM pathways of archaea contain two additional strictly irreversible enzymes, the nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase GAPN (Lorentzen et al. 2004), which is part of the aldehyde dehydrogenase superfamily (Reher and Schönheit 2006), and glyceraldehyde-3-phosphate:ferredoxin oxidoreductase GAPOR (Adams 1994). Both enzymes function unidirectionally in the catabolic direction under physiological conditions (Reher et al. 2007), catalysing the phosphate-independent conversion of glyceraldehyde-3-phosphate to 3-phosphoglycerate without the formation of 1,3-bisphosphoglycerate, as in the classical GAPDH-PGK reaction sequence; as a consequence, both enzymes do not couple the (exergonic) glyceraldehyde-3-phosphate oxidation to 3-phosphoglycerate with energy conservation as ATP (Fig. 1). GAPDHs can be divided into two classes: archaeal GAPDHs, which only act in the direction of gluconeogenesis (anabolic) and are phylogenetically ancestral, and bacterial GAPDHs, which are able to catalyse the reaction in both directions and were transferred from bacteria to halophilic archaea (Tästensen and Schönheit 2018). The enzymes of the archaeal EM pathway and ED pathway as well as related pathways were reviewed by Bräsen et al. (2014) and Siebers and Schönheit (2005).

Here, we focus on evolutionary aspects and investigate the distribution and phylogenies of glycolytic and gluconeogenic enzymes among the classical and modified EM pathway linking pyruvate with glucose in an evolutionary context. We aim to address the question of whether the data indicate a heterotrophic (glycolytic) or an autotrophic, gluconeogenic origin (CO_2 → pyruvate → glucose → glycogen) of the glycolytic pathway and how glycogen metabolism figures into the evolution of carbohydrate metabolism.

Universality of trunk glycolytic enzymes

For the present investigation, the assignment of the enzymes to bacteria or archaea is important if we want to look into the origin of the pathway extending back to the LUCA. We did not include metagenomic data, because they are still problematic when considering higher level taxonomy assignment of sequences (Zhang et al. 2025). Among the Reference Sequence genome database, there are many more bacteria represented than archaea, so we generated a prokaryotic dataset with balanced representation among archaeal and bacterial groups, with sampling as explained in the legend to Fig. 2. The enzymes chosen for study here included those outlined in Fig. 1.

In the study of ancient protein evolution, phylogenetic trees are typically used, but trees have the disadvantage that they only display genomes where the gene is present. In deep phylogeny, it is equally important to see where the gene is absent, in order to uncover deep conservation and lineage specific inventions. The presence/absence distribution of our current enzyme sample was determined by scoring archaeal and bacterial homologs of the query sequence set (Table S2) in the respective genomes. Displayed as a summary of complete, partially present, or absent pathways, the result is shown in Fig. 2.

Gluconeogenesis was scored as complete if a genome was found to possess at least one homolog of the following enzymes (otherwise, gluconeogenesis was considered partial): PPDK or PEPS, ENO, iPGM, or dPGM, PGK, GAPDH, TIM, either FBPA I/II and FBPase or bifunctional FBPAase, PGI, or cPGI or bifunctional PGI/PMI and PGLM. This requirement was fulfilled by 255 archaeal and 325 bacterial genomes. Methanobacteria was the only taxonomic group in which no complete gluconeogenesis was detected, on the other hand, Methanococci was the only group with a complete gluconeogenic pathway in each single member. Out of the 255 archaeal and 325 bacterial genomes harboring complete gluconeogenesis, 25 archaea and 251 bacteria were found

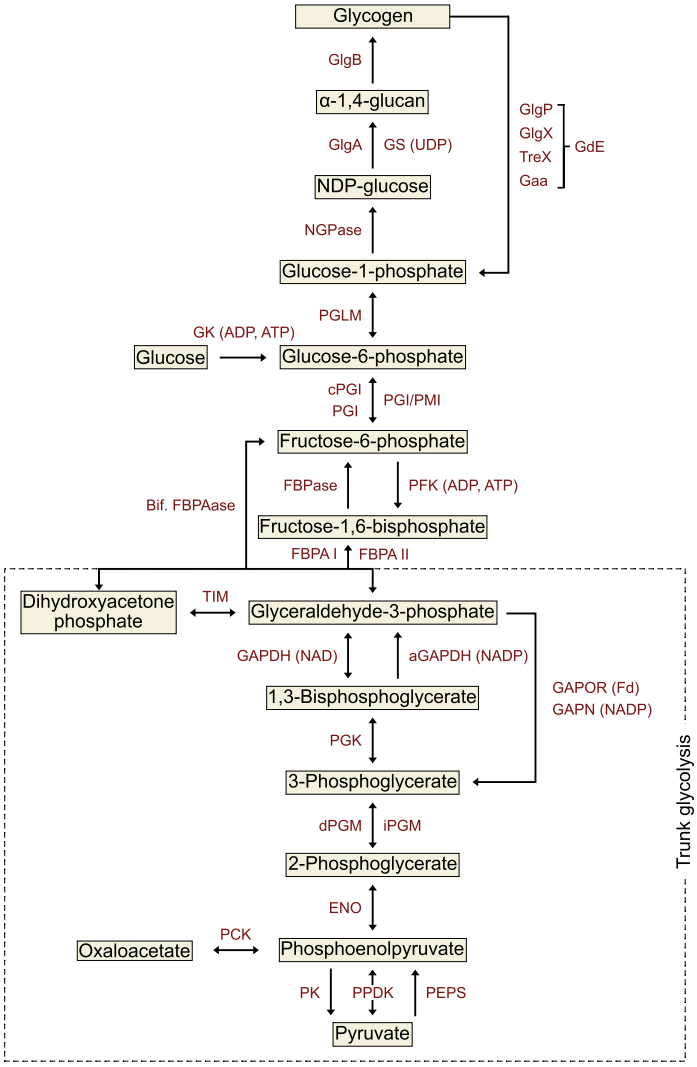


Figure 1. Schematic of the here investigated glycolytic/gluconeogenic pathway. The diagram visualizes the enzymatic conversion of pyruvate and oxaloacetate into polymeric C6 sugars and vice versa. Abbreviations display enzymes analysed in this work and their involvement in the corresponding anabolic and/or catabolic reactions. Trunk glycolysis is outlined by a dashed box. Changes of free energy for the reactions as calculated with eQuilibrator (Beber et al. 2023) and based on (apparent) equilibrium constants (Table S1) are given in Table 1. Reactions that are irreversible under physiological conditions are highly exergonic and indicated with unidirectional arrows in the figure. Protein identifiers as well as additional information of all query sequences of 29 different enzymes involved in glycolysis, gluconeogenesis, and GS are listed in Table S2. Enzymes of the modified archaeal EM pathway were added based on Bräsen et al. (2014). PK—pyruvate kinase; PEPS—phosphoenolpyruvate synthase; PPDK—pyruvate:phosphate dikinase; PCK—phosphoenolpyruvate carboxykinase; ENO—enolase; dPGM/iPGM—phosphoglycerate mutase (cofactor-dependent/-independent); PGK—phosphoglycerate kinase; GAPDH (NAD)—glyceraldehyde-3-phosphate dehydrogenase; aGAPDH (NADP)—archaeal GAPDH, anabolic; GAPN (NADP)—nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase; GAPOR (Fd)—glyceraldehyde-3-phosphate:ferredoxin oxidoreductase; TIM—triose-phosphate isomerase; FBPA I/II—fructose-1,6-bisphosphate aldolase class I/II; PFK (ATP/ADP)–6-phosphofructokinase; FBPase—fructose-bisphosphatase; Bif. FBPAase—bifunctional fructose-1,6-bisphosphate aldolase/phosphatase; PGI—phosphoglucose isomerase; cPGI—cupin-type phosphoglucose isomerase; PGI/PMI—bifunctional phosphoglucose/phosphomannose isomerase; GK (ATP/ADP)—glucokinase; PGLM—phosphoglucomutase; NGPase—NDP-glucose pyrophosphorylase; GlgA/GS (UDP)—glycogen synthase; GlgB–1,4-α-glucan-branching protein; GlgP—glycogen phosphorylase; GlgX/TreX—glycogen debranching enzyme; Gaa—glucoamylase; and GdE—glycogen degrading enzyme.

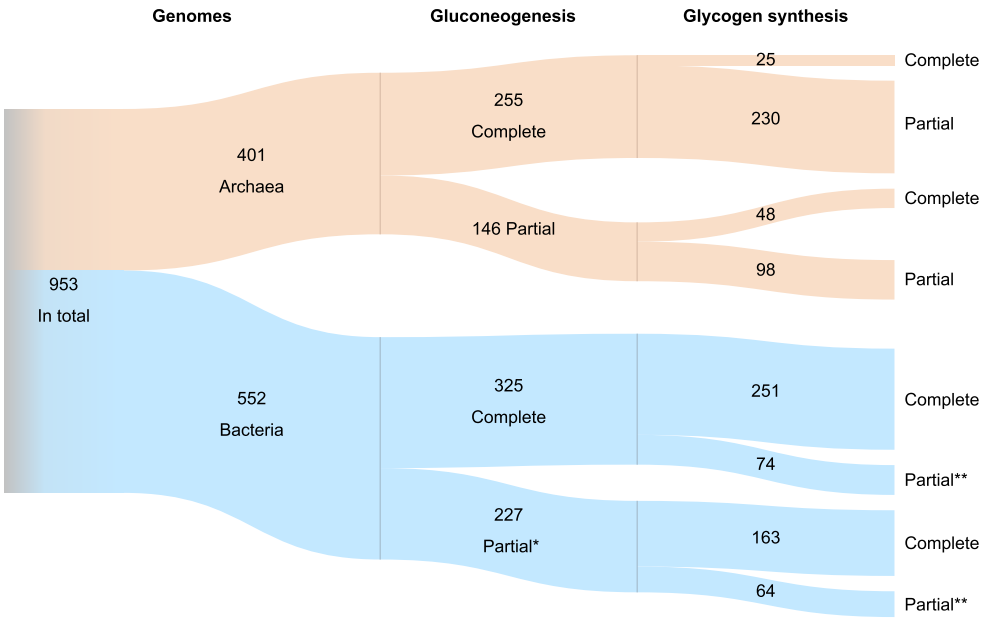


Figure 2. Numbers of complete or partially present pathways for gluconeogenesis and GS within the dataset of 953 prokaryotic genomes (401 archaea and 552 bacteria), sampled from all complete Reference Sequence (O'Leary et al. 2016) genomes of the National Center for Biotechnology Information (Sayers et al. 2022) available in May 2024. The data set represents the biggest genomes in terms of the number of protein sequences per archaeal species and bacterial family (Table S5) in order to avoid a reduction of the taxonomic spectrum. Archaeal classes and bacterial phyla with <10 representatives were classified and combined as "Other Archaea" and "Other Bacteria," respectively. Corresponding numbers for each analysed archaeal and bacterial taxonomic group can be found in Table S3. ∗The 227 bacterial genomes with partial gluconeogenesis include one species that completely lacks homologs of gluconeogenic enzymes, *Ca. Absonditicoccus praedator*. ∗∗Including 18 bacterial genomes in total in which no homologs of enzymes involved in GS were detected.

to have a complete GS pathway, defined here as the ability to link glucose molecules from either UDP- or ADP-glucose using at least one type of glycogen synthase and the 1,4-α-glucan-branching protein. Although only partial gluconeogenesis was detected in several lineages, 48 out of 146 archaea and 163 out of 227 bacteria were found to possess complete GS pathways. Eighteen bacterial genomes completely lacked homologs of enzymes involved in GS (UDP, GlgABC). Among genomes sampled, only one species, *Ca. Absonditicoccus praedator*, a parasitic bacterium that lacks most genes of pathways involved in core carbon metabolism (Yakimov et al. 2022) lacked homologs of gluconeogenic enzymes. Furthermore, we also investigated (not shown) the presence of homologous sequences of gluconeogenic enzymes in the archaeal symbiont *Nanoarchaeum equitans* (Huber et al. 2002) and symbionts of the genus *Rickettsia* (Perlman et al. 2006). No gluconeogenic enzymes were detected in *N. equitans* (GenBank AE017199.1) and only PPDK as well as PGLM were scored in three *Rickettsia* genomes (NCBI RefSeq GCF_964026435.1, GCF_964026455.1, and GCF_964030815.1).

Of course, the frequency of incomplete pathways in Fig. 2 can reflect any number of factors, including auxotrophies (enzyme is not required hence lost), low sequence conservation (enzyme present but not detected), or the existence of functionally equivalent enzymes or shunts involving unannotated enzymes or routes not included in our search set. This is most evident among pathways for GS in archaea, where the majority of archaea represented in Fig. 2 possess incomplete pathways, even though almost all of the archaea sampled possess a form of glycogen synthase (Fig. 3). This suggests that our search set for the pathway from glucose-6-phosphate to glycogen (Fig. 1) is incomplete. Such limitations are inherent to the kind of survey that we have performed here.

Displayed as a matrix of gene presence (columns) versus genomes (rows), the distribution of glycolysis, gluconeogenesis and GS is shown in Fig. 3. The trunk glycolytic enzymes (present in % of genomes) phosphoenolpyruvate synthase (87.6%), enolase (98.5%), cofactor-independent phosphoglycerate mutase (82.9%), phosphoglycerate kinase (98.9%), NAD-type or archaeal NADP-type glyceraldehyde-3-phosphate dehydrogenase (99.2%), and triosephosphate isomerase (98.6%) are universally (or nearly so) distributed and highly conserved across the prokaryotic sample. We detected at least one homolog of a fructose-1,6-bisphosphate aldolase—class I (Siebers et al. 2001), class II (Alefounder et al. 1989, Pickl et al. 2012), or the bifunctional (gluconeogenic) fructose-1,6-bisphosphate aldolase/bisphosphatase (Say and Fuchs 2010)—in 95.4% of the genomes sampled, either a phosphoglucose isomerase, a cupin-type phosphoglucose isomerase (Hansen et al. 2005) or a bifunctional phosphoglucose/phosphomannose isomerase (Hansen et al. 2004) in almost every genome (95.5%). The same is true for phosphoglucomutase (Cori and Cori 1936), found in 946 (99.3%) genomes. Given their universal distribution across bacteria and archaea, the data suggest the presence of these enzymes in the last common ancestor

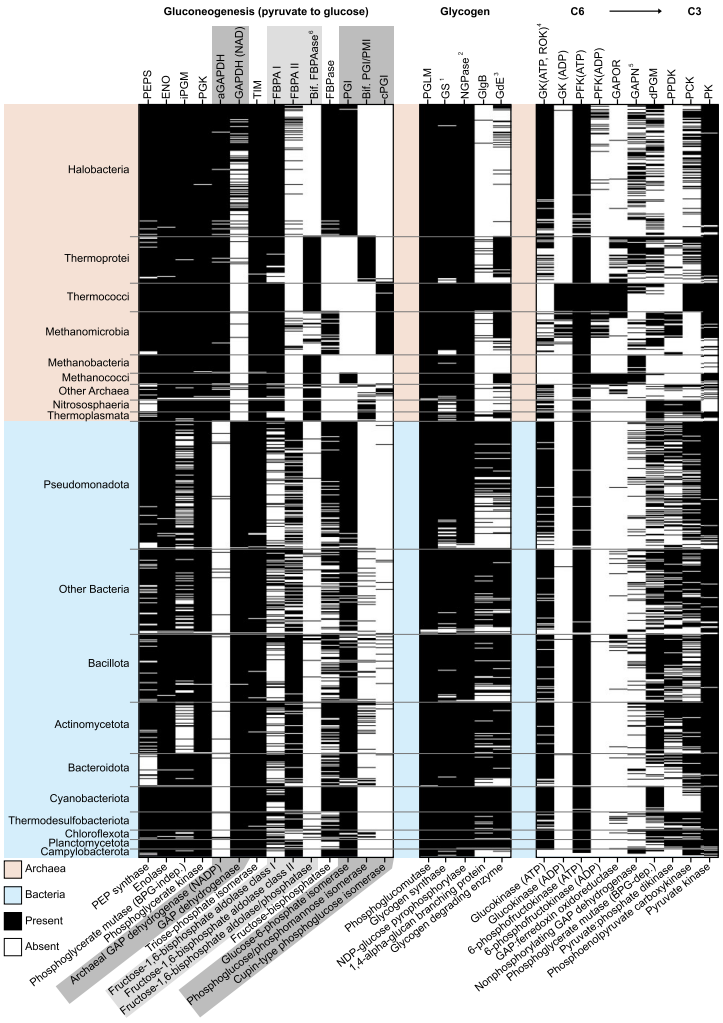


Figure 3. Presence–absence matrix of enzymes involved in gluconeogenesis, GS/degradation, and glycolysis. The upper and lower x-axis labels represent the analysed enzymes (abbreviations and full names), the y-axis shows the prokaryotic taxonomic spectrum analysed in this work. Each row displays the presence (black patches) or absence (white patches) of at least one homologous sequence of the corresponding enzyme in a single genome. The matrix is divided into three parts: The left pattern shows the distribution of enzymes involved in gluconeogenesis (pyruvate to glucose), the middle part displays the presence and absence of enzymes involved in GS/degradation, and the right panel demonstrates additional enzymes involved in C6 → C3 metabolism. Query sequences were blasted against 953 prokaryotic genomes using DIAMOND version 2.1.8.162 (Buchfink et al. 2021) with an e-value threshold of $\leq 1^{-10}$ and sequence identities of at least 25%. [1]Glycogen synthase includes homologs of glycogen synthase GlgA, glycogen synthase (UDP), and Rv3032 (Stadthagen et al. 2007). [2]NGPase entails homologous sequence of ADP-/UDP-/GDP-glucose pyrophosphorylases (Cifuente et al. 2024). [3]The column named GdE encompasses best hits against glycogen phosphorylase GlgP, glycogen-debranching enzyme GlgX, and glucoamylase of *Sulfolobus acidocaldarius* (Lee et al. 2021) as well as glycogen-debranching enzyme TreX of *Sulfolobus solfataricus* (Park et al. 2007a). [4]ATP-dependent glucokinases of Archaea and Bacteria belong to either glucokinases or ROK (repressor, ORF, kinase) family glucokinases (Hansen et al. 2002; Hansen and Schönheit, 2003). ATP-dependent glucokinases of halophilic archaea are not involved in glucose degradation, but convert glucose to glucose-6-phosphate as a starting reaction leading to the pentose phosphate pathway (Pickl and Schönheit, unpublished data). [5]As GAPN is part of the aldehyde dehydrogenase superfamily, sequences of other members of this superfamily were also detected via the homology search. In order to reduce the number of false positives, GAPN sequences displayed in this matrix were selected based on the phylogenetic tree shown in Fig. S1 (B). [6]The almost exclusive presence of bifunctional fructose bisphosphate aldolase/phosphatase in hyperthermophilic archaea is in accordance with its adaptation to a hyperthermophilic life style (Say and Fuchs 2010). Its bifunctionality ensures that heat-labile triosephosphates are quickly removed and trapped in stable fructose-6-phosphate, rendering gluconeogenesis unidirectional. Figures S2–S8 were created to distinguish between different ATP-PFKs, ATP-GKs, glycogen synthases, glycogen-branching enzymes, glucose pyrophosphorylases, PEP carboxykinases, and glycogen degrading enzymes.

of bacteria and archaea, which in modern schemes is the LUCA. A study using maximum-likelihood tree reconstruction models was able to trace only two of these enzymes, triose-phosphate isomerase and cupin-type phosphoglucose isomerase, to LUCA (Weiss et al. 2016), indicating that gene distributions themselves (Fig. 3) can contain robust information that likelihood models can miss.

Glycogen is one of the most common reserve compounds of cells (Wang et al. 2019). Early work detected glycogen in some members of the thermoacidophilic archaea *Sulfolobus*, *Thermoproteus*, *Desulfurococcus*, and *Thermococcus* (König et al. 1982) as well as in methanogens (Murray and Zinder 1987). The first step of GS is the formation of NDP-glucose, which is catalysed by NDP-glucose pyrophosphorylase. Subsequently, glycogen synthase links multiple molecules of NDP-glucose to linear chains, which are further branched with the help of 1,4-α-glucan-branching protein (Preiss 2014). We found a wide distribution of at least one glycogen synthase (in 93.7% of genomes) and NDP-glucose pyrophosphorylase (98.9%) across the prokaryotic sample (Fig. 3). The 1,4-α-glucan-branching protein GlgB, which enhances the density of glucose polymers (Cifuente et al. 2019), is widely dispersed across the bacterial sample and was detected in several archaeal genomes, in particular among the Thermoprotei, Thermococci, and Methanomicrobia. The broader distribution of GlgB in bacteria may relate to its involvement in a further pathway of GS (nonclassical pathway I), which links trehalose to glycogen metabolism and is common to bacterial species (Chandra et al. 2011).

The degradation of glycogen to glucose-1-phosphate (Fig. 1) proceeds via glycogen phosphorylase GlgP in combination with the debranching enzymes GlgX, TreX, and glucose-forming glucoamylase Gaa (Chandra et al. 2011, Lee et al. 2021, Cifuente et al. 2024).

The right panel of Fig. 3 shows the occurrence of homologs for additional enzymes involved in sugar phosphate metabolism. As previously reported (Dandekar et al. 1999, Ronimus and Morgan 2003, Bräsen et al. 2014), the remaining enzymes involved in C6 → C3 metabolism are, especially in the archaeal sample, more sparsely distributed. For example, methanogens of the Methanobacteria and Methanococci lack homologs of glucokinase and phosphoenolpyruvate carboxykinase. Although the modified EM pathway is mainly used by archaea (Selig et al. 1997, Siebers and Schönheit 2005, Bräsen et al. 2014), the presence of GAPN, which catalyses the irreversible oxidation of glyceraldehyde-3-phosphate to 3-phosphoglycerate, was reported in bacteria as well (Iddar et al. 2005). Homologs of GAPN were nearly universally detected, but other members of the aldehyde dehydrogenase superfamily were among the identified sequences, too. These sequences were removed based on the phylogeny shown in Fig. S1. In accordance with previous findings, homologs of ferredoxin-dependent GAPOR, a glycolysis-specific enzyme, were almost exclusively detected among Thermoprotei, Thermococci, and Methanococci (Heider et al. 1995, van der Oost et al. 1998, Park et al. 2007b), but also in some bacterial species, which might use a phylogenetically unrelated type of GAPOR in alternative glycolytic pathways (Scott et al. 2019).

Trunk glycolysis is conserved because of autotrophic amino acid metabolism

The conversions from triosephosphates to pyruvate (Fig. 1) are traditionally termed trunk glycolysis (Fothergill-Gillmore and Michels 1993). It has long been known that the enzymes of trunk glycolysis, enolase, phosphoglycerate mutase, phosphoglycerate kinase, glyceraldehyde-3-phosphate dehydrogenase, and triosephosphate isomerase, are reversible enzymes that are involved both in gluconeogenesis and glycolysis and are the most widely distributed subset of glycolytic enzymes (Fothergill-Gilmore and Michels 1993). This holds true for analyses of archaea as well. As shown in Fig. 3, the trunk glycolytic pathway is widely distributed and highly conserved among bacteria and archaea, suggesting that this segment of the EM pathway is more ancient than the enzymes of the upper half of glycolysis, i.e. glucose degradation to triosephosphates. The greater antiquity of trunk glycolysis over the C6 → C3 segment was suggested based on the kinetic properties of the corresponding enzymes (Schäfer and Schönheit 1993), consistent with suggestions that glycolysis originally operated in the gluconeogenic direction (Romano and Conway 1996, Ronimus and Morgan 2003). The enzymes of trunk glycolysis are highly conserved reactions, but there are alternative enzymes in prokaryotes for all of the substrate conversions in glycolysis and gluconeogenesis, with one exception: enolase. Enolase catalyses the stereospecific addition of a water molecule to the double bond in PEP, generating 2-phospho-D-glycerate. There are so far no alternative enzymes to enolase. One possible reason for its extreme conservation, despite catalysing a very simple and freely reversible reaction, is that in organisms that use the acetyl-CoA pathway of carbon and energy metabolism, enolase generates the very first chiral carbon (D-glycerate) in autotrophic metabolism (Martin and Russell 2007), a chirality that is retained throughout the rest of sugar phosphate metabolism (including D-ribose for nucleic acids). For autotrophs, enolase is the wrong reaction for tinkering in an evolutionary sense.

In the catabolic conversions of glucose to triosephosphate, enzymatic diversity for the same or similar reactions is greater than in the anabolic reactions of trunk glycolysis (Fig. 1). This might reflect the ease with which these reactions take place. For example, Degani and Halmann showed in 1967 that glucose-6-phosphate is converted to fructose-6-phosphate and subsequently to glyceraldehyde-3-phosphate and dihydroxyacetone in water at pH > 7–8 without enzymes (Degani and Halmann 1967). Similar nonenzymatic transformations of sugar phosphates in the catabolic direction, that is, nonenzymatic glycolytic-like glucose and related sugar phosphate degradations, have been rediscovered in recent years (Muchowska et al. 2020). This underscores the thermodynamic ease with which carbohydrate breakdown takes place in cells (Decker et al. 1970). By contrast, the synthesis of sugars from CO_2 using the acetyl-CoA pathway requires energy input in the form of 3 ATP per glucose in methanogens—0 ATP from H_2 and CO_2 to pyruvate, 2 ATP at the AMP forming PEP synthase reaction, and 1 ATP at the PGK reaction (Fig. 4)—or 4 ATP per glucose in acetogens, which require one additional ATP from H_2 and CO_2 to pyruvate at the synthesis of formyl tetrahydrofolate (Fuchs 2011). The downhill reactions from glucose to triosephosphate require little in the way of catalysis to go forward (Degani and Halmann 1967), such that alternative enzymes could easily arise in evolution.

The main evolutionary rationale, however, for the greater conservation of trunk glycolysis over "crown" glycolysis rests with the relationship of amino acid biosynthetic pathways to core carbon metabolism in organisms that use the acetyl-CoA pathway for carbon assimilation. This is shown in Fig. 4. Cells are roughly 50% protein by weight (Stouthamer 1973, Schönheit et al. 2016) and the acetyl-CoA pathway is the most ancient pathway of CO_2 fixation (Fuchs and Stupperich 1985, Fuchs 1986, 2011, Schönheit et al. 2016). The strongest evidence in favor of its antiquity

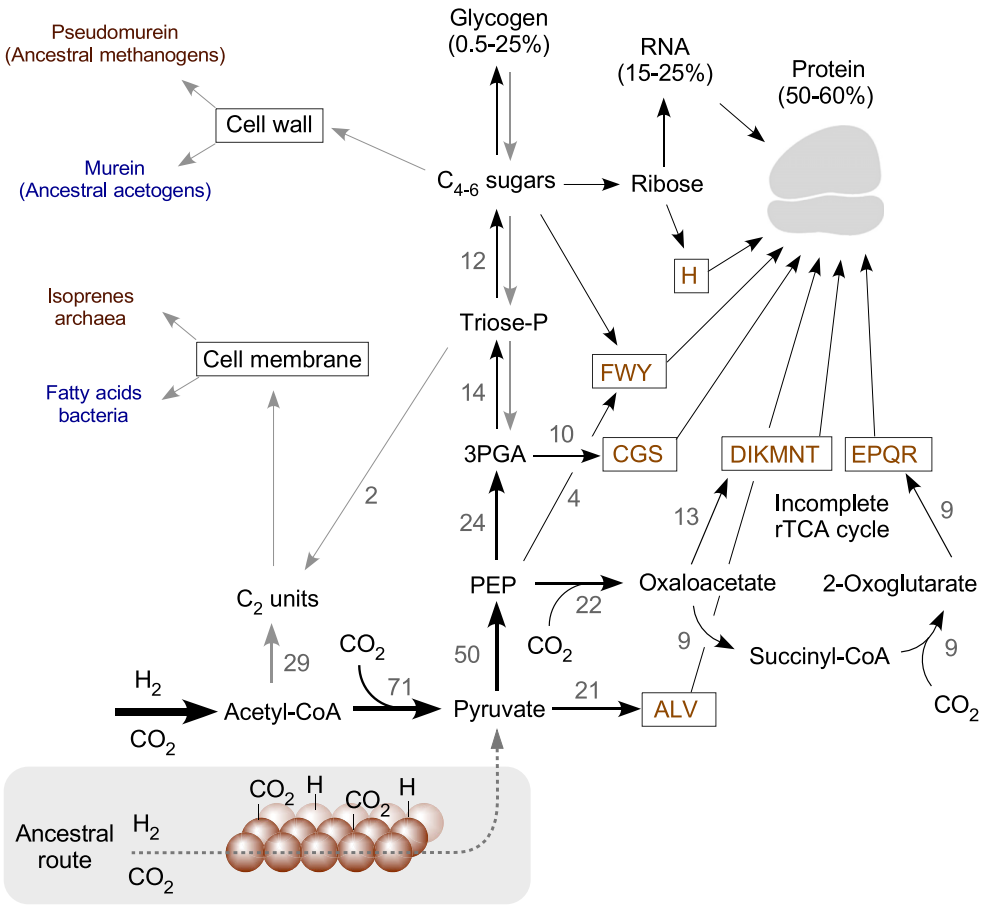


Figure 4. Primordial autotrophic metabolism. Black arrows represent primordial carbon flux, gray arrows represent later enzymatic steps of glycogen mobilization (see text). An arrow can indicate several enzymatic steps. Numbers next to arrows indicate the relative flux of carbon from CO_2 to key intermediates in organisms that use the acetyl-CoA pathway for CO_2 fixation as outlined by Fuchs (1986, 2011), based on 100 acetyl CoA units. The amino acid biosynthetic families (sepia font in one letter code) are boxed. Dotted arrow, grey box, and metal surfaces at lower left indicate metabolic origin from spontaneous synthesis of pyruvate from H_2 and CO_2 on native metal catalysts under hydrothermal vent conditions (Preiner et al. 2020, Martin 2020, Beyazay et al. 2023a, Mrnjavac et al. 2024a). Number ranges in parentheses indicate the approximate contribution of glycogen, RNA, and protein to dry weight of cells (see text and Schönheit et al. 2016). Synthesis of membrane lipids from acetyl CoA and synthesis of cell wall constituents in bacteria and archaea is indicated. While the cell wall of most archaea is made of S-layer proteins, a number of hydrogenotrophic methanogens without cytochromes, such as *Methanopyrus kandleri*, possess pseudomurein (Hartmann and König 1994, Albers and Meyer 2011), which contains N-acetylglucosamine as cell wall constituent. This ancient trait links methanogens without cytochromes to LUCA, because murein (bacteria) and pseudomurein (archaea) are synthesized by ancient, homologous enzymes that trace to LUCA (Subedi et al. 2021), indicating that LUCA was able to synthesize basic cell wall components. Hydrogenotrophic methanogens are ancestral in archaeal phylogeny (Williams et al. 2017, Mei et al. 2023), in agreement with the ancestral nature of methanogenic and acetogenic physiology in archaea and bacteria, respectively (Martin and Russell 2007, Sousa and Martin 2014). Enzymatic lipid synthesis (fatty acids in bacteria, isoprenes in archaea) does not trace to LUCA, which was able to utilize lipids synthesized by geochemical processes in the environment where it arose (Yu et al. 2025). The present paper indicates that ancestral bacteria and archaea were able to synthesize glycogen as a C6 carbon reserve.

over other CO_2-fixing pathways is that the acetyl-CoA pathway is unique in that (i) it operates in both bacteria and archaea (Fuchs 2011, Berg 2011), (ii) it requires no ATP hydrolysis in order to go forward (Fuchs 2011, Berg 2011) and (iii) its overall enzymatic reaction from H_2 and CO_2 to pyruvate can be replaced by native transition metal catalysts, Ni^0 in particular, in water (Preiner et al. 2020, Martin 2020), yielding up to 200 µM pyruvate under some conditions (Beyazay et al. 2023a), the physiological concentration of pyruvate in growing acetogen cells (Furdui and Ragsdale 2000). The ability of Ni^0 to efficiently catalyse the reactions of the acetyl-CoA pathway is significant because of the essential catalytic role of Ni^0 in two key enzymes of the pathway: carbon monoxide dehydrogenase and acetyl-CoA synthase (Doukov et al. 2002, Ragsdale 2008, Ragsdale and Pierce 2008, Biester et al. 2024).

Gluconeogenesis starts with pyruvate, and the acetyl-CoA pathway produces pyruvate (Fuchs 2011).

The schematic flux of carbon through central carbon metabolism of an autotrophic cell that uses the acetyl-CoA pathway is shown in Fig. 4. Pyruvate, PEP, and 3PGA are direct precursors for the universal pathways of amino acid synthesis (Wimmer et al. 2021), and PEP is furthermore the starting point of oxalacetate and 2-oxoglutarate synthesis via the incomplete reverse citric acid cycle (Fuchs and Stupperich 1978) for the synthesis of the aspartate and glutamate families of amino acids. Nucleotide bases are also derived from amino acids and intermediates of the acetyl-CoA pathway, as are cofactors (Wimmer et al. 2021). Of course, at the very beginning of biochemical evolution, amino acids had to be supplied by nonenzymatic reactions in the environment, which is demonstrable in the laboratory under conditions of serpentinizing hydrothermal vents, catalysed by Ni^0 (Kaur et al. 2024), and observed in modern hydrothermal vents in nature (Nobu et al. 2023). When metabolism catalysed by enzymes arose, the building blocks of catalysts were amino acids. The shift from transition metal catalysts to enzymatic catalysis was moderated by cofactors, which can interact with transition metals (Henriques Pereira et al. 2022, Mayer and Moran 2024, Schlikker et al. 2024) or contain transition metals themselves (Brabender et al. 2024, Modjewski et al. 2024). In any event, the transition from transition metal-catalysed reactions in prebiotic metabolism to enzyme-catalysed reactions in early metabolism required carbon backbones and amino acids, probably along the lines of reactions similar to Fig. 4.

Phylogenies trace gluconeogenesis to early biochemical evolution

While gene distributions (Fig. 3) reveal the absence of genes in lineages, providing insights into very early evolution, phylogenetic trees can uncover the degree of transfer and vertical evolution within enzyme families. There is a vast literature on the phylogeny of glycolytic enzymes in bacteria and archaea that extends back over many decades. In Table S4, we have tried to summarize some of the studies of these enzymes that include archaeal homologs, an exhaustive list goes beyond the scope of this paper.

In order to facilitate a current phylogenetic overview of enzymes involved in gluconeogenesis, glycolysis and GS, we generated multiple sequence alignments and unrooted gene trees using MAFFT L-INS-i version 7.471 (Katoh and Standley 2013) and RAxML (random seed: 12345; model: PROTCATWAG) version 8.2.12 (Stamatakis 2014), respectively, for the archaeal and bacterial sequences in the present study. The phylogenies of trunk glycolysis (Fig. 5), enolase, phosphoglycerate kinase, and triosephosphate isomerase reflect a clear immiscibility of archaeal and bacterial clades with very few horizontal gene transfers across the bacterial–archaeal divide. These transfers are found in Methanobacteriales, Nitrososphaerales, and Ca. *Prometheoarchaeum syntrophicum*, respectively. Given the existence of prokaryotic lateral gene transfer (LGT) and exchange via pangenomes and accessory genomes over 4 billion years of evolution (Trost et al. 2024), and given the otherwise high degree of interleaving observed for nine other enzymes surveyed in this study (Fig. 6), this is noteworthy. It indicates that these ancient enzymes, for whatever reason, cannot readily substitute for one another in the bacterial and archaeal lineages. This might reflect their involvement in other essential, but lineage specific functions within the cell ("moonlighting"). For example, enolase is an essential component of the RNA degradosome in bacteria (Morita et al. 2004) and is even considered as a cancer target in humans (Ma et al. 2023). Also, many glycolytic enzymes have moonlighting functions in eukaryotes (Kim and Dang 2005, Zhang et al. 2020, Pan et al. 2021, Bian et al. 2022, Gupta and Uversky 2023).

In contrast to its cofactor-dependent counterpart, cofactor-independent phosphoglycerate mutase is very well distributed within the data set (Fig. 3). The corresponding phylogeny of iPGM shows, like the trees of enolase, phosphoglycerate kinase, and triose-phosphate isomerase, a relatively clear separation of archaea and bacteria (except for few LGTs and a bacterial origin of haloarchaeal iPGM, see Fig. 8), possibly tracing this enzyme to early evolution, too.

The gene trees in this study clearly reflect the influence of LGT among prokaryotes. The unrooted gene trees displayed in Figs 5–8 reveal different degrees of LGT between archaea and bacteria, especially phylogenies of enzymes involved in GS/degradation show more (GS, NGPase, and GdE) or less (GlgB) widespread interleaving and no clear pattern separating archaeal and bacterial sequences. In addition, archaea and bacteria are intermixed in the phylogenies of PEPS, class I FBPA, PGLM, and PK (Fig. 6).

Several enzymes reflect transfer overlain upon domain specificity. For example, the bifunctional fructose-1,6-bisphosphate aldolase/phosphate FBPAase (Say and Fuchs 2010) and cupin-type phosphoglucose isomerase cPGI (Hansen et al. 2001, 2005, Verhees et al. 2001) are typically archaeal enzymes, with a few subsequent, but restricted transfers into the bacterial lineage (Fig. 7). The gluconeogenesis-specific bifunctional FBPAase has been fixed in bacteria of several phyla (Actinomycetota, Bacillota, Chloroflexota, Pseudomonadota, and Thermodesulfobacteriota). The converse picture is found for cofactor-dependent PGM, class II FBPA, FBPase, and phosphoglucose isomerase PGI, which are mainly present in bacterial species with few transfers to archaea (Fig. 8).

ATP-dependent glucokinase is more common among bacteria than its ADP-dependent counterpart (Verhees et al. 2003), which is, like ADP-dependent 6-phosphofructokinase, more widely distributed (Fig. 7) in archaeal species (Siebers and Schönheit 2005, Bräsen et al. 2014). Other than the GAPDH/PGK couple, whose universal distribution as well as the phylogenetic separation of archaeal and bacterial sequences reflect an ancient role in gluconeogenesis, the non-ATP conserving GAPN and GAPOR enzymes only function in the catabolic direction. Compared to GAPOR, which is largely archaea-specific, homologs of GAPN (and other members of the aldehyde dehydrogenase superfamily) are widely distributed across archaea and bacteria and the respective sequences of both domains are highly interleaved in the corresponding phylogeny. In contrast to the phylogenies of ENO, PGK, and TIM in trunk glycolysis, the remaining gene trees of enzymes involved in C6 → C3 metabolism show no consistent pattern separating archaea and bacteria.

Archaeal halophiles have bacterial sugar phosphate metabolism

A particularly prominent aspect of the current investigation is the position of archaeal halophiles. Haloarchaea are aerobic, respiring heterotrophs with a cytochrome-dependent respiratory chain. Evolutionarily they are derived from methanogens by the measure of standard phylogenetic markers (Makarova et al. 2010, Kelly et al. 2011, Sorokin et al. 2017, Mei et al. 2023). For the haloarchaeal glycolytic enzymes that branch with archaeal homologs (ENO, PGK, TIM, aGAPDH, ADP- and ATP-dep.

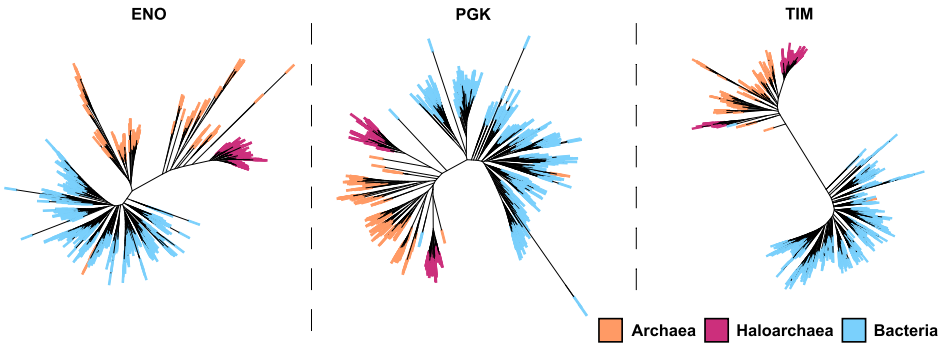


Figure 5. Overview of unrooted gene trees of glycolytic and gluconeogenic enzymes that show—with the exception of few gene transfers—a separation of archaea (orange, magenta) and bacteria (blue). Abbreviations of the corresponding enzymes are displayed above each tree and the arrangement of phylogenies (left to right) is equivalent to the order shown in the presence–absence matrix (Fig. 3). Phylogenies were visually modified utilizing the tool Interactive Tree of Life (Letunic and Bork 2021).

PFK, PEPS, and PGLM), the sister group (excluding "Other Archaea") of the halophiles are mainly Methanomicrobia, followed by Thermoprotei, Nitrososphaeria, and Thermococci. By contrast, phylogenies for eight glycolytic enzymes from halophilic archaea, namely cofactor-independent and cofactor-dependent PGM, GAPDH (NAD), class II FBPA, PGI, ATP-dependent GK, ATP-dependent PEP carboxykinase, and pyruvate kinase, present a consistent pattern: they each reflect a single origin of the haloarchaeal enzymes from a bacterial donor, notwithstanding a subsequent transfer in the case of cofactor-dependent PGM and to the exclusion of other archaea except in the case of ATP-dependent GK (Fig. 8 and Table 2). No other group of archaea or bacteria in the present sample (or any other previously published sample) shows the same pattern. In each case, the Haloarchaea appear to branch within a different bacterial clade, but that does not indicate that these sugar phosphate metabolism genes were acquired in independent events of gene transfer, because the bacterial genes have been undergoing LGT themselves over the last 4 billion years (Martin 1999, Nagies et al. 2020), such that the donor lineage does not need to correlate with the bacterial taxonomic groups within which the halophile enzymes branch today. A detailed look at the sister groups of halophilic archaea in our calculated gene trees (Fig. 8) reveals a mixed picture of bacterial donor lineages, species of the Actinomycetota, Bacillota, and Synergistota are most prominently found in the sister group of Haloarchaea (Table 2). The bacterial nature of individual haloarchaeal sugar phosphate enzymes was shown in earlier studies (Nelson-Sathi et al. 2012, Tästensen and Schönheit 2018, Kuprat et al. 2021), but the present analysis uncovers the overall recurrent pattern more clearly than previous studies do. Importantly, these transfers impact in most cases the Haloarchaea specifically, not other archaeal groups.

With the exception of TIM, ENO, and PGK, which appear more or less refractory to transfer across the bacterial–archaeal divide, Haloarchaea possess a bacterial glycolytic pathway. This underscores the findings of Nelson-Sathi et al. (2012), who reported that the haloarchaeal lineage is a direct descendant of the methanogen lineage, but was transformed at the base of the lineage by gene acquisition from a bacterial donor, transforming a strictly anaerobic H_2-dependent chemolithoautotroph into an aerobic heterotroph possessing (1) a bacterial cytochrome and quinone-dependent respiratory chain (Kennedy et al. 2001), and as we see in the present analysis, (2) a bacterial glycolytic pathway. This physiological transformation closely parallels that at the origin of the eukaryotic lineage under theories that involve an autotrophic host for the origin of mitochondria (Martin and Müller 1998, Müller et al. 2012, Martin et al. 2017, Zhang et al. 2025), except that in the haloarchaeal lineage no bacterial-derived organelle (the mitochondrion) persisted (Nelson-Sathi et al. 2012).

It should be noted that, besides enzymes of sugar phosphate metabolism—the metabolism of galactose (Tästensen et al. 2020) and of the deoxysugar L-rhamnose (Reinhardt et al. 2019) in Haloarchaea also showed many features of LGT events from bacteria. This holds also true for several enzymes of fructose degradation in *Haloferax volcanii*, involving e.g. a bacterial-type phosphotransferase system for the uptake of fructose (Pickl et al. 2012). In contrast, glucose-6-phosphate dehydrogenase from *H. volcanii* as part of the oxidative pentose phosphate pathway is of archaeal origin (Pickl and Schönheit 2015).

Overall, the phylogenies of trunk glycolysis, with some of the gene trees (enolase, phosphoglycerate kinase, and triosephosphate isomerase) showing nearly domain monophyly, and the universal distribution across archaea and bacteria, indicate an ancient gluconeogenic role of the corresponding enzymes in autotrophic metabolism starting from H_2 and CO_2 to generate glucose and glycogen via pyruvate, which is converted into phosphoenolpyruvate by PEP synthase. The unidirectional functionality of PEPS (Schäfer and Schönheit 1993, Hutchins et al. 2001, Zaparty et al. 2008, Kuprat et al. 2020) supports the utilization of this enzyme in a gluconeogenic manner under carbon dioxide-rich conditions (Sleep et al. 2011) prevailing in Earth's early history. This is additionally highlighted by the presence of homologs of PEPS and other gluconeogenic enzymes in hydrogen-dependent methanogens and acetogens, ancient prokaryotic groups inhabiting hydrothermal systems (Weiss et al. 2016, Colman et al. 2022, 2025) and the comparatively sparse distribution of glycolytic enzymes in those groups.

Most of the analysed genomes harbor homologs of at least one glycogen synthase and NDP-glucose pyrophosphorylase, providing the ability to synthesize linked glucose molecules as carbon

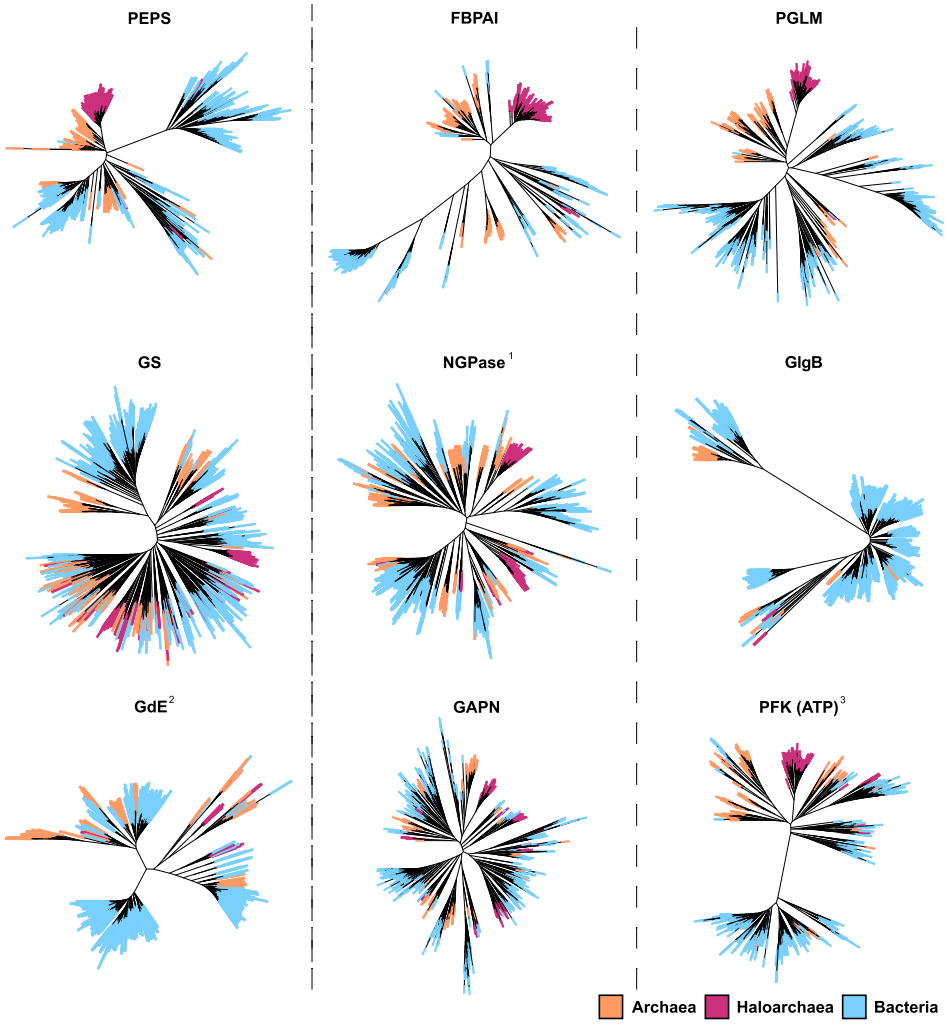


Figure 6. Overview of unrooted gene trees of glycolytic (glycogen → pyruvate) and gluconeogenic enzymes with different degrees of archaeal (orange, magenta) and bacterial (blue) interleaving. Abbreviations of the corresponding enzymes are displayed above each tree and the arrangement of phylogenies (top left to bottom right) is equivalent to the order shown in the presence–absence matrix (Fig. 3). In order to distinguish between GlgA and glycogen synthase (UDP) and to indicate homologs of Rv3032, an alternatively colored version of the GS phylogeny is displayed in Fig. S4. Glycogen-branching enzymes belong to different families of glycoside hydrolases (Suzuki and Suzuki 2016), the sequences of the unrooted GlgB phylogeny are colored according to their glycoside hydrolase family affiliation in Fig. S5. [1]As seen in Fig. 3, NGPase comprises different NDP-glucose pyrophosphorylases (Cifuente et al. 2024), which are separately colored in Fig. S6. [2]GdE comprises homologous sequence of glycogen phosphorylase GlgP, glycogen-debranching enzymes GlgX/TreX and glucoamylase Gaa, which are colored accordingly in Fig. S8. [3]The ATP-dependent phosphofructokinase encompasses two different families, PFK-A, which includes eukaryotic/bacterial ATP-PFKs as well as crenarchaeal PP_i-PFKs, and PFK-B, which belongs to the ribokinase superfamily and includes haloarchaeal 1-phosphofructokinases (Pickl et al. 2012) as well as archaeal 6-phosphofructokinases (Hansen and Schönheit 2001, Siebers and Schönheit 2005, Bräsen et al. 2014). The labels of the PFK (ATP) phylogeny were also colored according to their PFK family affiliation (Fig. S2.), which was investigated using the hmmscan tool of HMMER v3.3.2 (Eddy 2011) and the PFAM-A database (Punta et al. 2012). DIAMOND was used to detect false positive hits against GAPN that belong to the aldehyde dehydrogenase superfamily, too (Table S2). A query dataset of different members of the aldehyde dehydrogenase superfamily was created based on Reher and Schönheit (2006) and searched against all sequences present in the GAPN phylogeny. Phylogenies were visually modified utilizing the tool Interactive Tree of Life (Letunic and Bork 2021).

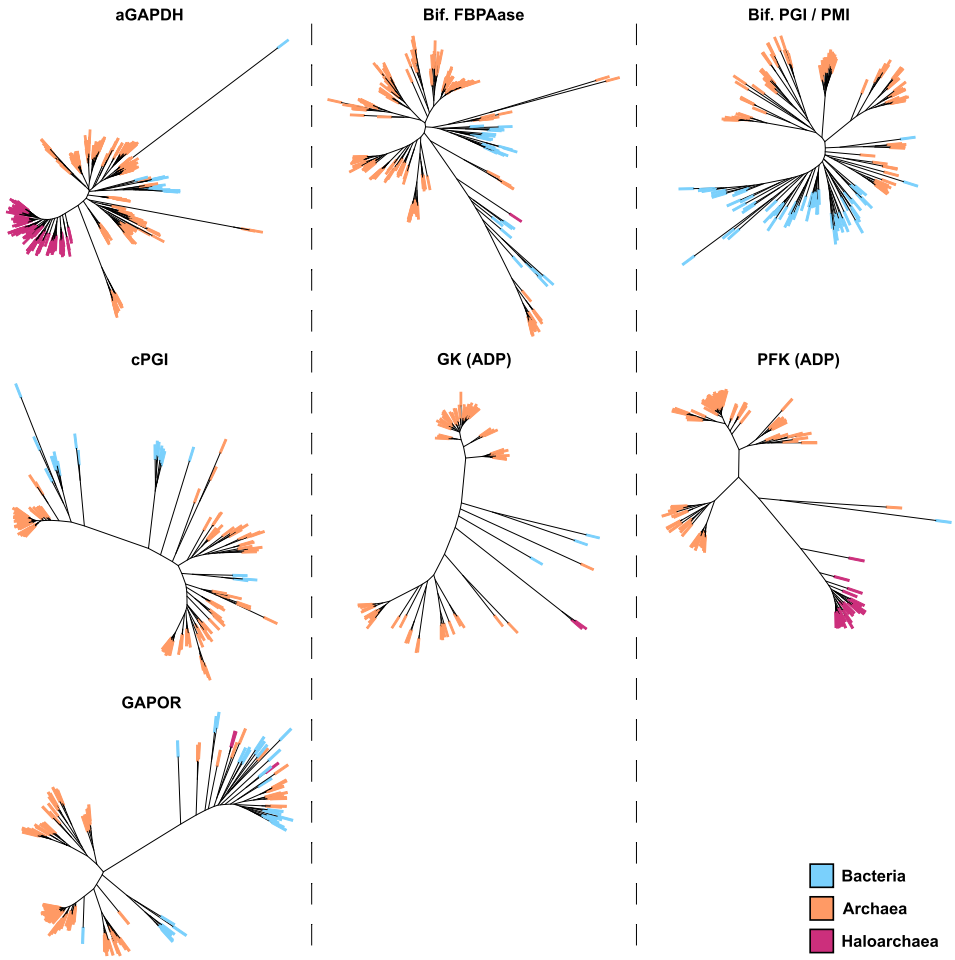


Figure 7. Overview of unrooted gene trees of archaeal (orange, magenta) glycolytic and gluconeogenic enzymes that were rarely transferred to bacterial species (blue). Abbreviations of the corresponding enzymes are displayed above each tree and the arrangement of phylogenies (top left to bottom right) is equivalent to the order shown in the presence–absence matrix (Fig. 3). The rooted phylogeny (not shown) of bifunctional PGMI/PMI indicates an origin of this enzyme in Thermoprotei. Phylogenies were visually modified utilizing the tool Interactive Tree of Life (Letunic and Bork 2021).

and energy reserves. Due to the lack of homologs of the branching enzyme GlgB in most archaea, this might represent an ancient route of connecting glucose molecules, whose packing density was optimized with the use of GlgB later in evolution. Moreover, it might also indicate that a novel, unknown type of branching enzyme, yet to be discovered, is operative in GS in these archaea (Suzuki and Suzuki 2016).

The data suggest a gluconeogenic origin of the classical and modified glycolytic EM pathway, which appears to be an evolutionarily late invention following the autotrophic, gluconeogenic formation of C3 and C6 polymers as well as chained carbons. The presence of glycolysis in strict autotrophs, such as methanogens, indicates that the origin of the glycolytic pathway involved the mobilizing utilization of glycogen that was autotrophically synthesized as an intracellular reserve. The conspicuously bacterial nature of sugar-phosphate metabolizing enzymes in Haloarchaea indicates the evolutionary transformation of a strictly anaerobic, methanogenic autotroph into an O_2-respiring heterotroph via gene transfer from bacteria (Nelson-Sathi et al. 2012). A gluconeogenic origin of carbon metabolism has been proposed many times (Fuchs 1986, Romano and Conway 1996, Ronimus and Morgan 2003, Say and Fuchs 2010), the present findings are in full agreement with that view and furthermore suggest a role of GS in early autotrophic prokaryotes.

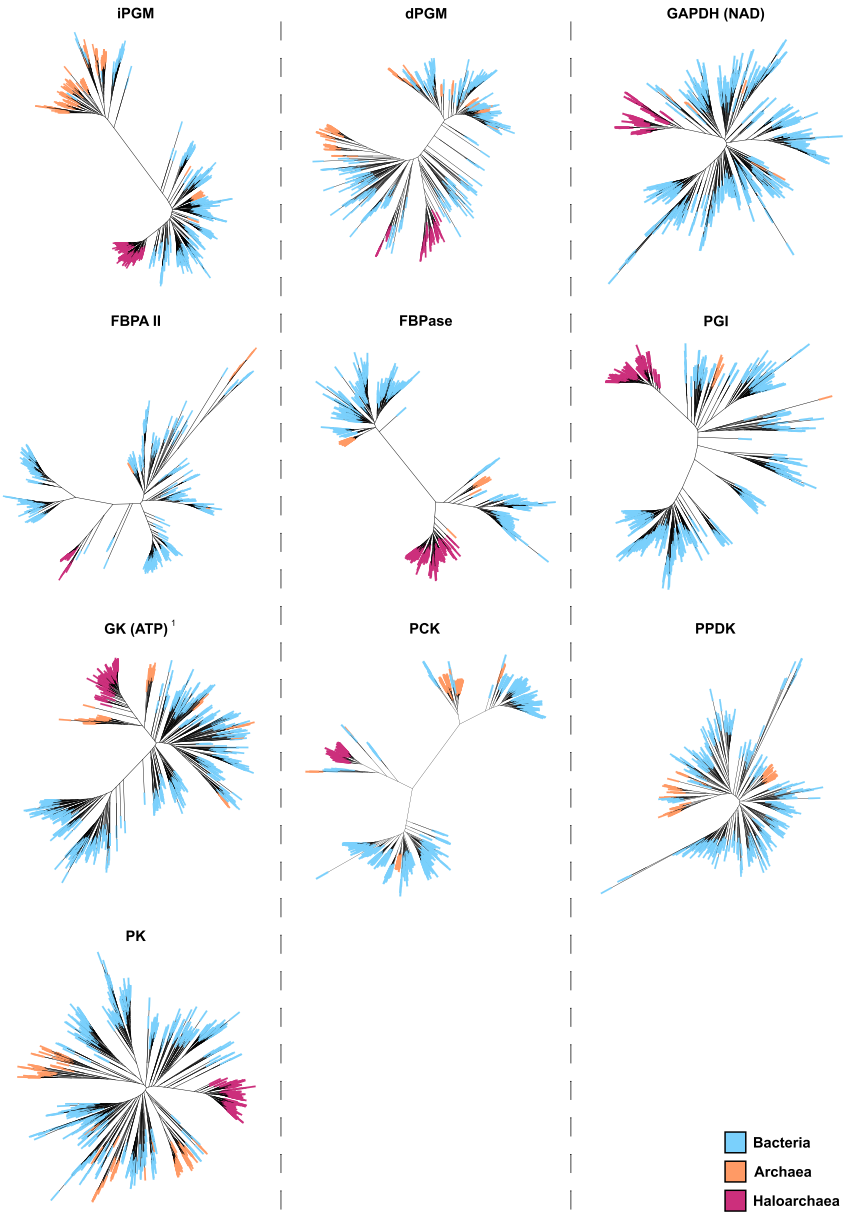


Figure 8. Unrooted gene trees of glycolytic and gluconeogenic enzymes that either show a bacterial origin of halophilic (magenta) enzymes or that are more common to bacterial species (blue) with few transfers to archaea (orange). Abbreviations of the corresponding enzymes are displayed above each tree and the arrangement of phylogenies (top left to bottom right) is equivalent to the order shown in the presence–absence matrix (Fig. 3). [1]HMMER and the PFAM-A database were used to differentiate between glucokinases and ROK (repressor, ORF, kinase) family glucokinases (Hansen et al. 2002, Hansen and Schönheit 2003) within the ATP-GK phylogeny (Fig. S3). [3]Phosphoenolpyruvate carboxykinases are either ATP- or GTP-dependent (Bräsen et al. 2014), the tips of the phylogeny were colored based on the NTP-dependency of each sequence (Fig. S7). Phylogenies were visually modified utilizing the tool Interactive Tree of Life (Letunic and Bork 2021).

Table 2. Glycolytic enzymes that were horizontally transferred from bacteria to halophilic archaea. If more than one bacterial taxon was found in the sister group of halophiles, the apparent bacterial donor lineage was stated as "mixed sister."

Enzyme	Apparent bacterial donor lineage(s)
Phosphoglycerate mutase, cofactor-independent	Actinomycetota
Phosphoglycerate mutase, cofactor-dependent	Mixed sister[a]
Glyeraldehyde-3-phosphate dehydrogenase	Bacillota
Fructose-1,6-bisphosphate aldolase class II	Mixed sister[b]
Phosphoglucose isomerase	Synergistota
Glucokinase, ATP-dependent	Mixed sister[c]
PEP carboxykinase, ATP-dependent	Mixed sister[d]
Pyruvate kinase	Actinomycetota

[a]Pseudomonadota, Bacillota, Chloroflexota, and Actinomycetota
[b]Ca. Cloacimonadota, Ca. Bipolaricaulota
[c]Actinomycetota and Deinococcota
[d]Bacillota, Planctomycetota, Coprothermobacterota, Synergistota, and Thermotogota

What do methanogens do with glycogen?

Methanogens have highly specialized carbon and energy metabolism in which carbon is assimilated via the acetyl-CoA pathway and ATP is synthesized through methanogenesis. They cannot grow on exogenous glucose or other sugars, raising the question as to how do methanogens obtain benefit from glycogen. Studies in various methanogens are consistent with a carbon storage function. König et al. (1985) showed that glycogen accumulated as cytosolic granules in *Methanolobus* and *Methanococcus* and became depleted upon starvation accompanied by methane production. Murray and Zinder (1987) found that the acetoclastic methanogen *Methanosarcina thermophila* accumulated reserve polysaccharide up to 20 mg/g protein when grown under excess carbon and nitrogen limitation. Pellerin et al. (1987) found that *Methanothrix* str. FE accumulates glycogen up to 2% of cell mass by dry weight as cytosolic particles of small polymers with ∼13 glucose each. Maitra et al. (2001) reported that *Methanococcoides methylutens* grown on methanol accumulates glycogen up to 36 mg/g protein and slowly generates methane upon subsequent starvation. Yu et al. (1994) studied glycogen metabolism in *Methanococcus maripaludis*, which grows on H_2 and CO_2 or formate and accumulates ∼0.3% glycogen by dry weight. They found high activities of ED enzymes and of enzymes involved in the nonoxidative pentose phosphate pathway, which supplies essential intermediates for nucleotide synthesis, whereby the synthesis of aromatic amino acids in methanogens differs from the classical erythrose-4-phosphate-dependent shikimate pathway, starting from aspartate semialdehyde and 6-deoxy-5-ketohexose 1-phosphate (White 2004).

A thorough study of GS and mobilization enzymes in *Methanosarcina acetivorans* was reported by Santiago-Martinez et al. (2016), who found that glycogen accumulated during growth and was converted into CH_4 and CO_2 in the stationary phase, also consistent with a storage function as reserve polysaccharide. Although the enzymatic machinery is available to generate ATP from glycogen in *M. acetivorans*, its ATP synthesis from glycogen proceeds via conversion into intermediates for methanogenesis, not via SLP. While methanogens can be engineered to grow for several cell divisions from SLP (Bock and Schönheit 1995, Schöne et al. 2022), none are known to use SLP for energy metabolism in nature. The apparent inability of methanogens to use external or internal glucose or pyruvate (Bock et al. 1994) for growth via SLP indicates that their carbon and energy metabolism are inextricably intertwined. Together with the status of methanogens as the most ancient lineages of archaea (Williams et al. 2017, Mei et al. 2023) and the geochemical evidence for methane production over 3.5 billion years ago (Ueno et al. 2006), the inability of methanogens to grow without methanogenesis—from either external or internal carbohydrates—is most easily understood as a holdover from the ancestrally chemiosmotic nature of the first prokaryotic cells (Martin and Russell 2007, Weiss et al. 2016, Schönheit et al. 2016) and the closure of redox-balanced stoichiometry in coupled carbon and energy metabolism in methanogenesis via the origin of flavin based electron bifurcation (Brabender et al. 2024, Martin and Kleinermanns 2024).

An observant reader posed an excellent question concerning the reasons underlying the different trajectories of the ability to use external glucose in methanogens and acetogens. These lie in the distant past and are, of course, obscure, but there are observations from which one can reason. We start with the circumstance that methanogens are, in all cases studied so far, never facultative: a cell is either a methanogen, obtaining ATP from methane synthesis, or not. Because evolution has already performed the experiment of introducing countless genes on countless occasions into methanogens, without ever having produced facultative methanogens (as far as we know so far), we can conclude that methanogenesis does not combine with any other form of energy metabolism. It can be replaced, as in the case of halophiles. But it does not produce physiologically useful combinations that would offer methanogens access to diverse substrates or environments, while preserving their ability to perform methanogenesis. At the same time, methanogenesis has not yet been shown to be transferrable to other lineages, in contrast to photosynthesis, for example, which also involves many genes but is readily transferred (Brinkmann et al. 2018). Methanogens can be forced to grow on pyruvate (Bock et al. 1994), or engineered to grow as acetogens (Schöne et al. 2022), but in nature they seem to be extreme specialists in terms of energy conservation. Methanogens are, in maize breeder's terms, poor combiners. Acetogens are, by contrast, extremely versatile in terms of their substrates to support carbon and energy metabolism (Müller 2003, Döning and Müller 2018).

Were one to try to pin this difference on one enzyme or step, the use of MtrA-H to generate an ion gradient for ATP synthesis in methanogens might be the prime suspect. MtrA-H accepts methyl groups from methyl-tetrahydromethanopterin, transfers them first to the Co atom in cobamide, then to CoM, and that 2-step reaction (each roughly –15kJ/mol) powers the pumping of a Na^+ ion across the membrane (Gottschalk and Thauer 2001). All methanogens, to our knowledge, use this coupling reaction in ATP synthesis. On the bottom line, energy metabolism in methanogens

runs on a short methyl transport chain. In acetogens, Rnf and Ech pump ions with the help of electron transfers: ferredoxin to NADH in the case of Rnf (Kumar et al. 2025), ferredoxin to H_2 in the case of Ech (Schoelmerich and Müller 2019). The methanogen strategy of the *methyl* transport chain does not readily lend itself well to energy conservation using carbohydrates, in contrast to methyl-containing environmental substrates, such as acetate or methylamines (which entail pyrrolysine for catabolism; Krzycki 2005). By contrast, the acetogen strategy of a short *electron* transport chain can be coupled to countless donors and acceptors, such that any oxidizable substrate, such as glucose, lactate, or any nonfermentable substrate, with a suitable acceptor can be plugged into the circuit, the circuit expanded and modified. Add cytochromes, and modern looking respiratory chains can start to emerge (Calisto and Pereira 2021). This is the kind of strategy that some acetogens have taken, while retaining the facultative ability to pump using acetate synthesis from H_2 and CO_2.

Thus, it is possible that methanogens were never able to compete effectively with fast carbohydrate fermenters or efficient carbohydrate respirers, although they occasionally became carbohydrate respirers through gene transfer, in which case they could compete well, but at the price of relinquishing methanogenesis (Nelson-Sathi et al. 2012). That is, methanogens clearly evolved the ability to store and utilize glucose from internal stores, but apparently utilizing it at their own intracellular pace, and were never able to effectively compete for carbohydrates as an environmental substrate.

Conclusions

The data indicate a gluconeogenic origin of the glycolytic EM pathway (Romano and Conway 1996, Ronimus and Morgan 2003). This is in line with autotrophic theories for the origin of metabolism emergent from the acetyl-CoA pathway of CO_2 reduction, a line of physiological reasoning with a very long tradition going back over 100 years (Mereschkowsky 1910, Fuchs and Stupperich 1985, Wächtershäuser 1992, Schönheit et al. 2016, Martin 2020).

Trunk glycolysis is the most conserved segment of the glycolytic pathway because of its function in the gluconeogenic direction to provide C3 (and larger) backbones for the synthesis of amino acids, bases, and cofactors. Trunk glycolysis is also the segment of sugar phosphate metabolism that is least often transferred between bacteria and archaea, for reasons that are not immediately evident, but almost certainly relate to function.

LGT between bacteria and archaea (and implicitly within both domains, though not the subject of this study) is otherwise widespread among glycolytic/gluconeogenic enzymes, a reflection of their widespread interchangeability and utility as providers of carbon backbones for carbon and energy metabolism.

The Haloarchaea represent a unique case in the evolution of sugar phosphate metabolism in that they obtained almost their entire glycolytic pathway from bacteria in what was very likely a symbiotic association (Nelson-Sathi et al. 2012) that paralleled the origin of the bacterial glycolytic pathway of eukaryotes (Martin and Müller 1998). There are earlier (Nelson-Sathi et al. 2015) and more recent reports (Baker et al. 2025) that the origin of several archaeal groups corresponds to gene acquisitions from bacteria.

The universal distribution of glycogen synthases among bacteria and archaea indicates that glycogen was present at least in the last common ancestors of bacteria and archaea, and possibly in the LUCA. The universal distribution of glycogen and glycogen synthases among strictly autotrophic methanogens explains the otherwise puzzling observation that methanogens can obtain carbon and energy from glucose that they synthesize and accumulate as an intracellular polymer (Gonzales-Ordenes et al. 2024), but cannot import it from the environment or use exogenous glucose for growth (Bock et al. 1994, Schöne et al. 2022, Richter et al. 2024). Thus, glycogen formed within cells might be considered as one of the first heterotrophic substrates in the evolution of central carbon metabolism.

The ability of methanogens to synthesize glucose from H_2 and CO_2, to sequester glucose as intracellular glycogen, and to utilize it as a self-made carbon source, coupled with their inability to utilize environmental glucose is a strong line of physiological evidence in favor of the autotrophic origin of metabolism. The observation that native transition metals, in particular nickel, can synthesize pyruvate at physiological concentrations under the conditions of serpentinizing hydrothermal vents (Beyazay et al. 2023a) from H_2 and CO_2 is an independent and purely chemical line of evidence in favor of autotrophic origins. The simplest interpretation of the current observations is that metabolism really did emerge along the lines of the acetyl-CoA pathway and that methanogens and acetogens, H_2-dependent chemolithoautotrophs that use the acetyl-CoA pathway in carbon and energy metabolism, really are the starting points of biological evolution (Martin and Russell 2007, Weiss et al. 2016, Colman et al. 2025) in the bacterial and archaeal domains, respectively.

Acknowledgments

Computational infrastructure and support were provided by the Centre for Information and Media Technology at Heinrich Heine University Düsseldorf. We thank Loraine Schwander for preparing the balanced prokaryotic dataset as well as Rebecca Gerhards, Jessica Wimmer, Nils Kapust, and Michael Knopp for discussions and advice on computational work.

Supplementary data

Supplementary data is available at *FEMSRE* online.

Conflict of interest: None declared.

Funding

This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement number 101018894). W.F.M. thanks the ERC (101018894) for funding.

References

Adams MWW. Biochemical diversity among sulfur-dependent, hyperthermophilic microorganisms. *FEMS Microbiol Rev* 1994;**15**:261–7. https://doi.org/10.1111/j.1574-6976.1994.tb00139.x.

Albers SV, Meyer BH. The archaeal cell envelope. *Nat Rev Microbiol* 2011;**9**:414–26. https://doi.org/10.1038/nrmicro2576.

Alefounder PR, Baldwin SA, Perham RN *et al.* Cloning, sequence analysis and over-expression of the gene for the class II fructose 1,6-bisphosphate aldolase of *Escherichia coli*. *Biochem J* 1989;**257**:529–34. https://doi.org/10.1042/bj2570529.

Allen JF, Thake B, Martin WF. Nitrogenase inhibition limited oxygenation of Earth's proterozoic atmosphere. *Trends Plant Sci* 2019;**24**:1022–103. https://doi.org/10.1016/j.tplants.2019.07.007.

Baker BA, McCarthy CGP, López-García P *et al.* Phylogenomic analyses indicate the archaeal superphylum DPANN originated from free-living euryarchaeal-like ancestors. *Nat Microbiol* 2025;**10**:1593–604. https://doi.org/10.1038/s41564-025-02024-5.

Beber ME, Gollub MG, Mozaffari D *et al.* eQuilibrator 3.0: a database solution for thermodynamic constant estimation. *Nucleic Acids Res* 2023;**50**:D603–9. https://doi.org/10.1093/nar/gkab1106.

Belthle KS, Tüysüz H. Linking catalysis in biochemical and geochemical CO_2 fixation at the emergence of life. *ChemCatChem* 2023;**15**:e202201462. https://doi.org/10.1002/cctc.202201462.

Berg IA. Ecological aspects of the distribution of different autotrophic CO_2 fixation pathways. *Appl Environ Microbiol* 2011;**77**:1925–36. https://doi.org/10.1128/AEM.02473-10.

Beyazay T, Ochoa-Hernández C, Song Y *et al.* Influence of composition of nickel-iron nanoparticles for abiotic CO_2 conversion to early prebiotic organics. *Angew Chem* 2023a;**62**:e202218189. https://doi.org/10.1002/anie.202218189.

Beyazay T, Martin WF, Tüysüz H. Direct synthesis of formamide from CO_2 and H_2O with nickel-iron nitride heterostructures under mild hydrothermal conditions. *J Am Chem Soc* 2023b;**145**:19768–79. https://doi.org/10.1021/jacs.3c05412.

Beyazay T, Belthle KS, Farès C *et al.* Ambient temperature CO_2 fixation to pyruvate and subsequently to citramalate over iron and nickel nanoparticles. *Nat Commun* 2023c;**14**:570. https://doi.org/10.1038/s41467-023-36088-w.

Bian X, Jiang H, Meng Y *et al.* Regulation of gene expression by glycolytic and gluconeogenic enzymes. *Trends Cell Biol* 2022;**32**:786–99. https://doi.org/10.1016/j.tcb.2022.02.003.

Biester A, Grahame DA, Drennan CL. Capturing a methanogenic carbon monoxide dehydrogenase/acetyl-CoA synthase complex via cryogenic electron microscopy. *Proc Nat Acad Sci USA* 2024;**121**:e2410995121. https://doi.org/10.1073/pnas.2410995121.

Bock AK, Prieger-Kraft A, Schönheit P. Pyruvate—a novel substrate for growth and methane formation in *Methanosarcina barkeri*. *Arch Microbiol* 1994;**161**:33–46.

Bock AK, Schönheit P. Growth of *Methanosarcina barkeri* (Fusaro) under nonmethanogenic conditions by the fermentation of pyruvate to acetate: ATP synthesis via the mechanism of substrate level phosphorylation. *J Bacteriol* 1995;**177**:2002–7. https://doi.org/10.1128/jb.177.8.2002-2007.1995.

Brabender M, Henriques Pereira DP, Mrnjavac N *et al.* Ferredoxin reduction by hydrogen with iron functions as an evolutionary precursor of flavin-based electron bifurcation. *Proc Nat Acad Sci USA* 2024;**121**:e2318969121. https://doi.org/10.1073/pnas.2318969121.

Bräsen C, Esser D, Rauch B *et al.* Carbohydrate metabolism in archaea: current insights into unusual enzymes and pathways and their regulation. *MMBR* 2014;**78**:89–175. https://doi.org/10.1128/MMBR.00041-13.

Brinkmann H, Göker M, Koblížek M *et al.* Horizontal operon transfer, plasmids, and the evolution of photosynthesis in Rhodobacteraceae. *ISME J* 2018;**12**:1994–2010. https://doi.org/10.1038/s41396-018-0150-9.

Buchfink B, Reuter K, Drost HG. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat Methods* 2021;**18**:366–8. https://doi.org/10.1038/s41592-021-01101-x.

Budgen N, Danson MJ. Metabolism of glucose via a modified Entner-Doudoroff pathway in the thermoacidophilic archaebacterium, *Thermoplasma acidophilum*. *FEBS Lett* 1986;**196**:207–10. https://doi.org/10.1016/0014-5793(86)80247-2.

Calisto F, Pereira MM. Modularity of membrane-bound charge-translocating protein complexes. *Biochem Soc Trans* 2021;**49**:2669–85. https://doi.org/10.1042/BST20210462.

Chandra G, Chater KF, Bornemann S. Unexpected and widespread connections between bacterial glycogen and trehalose metabolism. *Microbiology* 2011;**157**:1565–72. https://doi.org/10.1099/mic.0.044263-0.

Chen X, Schreiber K, Appel J *et al.* The Entner-Doudoroff pathway is an overlooked glycolytic route in cyanobacteria and plants. *Proc Nat Acad Sci USA* 2016;**113**:5441–6. https://doi.org/10.1073/pnas.1521916113.

Cifuente JO, Colleoni C, Kalscheuer R *et al.* Architecture, function, regulation, and evolution of α-glucans metabolic enzymes in prokaryotes. *Chem Rev* 2024;**124**:4863–934. https://doi.org/10.1021/acs.chemrev.3c00811.

Cifuente JO, Comino N, Trastoy B *et al.* Structural basis of glycogen metabolism in bacteria. *Biochem J* 2019;**476**:2059–92. https://doi.org/10.1042/BCJ20170558.

Colman DR, Kraus EA, Thieringer PH *et al.* Deep-branching acetogens in serpentinized subsurface fluids of Oman. *Proc Nat Acad Sci USA* 2022;**119**:e2206845119. https://doi.org/10.1073/pnas.2206845119.

Colman DR, Templeton AS, Spear JR *et al.* Microbial ecology of serpentinite-hosted ecosystems. *ISME J* 2025;**19**:wraf029. https://doi.org/10.1093/ismejo/wraf029.

Cori CF, Cori GT. Mechanism of formation of hexosemonophosphate in muscle and isolation of a new phosphate ester. *Exp Biol Med* 1936;**34**:702–5. https://doi.org/10.3181/00379727-34-8759P.

Dandekar T, Schuster S, Snel B *et al.* Pathway alignment: application to the comparative analysis of glycolytic enzymes. *Biochem J* 1999;**343**:115–24. https://doi.org/10.1042/bj3430115.

Decker K, Jungermann K, Thauer R. Energy production in anaerobic organisms. *Angew Chem* 1970;**9**:138–58. https://doi.org/10.1002/anie.197001381.

De Duve C. *Blueprint for a Cell: The Nature and Origin of Life*. Burlington, NC: Neil Patterson Publishers, 1991.

Degani C, Halmann M. Chemical evolution of carbohydrate metabolism. *Nature* 1967;**216**:1207. https://doi.org/10.1038/2161207a0.

Döning J, Müller V. Alanine, a novel growth substrate for the acetogenic bacterium *Acetobacterium woodii*. *Appl Environ Microbiol* 2018;**84**:e02023–18.

Doukov TI, Iverson TM, Seravalli J *et al.* A Ni-Fe-Cu center in a bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase. *Science* 2002;**298**:567–72. https://doi.org/10.1126/science.1075843.

Eddy SR Accelerated profile HMM searches. *PLoS Comput Biol* 2011;**7**:e1002195. https://doi.org/10.1371/journal.pcbi.1002195.

Egan AJF, Cleverly RM, Peters K *et al.* Regulation of bacterial cell wall growth. *FEBS J* 2017;**284**:851–67. https://doi.org/10.1111/febs.13959.

Flamholz A, Noor E, Bar-Even A *et al.* Glycolytic strategy as a trade-off between energy yield and protein cost. *Proc Nat Acad Sci USA* 2013;**110**:10039–44. https://doi.org/10.1073/pnas.1215283110.

Fothergill-Gilmore LA, Michels PAM. Evolution of glycolysis. *Prog Biophys Mol Biol* 1993;**59**:105–235. https://doi.org/10.1016/0079-6107(93)90001-Z.

Fuchs G, Stupperich E. Evidence for an incomplete reductive carboxylic acid cycle in *Methanobacterium thermoautotrophicum*. *Arch Microbiol* 1978;**118**:121–25. https://doi.org/10.1007/BF00406084.

Fuchs G, Stupperich E. Evolution of autotrophic CO_2 fixation. In: Schleifer KH, Stackebrandt E (eds), *Evolution of Prokaryotes*. London: Academic Press, 1985, 235–25.

Fuchs G. CO_2 fixation in acetogenic bacteria: variations on a theme. *FEMS Microbiol Rev* 1986;**2**:181–213. https://doi.org/10.1111/j.1574-6968.1986.tb01859.x.

Fuchs G. Alternative pathways of carbon dioxide fixation: insights into the early evolution of life? *Annu Rev Microbiol* 2011;**65**:631–58. https://doi.org/10.1146/annurev-micro-090110-102801.

Furdui C, Ragsdale SW. The role of pyruvate ferredoxin oxidoreductase in pyruvate synthesis during autotrophic growth by the Wood-Ljungdahl pathway. *J Biol Chem* 2000;**275**:28494–9. https://doi.org/10.1074/jbc.M003291200.

Gonzalez-Ordenes F, Herrera-Soto N, Muñoz SM *et al.* Glycogen metabolism in methanogens: a key pathway for metabolic response to nutrient availability. *Biochem Biophys Res Commun* 2024;**739**:150978. https://doi.org/10.1016/j.bbrc.2024.150978.

Gottschalk G, Thauer RK. The Na(+)-translocating methyltransferase complex from methanogenic archaea. *Biochim Biophys Acta* 2001;**1505**:28–36. https://doi.org/10.1016/S0005-2728(00)00274-7.

Gupta MN, Uversky VN. Moonlighting enzymes: when cellular context defines specificity. *CMLS* 2023;**80**:130. 24. https://doi.org/10.1007/s00018-023-04781-0.

Hansen T, Oehlmann M, Schönheit. P. Novel type of glucose-6-phosphate isomerase in the hyperthermophilic archaeon *Pyrococcus furiosus*. *J Bacteriol* 2001;**183**:3428–35. https://doi.org/10.1128/JB.183.11.3428-3435.2001.

Hansen T, Reichstein B, Schmid R *et al.* The first archaeal ATP-dependent glucokinase, from the hyperthermophilic crenarchaeon *Aeropyrum pernix*, represents a monomeric, extremely thermophilic ROK glucokinase with broad hexose specificity. *J Bacteriol* 2002;**184**:5955–65. https://doi.org/10.1128/JB.184.21.5955-5965.2002.

Hansen T, Schlichting B, Felgendreher M *et al.* Cupin-type phosphoglucose isomerases (Cupin-PGIs) constitute a novel metal-dependent PGI family representing a convergent line of PGI evolution. *J Bacteriol* 2005;**187**:1621–31. https://doi.org/10.1128/JB.187.5.1621-1631.2005.

Hansen T, Schönheit P. Sequence, expression, and characterization of the first archaeal ATP-dependent 6-phosphofructokinase, a non-allosteric enzyme related to the phosphofructokinase-B sugar kinase family, from the hyperthermophilic crenarchaeote *Aeropyrum pernix*. *Arch Microbiol* 2001;**177**:62–9. https://doi.org/10.1007/s00203-001-0359-1.

Hansen T, Schönheit P. ATP-dependent glucokinase from the hyperthermophilic bacterium *Thermotoga maritima* represents an extremely thermophilic ROK glucokinase with high substrate specificity. *FEMS Microbiol Lett* 2003;**226**:405–11. https://doi.org/10.1016/S0378-1097(03)00642-6.

Hansen T, Wendorff D, Schönheit P. Bifunctional phosphoglucose/phosphomannose isomerases from the Archaea *Aeropyrum pernix* and *Thermoplasma acidophilum* constitute a novel enzyme family within the phosphoglucose isomerase superfamily. *J Biol Chem* 2004;**279**:2262–72. https://doi.org/10.1074/jbc.M309849200.

Hartmann E, König H. A novel pathway of peptide biosynthesis found in methanogenic archaea. *Arch Microbiol* 1994;**162**:430–2. https://doi.org/10.1007/BF00282108.

Heider J, Ma K, Adams MW. Purification, characterization, and metabolic function of tungsten-containing aldehyde ferredoxin oxidoreductase from the hyperthermophilic and proteolytic archaeon *Thermococcus* strain ES-1. *J Bacteriol* 1995;**177**:4757–64. https://doi.org/10.1128/jb.177.16.4757-4764.1995.

Henriques Pereira DP, Leethaus J, Beyazay T *et al.* Role of geochemical protoenzymes (geozymes) in primordial metabolism: specific abiotic hydride transfer by metals to the biological redox cofactor NAD^+. *FEBS J* 2022;**289**:3148–62. https://doi.org/10.1111/febs.16329.

Hensel R, Zwickl P, Fabry S *et al.* Sequence comparison of glyceraldehyde-3-phosphate dehydrogenases from the three urkingdoms: evolutionary implication. *Can J Microbiol* 1989;**35**:81–5. https://doi.org/10.1139/m89-012.

Huber H, Hohn MJ, Rachel R *et al.* A new phylum of archaea represented by a nanosized hyperthermophilic symbiont. *Nature* 2002;**417**:63–7. https://doi.org/10.1038/417063a.

Hutchins AM, Holden JF, Adams MW. Phosphoenolpyruvate synthetase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J Bacteriol* 2001;**183**:709–15. https://doi.org/10.1128/JB.183.2.709-715.2001.

Iddar A, Valverde F, Assobhei O *et al.* Widespread occurrence of non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase among Gram-positive bacteria. *Int Microbiol* 2005;**8**:251–8.

Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013;**30**:772–80. https://doi.org/10.1093/molbev/mst010.

Kaur H, Rauscher SA, Werner E *et al.* A prebiotic Krebs cycle analog generates amino acids with H_2 and NH_3 over nickel. *Chem* 2024;**10**:1528–40. https://doi.org/10.1016/j.chempr.2024.02.001.

Keller A, Schink B, Müller N. Energy-conserving enzyme systems active during syntrophic acetate oxidation in the thermophilic bacterium *Thermacetogenium phaeum*. *Front Microbiol* 2019;**10**:2785. https://doi.org/10.3389/fmicb.2019.02785.

Keller MA, Turchyn AV, Ralser M. Non-enzymatic glycolysis and pentose phosphate pathway-like reactions in a plausible Archean ocean. *Mol Syst Biol* 2014;**10**:725. 25. https://doi.org/10.1002/msb.20145228.

Kelly S, Wickstead B, Gull K. Archaeal phylogenomics provides evidence in support of a methanogenic origin of the archaea and a thaumarchaeal origin for the eukaryotes. *Proc Biol Sci* 2011;**278**:1009–18.

Kennedy SP, Ng WV, Salzberg SL *et al.* Understanding the adaptation of *Halobacterium* species NRC-1 to its extreme environment through computational analysis of its genome sequence. *Genome Res* 2001;**11**:1641–50. https://doi.org/10.1101/gr.190201.

Kim JW, Dang CV. Multifaceted roles of glycolytic enzymes. *Trends Biochem Sci* 2005;**30**:142–50. https://doi.org/10.1016/j.tibs.2005.01.005.

König H, Nusser E, Stetter KO. Glycogen in *Methanolobus* and *Methanococcus*. *FEMS Microbiol Lett* 1985;**28**:265–9. https://doi.org/10.1111/j.1574-6968.1985.tb00803.x.

König H, Skorko R, Zillig W *et al.* Glycogen in thermoacidophilic Acaebacteria of the genera *Sulfolobus*, *Thermoproteus*, *Desulfurococcus* and *Thermococcus*. *Arch Microbiol* 1982;**132**:297–303.

Krzycki JA. The direct genetic encoding of pyrrolysine. *Curr Opin Microbiol* 2005;**8**:706–12. https://doi.org/10.1016/j.mib.2005.10.009.

Kumar A, Roth J, Kim H *et al.* Molecular principles of redox-coupled sodium pumping of the ancient Rnf machinery. *Nat Commun* 2025;**16**:2302. https://doi.org/10.1038/s41467-025-57375-8.

Kuprat T, Johnsen U, Ortjohann M *et al.* Acetate metabolism in archaea: characterization of an acetate transporter and of enzymes involved in acetate activation and gluconeogenesis in *Haloferax volcanii*. *Front Microbiol* 2020;**11**:604926. https://doi.org/10.3389/fmicb.2020.604926.

Kuprat T, Ortjohann M, Johnsen U *et al.* Glucose metabolism and acetate switch in archaea: the enzymes in *Haloferax volcanii*. *J Bacteriol* 2021;**203**:e00690–20. 23. https://doi.org/10.1128/JB.00690-20.

Lee A, Bae E, Park J *et al.* Identification of the genes related to the glycogen metabolism in hyperthermophilic archaeon, *Sulfolobus acidocaldarius*. *Front Microbiol* 2021;**12**:661053. https://doi.org/10.3389/fmicb.2021.661053.

Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 2021;**49**:W293–6. https://doi.org/10.1093/nar/gkab301.

Lipmann F. Metabolic generation and utilization of phosphate bond energy. In: Nord F.F., Werkman C.H. (eds), *Advances in Enzymology and Related Areas of Molecular Biology*. Hoboken, NJ: Wiley, 1941.

Lorentzen E, Hensel R, Knura T *et al.* Structural basis of allosteric regulation and substrate specificity of the non-phosphorylating glyceraldehyde 3-6 phosphate dehydrogenase from *Thermoproteus tenax*. *J Mol Biol* 2004;**341**:815–28. https://doi.org/10.1016/j.jmb.2004.05.032.

Ma Q, Jiang H, Ma L *et al.* The moonlighting function of glycolytic enzyme enolase-1 promotes choline phospholipid metabolism and tumor cell proliferation. *Proc Nat Acad Sci USA* 2023;**120**:e2209435120. https://doi.org/10.1073/pnas.2209435120.

Maitra PK, Bhosale SB, Kshirsagar DC *et al.* Metabolite and enzyme profiles of glycogen metabolism in *Methanococcoides methylutens*. *FEMS Microbiol Lett* 2001;**198**:23–9. https://doi.org/10.1111/j.1574-6968.2001.tb10614.x.

Makarova KS, Yutin N, Bell SD *et al.* Evolution of diverse cell division and vesicle formation systems in archaea. *Nat Rev Microbiol* 2010;**8**:731–41. https://doi.org/10.1038/nrmicro2406.

Martin WF, Baross J, Kelley D *et al.* Hydrothermal vents and the origin of life. *Nat Rev Microbiol* 2008;**6**:805–14. https://doi.org/10.1038/nrmicro1991.

Martin WF, Kleinermanns K. *The Geochemical Origin of Microbes*. Boca Raton, FL: CRC Press, 2024.

Martin WF, Müller M. The hydrogen hypothesis for the first eukaryote. *Nature* 1998;**392**:37–41. https://doi.org/10.1038/32096.

Martin WF, Russell MJ. On the origin of biochemistry at an alkaline hydrothermal vent. *Philo Trans R Soc Lond B Biol Sci* 2007;**362**:1887–925. https://doi.org/10.1098/rstb.2006.1881.

Martin WF, Schnarrenberger C. The evolution of the Calvin cycle from prokaryotic to eukaryotic chromosomes: a case study of functional redundancy in ancient pathways through endosymbiosis. *Curr Genet* 1997;**32**:1–18. https://doi.org/10.1007/s002940050241.

Martin WF, Tielens AGM, Mentel M *et al.* The physiology of phagocytosis in the context of mitochondrial origin. *MMBR* 2017;**81**:e00008–17. 14. https://doi.org/10.1128/MMBR.00008-17.

Martin WF. 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%3c99::AID-BIES3%3e3.0.CO;2-B.

Martin WF. Older than genes: the acetyl CoA pathway and origins. *Front Microbiol* 2020;**11**:817. 4. https://doi.org/10.3389/fmicb.2020.00817.

Mayer F, Müller V. Adaptations of anaerobic archaea to life under extreme energy limitation. *FEMS Microbiol Rev* 2014;**38**:449–72. https://doi.org/10.1111/1574-6976.12043.

Mayer RJ, Moran J. Metal ions turn on a stereoselective nonenzymatic reduction of keto acids by the coenzyme NADH. *Chem* 2024;**10**:2564–76. https://doi.org/10.1016/j.chempr.2024.05.007.

Mei R, Kaneko M, Imachi H *et al.* The origin and evolution of methanogenesis and archaea are intertwined. *PNAS Nexus* 2023;**2**:pgad023. https://doi.org/10.1093/pnasnexus/pgad023.

Mereschkowsky C. Theorie der zwei Plasmaarten als Grundlage der Symbiogenesis, einer neuen Lehre von der Entstehung der Organismen. *Biol Centralbl* 1910;**30**:278–88.

Modjewski LD, Karavaeva V, Mrnjavac N *et al.* Evidence for corrin biosynthesis in the last universal common ancestor. *FEBS J* 2024;**292**:827–50. https://doi.org/10.1111/febs.17367.

Morita T, Kawamoto H, Mizota I *et al.* Enolase in the RNA degradosome plays a crucial role in the rapid decay of glucose transporter mRNA in the response to phosphosugar stress in *Escherichia coli* *Mol Microbiol* 2004;**54**:1063–75. https://doi.org/10.1111/j.1365-2958.2004.04329.x.

Mrnjavac N, Degli Esposti M, Mizrahi I *et al.* Three enzymes governed the rise of O_2 on Earth. *Biochim Biophys Acta Bioenerg* 2024a;**1865**:149495. https://doi.org/10.1016/j.bbabio.2024.149495.

Mrnjavac N, Schwander L, Brabender M *et al.* Chemical antiquity in metabolism. *Acc Chem Res* 2024b;**57**:2267–78. https://doi.org/10.1021/acs.accounts.4c00226.

Muchowska KB, Varma SJ, Moran J. Nonenzymatic metabolic reactions and life's origins. *Chem Rev* 2020;**120**:7708–44. https://doi.org/10.1021/acs.chemrev.0c00191.

Müller M, Mentel M, van Hellemond JJ *et al.* Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *MMBR* 2012;**76**:444–95.

Müller V. Energy conservation in acetogenic bacteria. *Appl Environ Microbiol* 2003;**69**:6345–53.

Murray PA, Zinder SH. Polysaccharide reserve material in the acetotrophic methanogen, *Methanosarcina thermophila* strain TM-1: accumulation and mobilization. *Arch Microbiol* 1987;**147**:109–16. https://doi.org/10.1007/BF00415270.

Nagies FSP, Brueckner J, Tria FDK *et al.* A spectrum of verticality across genes. *PLoS Genet* 2020;**16**:e1009200. https://doi.org/10.1371/journal.pgen.1009200.

Nelson-Sathi S, Dagan T, Landan G *et al.* Acquisition of 1,000 eubacterial genes physiologically transformed a methanogen at the origin of Haloarchaea. *Proc Nat Acad Sci USA* 2012;**109**:20537–42. https://doi.org/10.1073/pnas.1209119109.

Nelson-Sathi S, Sousa FL, Roettger M *et al.* Origins of major archaeal clades correspond to gene acquisitions from bacteria. *Nature* 2015;**517**:77–80. https://doi.org/10.1038/nature13805.

Nobu MK, Nakai R, Tamazawa S *et al.* Unique H_2-utilizing lithotrophy in serpentinite-hosted systems. *ISME J* 2023;**17**:95–104. https://doi.org/10.1038/s41396-022-01197-9.

O'Leary NA, Wright MW, Brister JR *et al.* Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* 2016;**44**:D733–45.

Pan C, Li B, Celeste Simon M. Moonlighting functions of metabolic enzymes and metabolites in cancer. *Mol Cell* 2021;**81**:3760–74. https://doi.org/10.1016/j.molcel.2021.08.031.

Park HS, Park JT, Kang HK *et al.* TreX from *Sulfolobus solfataricus* ATCC 35092 displays isoamylase and 4-alpha-glucanotransferase activities. *Biosci Biotechnol Biochem* 2007a;**71**:1348–52. https://doi.org/10.1271/bbb.70016.

Park MO, Mizutani T, Jones RJ. Glyceraldehyde-3-phosphate ferredoxin oxidoreductase from *Methanococcus maripaludis*. *J Bacteriol* 2007b;**189**:7281–9. https://doi.org/10.1128/JB.00828-07.

Pellerin P, Gruson B, Prensier G *et al.* Glycogen in *Methanothrix* *Arch Microbiol* 1987;**146**:377–81. https://doi.org/10.1007/BF00410939.

Perlman SJ, Hunter MS, Zchori-Fein E. The emerging diversity of *Rickettsia*. *Proc Biol Sci* 2006;**273**:2097–106.

Pickl A, Johnsen U, Schönheit P. Fructose degradation in the haloarchaeon *Haloferax volcanii* involves a bacterial type phosphoenolpyruvate-dependent phosphotransferase system, fructose-1-phosphate kinase, and class II fructose-1,6-bisphosphate aldolase. *J Bacteriol* 2012;**194**:3088–97. https://doi.org/10.1128/JB.00200-12.

Pickl A, Schönheit P. The oxidative pentose phosphate pathway in the haloarchaeon *Haloferax volcanii* involves a novel type of glucose-6-phosphate dehydrogenase–the archaeal Zwischenferment. *FEBS Lett* 2015;**589**:1105–11. https://doi.org/10.1016/j.febslet.2015.03.026.

Preiner M, Igarashi K, Muchowska KB *et al.* A hydrogen-dependent geochemical analogue of primordial carbon

and energy metabolism. *Nat Ecol Evol* 2020;**4**:534–42. https://doi.org/10.1038/s41559-020-1125-6.
Preiss J. Glycogen: biosynthesis and regulation. *EcoSal Plus* 2014;**6**:1. https://doi.org/10.1128/ecosalplus.esp-0015-2014.
Punta M, Coggill PC, Eberhardt RY *et al.* The Pfam protein families database. *Nucleic Acids Res* 2012;**40**:D290–301. https://doi.org/10.1093/nar/gkr1065.
Ragsdale SW, Pierce E. Acetogenesis and the Wood-Ljungdahl pathway of CO_2 fixation. *Biochim Biophys Acta* 2008;**1784**:1873–98. https://doi.org/10.1016/j.bbapap.2008.08.012.
Ragsdale SW. Enzymology of the Wood-Ljungdahl pathway of acetogenesis. *Ann N Y Acad Sci* 2008;**1125**:129–36. https://doi.org/10.1196/annals.1419.015.
Reher M, Fuhrer T, Bott M *et al.* The nonphosphorylative Entner-Doudoroff pathway in the thermoacidophilic euryarchaeon *Picrophilus torridus* involves a novel 2-Keto-3-deoxygluconate-specific aldolase. *J Bacteriol* 2010;**192**:964–74. https://doi.org/10.1128/JB.01281-09.
Reher M, Gebhard S, Schönheit P. Glyceraldehyde-3-phosphate ferredoxin oxidoreductase (GAPOR) and nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase (GAPN), key enzymes of the respective modified Embden-Meyerhof pathways in the hyperthermophilic crenarchaeota *Pyrobaculum aerophilum* and *Aeropyrum pernix*. *FEMS Microbiol Lett* 2007;**273**:196–205.
Reher M, Schönheit P. Glyceraldehyde dehydrogenases from the thermoacidophilic euryarchaeota *Picrophilus torridus* and *Thermoplasma acidophilum*, key enzymes of the non-phosphorylative Entner-Doudoroff pathway, constitute a novel enzyme family within the aldehyde dehydrogenase superfamily. *FEBS Lett* 2006;**580**:1198–204.
Reinhardt A, Johnsen U, Schönheit P. L-rhamnose catabolism in archaea. *Mol Microbiol* 2019;**111**:1093–108. https://doi.org/10.1111/mmi.14213.
Richter M, Sattler C, Schöne C *et al.* Pyruvate-dependent growth of *Methanosarcina acetivorans*. *J Bacteriol* 2024;**206**:e0036323. https://doi.org/10.1128/jb.00363-23.
Romano AH, Conway T. Evolution of carbohydrate metabolic pathways. *Res Microbiol* 1996;**147**:448–55. https://doi.org/10.1016/0923-2508(96)83998-2.
Ronimus RS, Morgan HW. Distribution and phylogenies of enzymes of the Embden-Meyerhof-Parnas pathway from archaea and hyperthermophilic bacteria support a gluconeogenic origin of metabolism. *Archaea* 2003;**1**:199–221. https://doi.org/10.1155/2003/162593.
Sakuraba H, Ohshima T. Novel energy metabolism in anaerobic hyperthermophilic archaea: a modified Embden-Meyerhof pathway. *J Biosci Bioeng* 2002;**93**:441–8. https://doi.org/10.1016/S1389-1723(02)80090-1.
Santiago-Martínez MG, Encalada R, Lira-Silva E *et al.* The nutritional status of *Methanosarcina acetivorans* regulates glycogen metabolism and gluconeogenesis and glycolysis fluxes *FEBS J* 2016;**283**:1979–99. https://doi.org/10.1111/febs.13717.
Sato T, Atomi H. Novel metabolic pathways in archaea. *Curr Opin Microbiol* 2011;**14**:307–14. https://doi.org/10.1016/j.mib.2011.04.014.
Sattler C, Richter M, Rother M. Physiological barriers for glucose utilization in *Methanosarcina acetivorans*. *Biorxiv* 2024. https://doi.org/10.1101/2024.08.21.609014.
Say RF, Fuchs G. Fructose 1,6-bisphosphate aldolase/phosphatase may be an ancestral gluconeogenic enzyme. *Nature* 2010;**464**:1077–81. https://doi.org/10.1038/nature08884.
Sayers EW, Bolton EE, Brister JR *et al.* Database resources of the national center for biotechnology information. *Nucleic Acids Res* 2022;**50**:D20–6. https://doi.org/10.1093/nar/gkab1112.
Schäfer T, Schönheit P. Gluconeogenesis from pyruvate in the hyperthermophilic archaeon *Pyrococcus furiosus*: involvement of reactions of the Embden-Meyerhof pathway. *Arch Microbiol* 1993;**159**:354–63. https://doi.org/10.1007/BF00290918.
Schlikker ML, Brabender M, Schwander L *et al.* Conversion of pyridoxal to pyridoxamine with NH_3 and H_2 on nickel generates a protometabolic nitrogen shuttle under serpentinizing conditions. *FEBS J* 2024;**292**:3041–55. https://doi.org/10.1111/febs.17357.
Schoelmerich MC, Müller V. Energy conservation by a hydrogenase-dependent chemiosmotic mechanism in an ancient metabolic pathway. *Proc Nat Acad Sci USA* 2019;**116**:6329–34. https://doi.org/10.1073/pnas.1818580116.
Schöne C, Poehlein A, Jehmlich N *et al.* Deconstructing *Methanosarcina acetivorans* into an acetogenic archaeon. *Proc Nat Acad Sci USA* 2022;**119**:e2113853119.
Schönheit P, Buckel W, Martin WF. On the origin of heterotrophy. *Trends Microbiol* 2016;**24**:12–25. https://doi.org/10.1016/j.tim.2015.10.003.
Schönheit P, Schäfer T. Metabolism of hyperthermophiles. *World J Microbiol Biotechnol* 1995;**11**:26–57. https://doi.org/10.1007/BF00339135.
Schwander L, Brabender M, Mrnjavac N *et al.* Serpentinization as the source of energy, electrons, organics, catalysts, nutrients and pH gradients for the origin of LUCA and life. *Front Microbiol* 2023;**14**:1257597. https://doi.org/10.3389/fmicb.2023.1257597.
Scott IM, Rubinstein GM, Poole FL, 2nd *et al.* The thermophilic biomass-degrading bacterium *Caldicellulosiruptor bescii* utilizes two enzymes to oxidize glyceraldehyde 3-phosphate during glycolysis. *J Biol Chem* 2019;**294**:9995–10005. https://doi.org/10.1074/jbc.RA118.007120.
Selig M, Xavier KB, Santos H *et al.* Comparative analysis of Embden-Meyerhof and Entner-Doudoroff glycolytic pathways in hyperthermophilic archaea and the bacterium *Thermotoga*. *Arch Microbiol* 1997;**167**:217–32. https://doi.org/10.1007/BF03356097.
Siebers B, Brinkmann H, Dörr C *et al.* Archaeal fructose-1,6-bisphosphate aldolases constitute a new family of archaeal type class I aldolase. *J Biol Chem* 2001;**276**:28710–8. https://doi.org/10.1074/jbc.M103447200.
Siebers B, Hensel R. Glucose catabolism of the hyperthermophilic archaeum *Thermoproteus tenax*. *FEMS Microbiol Lett* 1993;**111**:1–7. https://doi.org/10.1111/j.1574-6968.1993.tb06353.x.
Siebers B, Schönheit P. Unusual pathways and enzymes of central carbohydrate metabolism in archaea. *Curr Opin Microbiol* 2005;**8**:695–705. https://doi.org/10.1016/j.mib.2005.10.014.
Sleep NH, Bird DK, Pope EC. Serpentinite and the dawn of life. *Philo Trans R Soc Lond B Biol Sci* 2011;**366**:2857–69. https://doi.org/10.1098/rstb.2011.0129.
Sleep NH. Geological and geochemical constraints on the origin and evolution of life. *Astrobiology* 2018;**18**:1199–219. https://doi.org/10.1089/ast.2017.1778.
Sorokin DY, Makarova KS, Abbas B *et al.* Discovery of extremely halophilic, methyl-reducing euryarchaea provides insights into the evolutionary origin of methanogenesis. *Nat Microbiol* 2017;**2**:17081. https://doi.org/10.1038/nmicrobiol.2017.81.
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**:964–81. https://doi.org/10.1016/j.bbabio.2014.02.001.
Stadthagen G, Sambou T, Guerin M *et al.* Genetic basis for the biosynthesis of methylglucose lipopolysaccharides in *Mycobacterium tuberculosis*. *J Biol Chem* 2007;**282**:27270–6. https://doi.org/10.1074/jbc.M702676200.

Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014;**30**:1312–3.
Stouthamer AH. A theoretical study on the amount of ATP required for synthesis of microbial cell material. *Antonie Van Leeuwenhoek* 1973;**39**:545–65. https://doi.org/10.1007/BF02578899.
Subedi BP, Martin WF, Carbone V *et al.* Archaeal pseudomurein and bacterial murein cell wall biosynthesis share a common evolutionary ancestry. *FEMS Microbes* 2021;**2**:xtab012. https://doi.org/10.1093/femsmc/xtab012.
Suzuki E, Suzuki R. Distribution of glucan-branching enzymes among prokaryotes. *CMLS* 2016;**73**:2643–60. https://doi.org/10.1007/s00018-016-2243-9.
Tästensen JB, Johnsen U, Reinhardt A *et al.* D-galactose catabolism in archaea: operation of the DeLey-Doudoroff pathway in *Haloferax volcanii*. *FEMS Microbiol Lett* 2020;**367**:fnaa029. https://doi.org/10.1093/femsle/fnaa029.
Tästensen JB, Schönheit P. Two distinct glyceraldehyde-3-phosphate dehydrogenases in glycolysis and gluconeogenesis in the archaeon *Haloferax volcanii*. *FEBS Lett* 2018;**592**:1524–34. https://doi.org/10.1002/1873-3468.13037.
Trost K, Knopp MR, Wimmer JLE *et al.* A universal and constant rate of gene content change traces pangenome flux to LUCA. *FEMS Microbiol Lett* 2024;**371**:fnae068. https://doi.org/10.1093/femsle/fnae068.
Ueno Y, Yamada K, Yoshida N *et al.* Evidence from fluid inclusions for microbial methanogenesis in the early archaean era. *Nature* 2006;**440**:516–9. https://doi.org/10.1038/nature04584.
van der Oost J, Schut G, Kengen SW *et al.* The ferredoxin-dependent conversion of glyceraldehyde-3-phosphate in the hyperthermophilic archaeon *Pyrococcus furiosus* represents a novel site of glycolytic regulation. *J Biol Chem* 1998;**273**:28149–54. https://doi.org/10.1074/jbc.273.43.28149.
Verhees CH, Huynen MA, Ward DE *et al.* The phosphoglucose isomerase from the hyperthermophilic archaeon *Pyrococcus furiosus* is a unique glycolytic enzyme that belongs to the cupin superfamily. *J Biol Chem* 2001;**276**:40926–32. https://doi.org/10.1074/jbc.M104603200.
Verhees CH, Kengen SWM, Tuininga JE *et al.* The unique features of glycolytic pathways in archaea. *Biochem J* 2003;**375**:231–46. https://doi.org/10.1042/bj20021472.
Wächtershäuser G. Groundworks for an evolutionary biochemistry: the iron-sulphur world. *Prog Biophys Mol Biol* 1992;**58**:85–201. https://doi.org/10.1016/0079-6107(92)90022-X.
Wald G. The origins of life. *Proc Nat Acad Sci USA* 1964;**52**:595–611. https://doi.org/10.1073/pnas.52.2.595.
Walker JE. ATP synthesis by rotary catalysis (Nobel lecture). *Angew Chem Engl Ed* 1998;**17**:2308–19. https://doi.org/10.1002/(SICI)1521-3773(19980918)37:17%3c2308::AID-ANIE2308%3e3.0.CO;2-W.
Wang L, Liu Q, Wu X *et al.* Bioinformatics analysis of metabolism pathways of archaeal energy reserves. *Sci Rep* 2019;**9**:1034. https://doi.org/10.1038/s41598-018-37768-0.
Weiss M, Sousa FL, Mrnjavac N *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.
White RH. L-Aspartate semialdehyde and a 6-deoxy-5-ketohexose 1-phosphate are the precursors to the aromatic amino acids in *Methanocaldococcus jannaschii*. *Biochemistry* 2004;**43**:7618–27. https://doi.org/10.1021/bi0495127.
Wilkinson JF. The problem of energy-storage compounds in bacteria. *Exp Cell Res* 1959;**7**:111–30. https://doi.org/10.1016/0014-4827(59)90237-X.
Williams TA, Szöllősi GJ, Spang A *et al.* Integrative modeling of gene and genome evolution roots the archaeal tree of life. *Proc Natl Acad Sci USA* 2017;**114**:E4602–11. https://doi.org/10.1073/pnas.1618463114.
Wimmer JLE, do Nascimento Vieira A, Xavier JC *et al.* The autotrophic core: an ancient network of 404 reactions converts H_2, CO_2, and NH_3 into amino acids, bases, and cofactors. *Microorganisms* 2021;**9**:458. https://doi.org/10.3390/microorganisms9020458.
Yakimov MM, Merkel AY, Gaisin VA *et al.* Cultivation of a vampire: 'Candidatus *Absconditicoccus praedator*'. *Environ Microbiol* 2022;**24**:30–49. https://doi.org/10.1111/1462-2920.15823.
Yonath A. Polar bears, antibiotics, and the evolving ribosome (Nobel Lecture). *Angew Chem* 2010;**49**:4341–54.
Yu F, Fei J, Jia Y *et al.* Chemiosmotic ATP synthesis by minimal protocells. *Cell Rep Phys Sci* 2025;**6**:102461. https://doi.org/10.1016/j.xcrp.2025.102461.
Yu JP, Ladapo J, Whitman WB. Pathway of glycogen metabolism in *Methanococcus maripaludis*. *J Bacteriol* 1994;**176**:325–32. https://doi.org/10.1128/jb.176.2.325-332.1994.
Zaparty M, Tjaden B, Hensel R *et al.* The central carbohydrate metabolism of the hyperthermophilic crenarchaeote *Thermoproteus tenax*: pathways and insights into their regulation. *Arch Microbiol* 2008;**190**:231–45. https://doi.org/10.1007/s00203-008-0375-5.
Zhang J, Feng X, Li M *et al.* Deep origin of eukaryotes outside Heimdallarchaeia within Asgardarchaeota. *Nature* 2025;**642**:990–8. https://doi.org/10.1038/s41586-025-08955-7.
Zhang Y, Kouril T, Snoep JL *et al.* The peculiar glycolytic pathway in hyperthermophilic archaea: understanding its whims by experimentation in silico. *Int J Mol Sci* 2017;**18**:876. https://doi.org/10.3390/ijms18040876.
Zhang Y, Sampathkumar A, Kerber SML *et al.* A moonlighting role for enzymes of glycolysis in the co-localization of mitochondria and chloroplasts. *Nat Commun* 2020;**11**:4509. https://doi.org/10.1038/s41467-020-18234-w.

Received 11 July 2025; **revised** 17 September 2025; **accepted** 25 September 2025

6 Zusammenfassung der Ergebnisse

Für die Synthese von Corrinen, die neben anderen Funktionen eine wichtige Rolle im Acetyl-CoA-Weg zur Fixierung von Kohlenstoffdioxid als Cofaktor des Corrinoid-Eisen/Schwefel-Proteins (CoFeS) einnehmen, existieren anaerobe und aerobe Routen in Prokaryoten, die sich im Wesentlichen in ihrer Abhängigkeit von O_2 unterscheiden (Bryan *et al*., 2020). Zudem werden sie durch den Zeitpunkt der Cobalt-Insertion charakterisiert, welche entweder früh (anaerob) oder spät (aerob) in der Synthese erfolgt. Im Kontext der frühen Evolution ist in erster Linie der anaerobe Weg von Bedeutung, da Sauerstoff das Ergebnis cyanobakterieller Photosynthese ist und kein Kennzeichen der geo- und biochemischen Umgebung darstellt, in welcher frühe prokaryotische Gruppen sowie der Acetyl-CoA-Weg ihren Ursprung finden (Martin und Russell, 2007; Martin und Sousa, 2015; Slotznick *et al*., 2022). So stellen O_2-abhängige Varianten biosynthetischer Enzyme vermutlich spätere Innovationen dar, die auf anaeroben Stoffwechselwegen basieren und der Adaption an sauerstoffreiche Umgebungen dienten (Mrnjavac *et al*., 2024b).

Da einige Erkenntnisse darauf schließen lassen, dass methanogene Archaeen und acetogene Bakterien die wohl ältesten prokaryotischen Gruppen bilden (Decker *et al*., 1970; Weiss *et al*., 2016) und diese den Acetyl-CoA-Weg mit dem damit verbundenen CoFeS besitzen (Fuchs 2011), fokussierten sich die in **Publikation I** dargestellten Analysen auf Gene der anaeroben Cobalaminsynthese sowie der unteren Liganden 5-Hydroxybenzimidazol und 5,6-Dimethylbenzimidazol. Ein Großteil der untersuchten Enzyme stammen aus *Methanococcus maripaludis* und *Moorella thermoacetica*, zwei Modellorganismen, welche Methanogene und Acetogene repräsentieren.

Strukturelle Alignments der herangezogenen Proteinsequenzen lassen auf eine Homologie der anaeroben Cobalaminsynthese in diesen anzestralen prokaryotischen Gruppen schließen, da sie in nahezu allen Fällen, mit Ausnahme der Adenosylcobinamidkinasen CobU und CobY, starke Übereinstimmungen zeigen. Zusätzlich weisen durch AlphaFold (Jumper *et al*., 2021) berechnete Proteinstrukturen des im Acetyl-CoA-Weg aktiven CoFeS, der Acetyl-CoA-Synthase und der Kohlenmonoxid-Dehydrogenase hohe Übereinstimmung zwischen den acetogenen und methanogenen Varianten auf (Adam *et al*., 2018). Sequenzvergleiche der CoFeS Untereinheiten Delta und Gamma lassen ebenfalls auf Homologie schließen, sie resultieren in globalen Identitäten von 31,4 % bzw. 37,8 %. Im Gegensatz dazu zeigen

strukturelle Vergleiche der archaeellen und bakteriellen Methylsynthesewege des Acetyl-CoA-Wegs einige nicht-homologe Abschnitte, was unabhängige Ursprünge dieser beiden Routen in Prokaryoten unterstützt (Sousa und Martin, 2014).

Homologiesuchen der hier zugrundeliegenden Sequenzen in einem Datensatz methanogener und acetogener Referenzgenome ergeben, dass die meisten Enzyme, die an der anaeroben Synthese von Cobalamin fungieren, über die acetogene-methanogene Grenze hinaus konserviert sind. Eine Vielzahl methanogener Sequenzen weist eine lokale Identität von mindestens 25 % in acetogenen Spezies auf, umgekehrt ist das gleiche Muster zu beobachten. Dennoch existieren, wie auch bereits in der Strukturanalyse beschrieben, einige Ausnahmen. So bestätigen Sequenzalignments des bakteriellen CobU und des archaeellen CobY den zuvor hervorgehobenen strukturellen Vergleich, sie weisen keine Homologie auf (Rodionov *et al*., 2003). Eine weitere Auffälligkeit ist die Abwesenheit der Adenosyltransferase CobA in methanogenen Organismen, sie koordiniert die Verknüpfung einer Adenosylgruppe mit dem zentralen Co-Atom von Cobalamin (Vitols *et al*., 1966). Die Beobachtung, dass viele methanogene Archaeen nicht über Adenosylcobalamin verfügen (Stupperich und Kräutler, 1988), wird durch dieses Resultat bestätigt. Die Enzyme der Ligandensynthese sind weniger gleichmäßig im Datensatz verteilt, was die Varianz an unteren Liganden in autotrophen Prokaryoten widerspiegelt (Stupperich und Kräutler, 1988; Stupperich *et al*., 1988). BzaF, welches die Produktion des unteren Liganden 5-Hydroxybenzimidazol katalysiert (Hazra *et al*., 2015), weist ein universelles Vorkommen in den analysierten Genomen auf.

Die Ergebnisse einer weiteren sequenzbasierten Analyse heben hervor, dass zwölf Enzyme der anaeroben Cobalaminsynthese Homologe in anderen biosynthetischen Wegen aufweisen, die auf LUCA zurückzuführen sind, darunter die Synthese von Cofaktoren und Aminosäuren (Weiss *et al*., 2016). Diese Beobachtung lässt vermuten, dass einige Enzyme der Cobalaminsynthese aus bereits vorhandenen anzestralen, weniger komplexen Synthesewegen übernommen wurden.

Einen weiteren, zentralen Aspekt von **Publikation I** bildet die phylogenetische Rekonstruktion der zuvor erwähnten Enzyme, wobei Sequenzalignments der zuvor detektierten Homologe als Basis der Berechnungen dienen. In dieser Analyse kristallisieren sich drei verschiedene Klassen von ungewurzelten Genbäumen heraus, die zusammengenommen eine nahezu vollständige Präsenz der anaeroben Cobalaminsynthese im letzten gemeinsamen Vorfahren allen Lebens, LUCA, unterstützen.

Die erste Klasse bilden Phylogenien, die eine archaeelle und bakterielle Monophylie aufweisen und dadurch die entsprechenden Enzyme auf den gemeinsamen Vorfahren der

Archaeen und Bakterien zurückführen. Insgesamt wurden neun Enzyme dieser Kategorie zugeordnet, davon sechs Enzyme der Cobalaminsynthese sowie drei Enzyme, die an der Synthese des unteren Liganden 5-Hydroxybenzimidazol beteiligt sind. Zwei dieser Enzyme finden ebenfalls in der Synthese von Sirohäm Verwendung, ein Cofaktor, der bereits in früheren Berichten LUCA zugeschrieben wurde (Weiss *et al*., 2016).

Die zweite Klasse umfasst phylogenetische Bäume, in denen Archaeen und Bakterien erneut nahezu in Monophylie auftreten, diese allerdings durch wenige laterale Gentransfers (LGT) aufgebrochen wird. So ordnen sich einige methanogene und acetogene Sequenzen in sieben Genbäumen innerhalb der jeweils anderen prokaryotischen Gruppe ein. Zusammen mit den Phylogenien der ersten Klasse deuten diese trotz weniger LGTs, wovon im Durchschnitt in etwa 42,5 % der Gene prokaryotischer Spezies betroffen sind (Dmitrijeva *et al*., 2024), auf einen Ursprung in LUCA hin. Eine Untersuchung der von LGT betroffen Sequenzen mithilfe eines für diesen genomischen Datensatz generierten prokaryotischen Stammbaums lässt vermuten, dass womöglich vier Enzyme der anaeroben Cobalaminsynthese in einem einzigen Transferereignis von acetogenen Bakterien auf den Vorfahren einer Gruppe methanogener Spezies übertragen wurde.

Die dritte und letzte Klasse beinhaltet Phylogenien, die eine hohe Durchmischung acetogener und methanogener Sequenzen aufweist, neun Genbäume werden dieser Kategorie zugeordnet. Die strukturelle und funktionelle Homologie dieser Enzyme lassen darauf schließen, dass (i) die Funktionen dieser Enzyme bereits in den anzestralen Versionen der Gene verankert waren und dass (ii) Mitglieder dieser funktionellen Enzymfamilien bereits im gemeinsamen Vorfahren von Archaeen und Bakterien vorhanden waren.

Zusammenfassend unterstützen die in **Publikation** I beschriebenen Analysen einen Ursprung der anaeroben Cobalaminsynthese in dem letzten gemeinsamen Vorfahren allen Lebens, LUCA, unter der Berücksichtigung der wichtigen Rolle, die Cobamide im Acetyl-CoA-Weg zur CO_2-Fixierung einnehmen, obwohl die Komplexität der Synthese von Cobalamin (Hazra *et al*., 2015; Bryant *et al*., 2020) auf den ersten Blick nicht in das Gesamtbild der frühen, autotrophen Evolution passt. Das zwischen Acetogenen und Methanogenen homologe CoFeS und die damit verbundene Funktion im anzestralen Acetyl-CoA-Weg (Fuchs und Stupperich, 1985; Fuchs 2011), dessen Enzyme durch metallische Katalysatoren ersetzt werden können (Preiner *et al*., 2020; Beyazay *et al*., 2023), spricht für eine frühe evolutionäre Einordnung der Cobalaminsynthese in eine Übergangsphase zwischen Geochemie und Biochemie. Die in dieser Arbeit bestätigte konvergente Evolution von Methylsynthesewegen in Archaeen und Bakterien (Sousa und Martin, 2014) verstärkt die Annahme, dass LUCA und

CoFeS Methylgruppen aus der direkten geochemischen Umgebung mobilisierten (Weiss *et al*., 2016). Diese Publikation verbindet somit die anaerobe Synthese von Cobalamin mit Phasen früher biochemischer Evolution, in der Reaktionen von Enzymen, Cofaktoren und Übergangsmetallen angetrieben wurden (Mrnjavac *et al*., 2024a).

Publikation II beschäftigt sich mit der evolutionären Einordnung glykolytischer Stoffwechselwege, die in zahlreichen Varianten existieren. So dienen der Embden-Meyerhof (EM) Weg, der Entner-Doudoroff Weg sowie der Pentosephosphatweg dem Abbau von Glucose und bilden zusammen mit dem Metabolismus von Glykogen zentrale Aspekte prokaryotischen Stoffwechsels (Bräsen *et al*., 2014; Cifuente *et al*., 2024). Der als klassische Glykolyse bezeichnete EM Weg wandelt ein Molekül Glucose in zwei Moleküle Pyruvat mit einem Nettoenergiegewinn von zwei ATP um, wobei diese Route in Vorbereitungsphase (Glucose → 2 Triosephosphate) und die evolutionär deutlich stärker konservierte (Fothergill-Gilmore und Michels, 1993) Energiegewinnungsphase (Triosephosphate → Pyruvat) unterteilt wird. Interessanterweise besitzen einige autotrophe, H_2-abhängige methanogene Archaeen glykolytische Enzyme, obwohl sie nicht über die Eigenschaft verfügen, extrazelluläre Glucose als Wachstumssubstrat zu nutzen (Yu *et al*., 1994; Siebers und Schönheit, 2005; Sattler *et al*., 2024). Bis auf wenige Ausnahmen sind die Enzyme des EM Wegs neben ihrer katabolen Aktivität in der Lage, die Reaktionen reversibel, anabol zu katalysieren (Bräsen *et al*., 2014), sodass in der Gluconeogenese Pyruvat in Glucose umgewandelt wird. Diese kann durch eine Reihe weiterer Synthese- und Abbauwege zu Glykogen verkettet werden, eine Kohlenstoff- und Energiereserve, die in allen drei Domänen des Lebens zu finden ist (Wilkinson *et al*., 1959; Wang *et al*., 2019).

Publikation II behandelt ferner die Frage, ob die Glykolyse einen heterotrophen Ursprung besitzt, wie einst von Christian de Duve (1991), George Wald (1964) und weiteren Wissenschaftlern vorgeschlagen (Keller *et al*., 2014), oder ob sie sich ursprünglich autotroph, gluconeogenetisch entwickelte (Ronimus und Morgan, 2003). Dazu werden die Verteilung und Phylogenien prokaryotischer Enzyme des klassischen und, hauptsächlich von Archaeen genutzten, modifizierten EM Wegs sowie fünf Enzyme des klassischen Glykogenmetabolismus in einem taxonomisch ausbalancierten Datensatz archaeeller und bakterieller Referenzgenome bioinformatisch untersucht. Die Identifizierung homologer Sequenzen der genannten Enzyme in besagten Genomen dient der phylogenetischen Analyse als Grundlage. Insbesondere die Enzyme der Energiegewinnungsphase (Triosephosphatisomerase, Glycerinaldehyd-3-phosphat-Dehydrogenase, Phosphoglyceratkinase, Cofaktor-unabhängige Phosphoglyceratmutase, Enolase und Phosphoenolpyruvatsynthase) sind nahezu universell

vertreten. Zudem weisen Fructose-1,6-bisphosphat-Aldolase, Phosphoglucoseisomerase und Phosphoglucomutase eine weite Verbreitung in den analysierten Prokaryoten auf. Das universelle Vorkommen dieser in der Gluconeogenese aktiven Enzyme lässt vermuten, dass sie eine anzestrale Funktion in der frühen Evolutionsgeschichte der Prokaryoten ausübten. Zwei dieser Enzyme (Cupin-Typ Phosphoglucoseisomerase und Triosephosphatisomerase) wurden in früheren Berichten bereits mit LUCA in Verbindung gebracht (Weiss *et al*., 2016).

Darüber hinaus sind einige Enzyme des Glykogenstoffwechsels in einem Großteil der untersuchten Genome vorhanden, vor allem Glykogen-Synthase und NDP-Glucose-Pyrophosphorylase, welche die ersten beiden Schritte in der klassischen Umwandlung von Glucose in Glykogen katalysieren (Preiss *et al*., 2014). Das Verzweigungsenzym GlgB, welches Glykogen weiter verdichtet und ebenfalls in einer anderen Route der Glykogensynthese beteiligt ist (Chandra *et al*., 2011), weist eine weite Verbreitung in der bakteriellen Fraktion auf. Die Abwesenheit dieses Enzyms in den meisten Archaeen lässt vermuten, dass (i) in diesen Spezies Glucosemoleküle entweder zu simpleren Formen verkettet werden oder dass (ii) eine bislang nicht entdeckte archaeelle Alternative dieses Enzyms existiert (Suzuki und Suzuki, 2016). Homologe von Enzymen, die hauptsächlich glykolytische Funktionen innehaben, sind, insbesondere in methanogenen Archaeen, weniger weit verbreitet.

Phylogenetische Rekonstruktionen der hier vorliegenden Enzymsequenzen ermöglichen einen zusätzlichen Einblick in deren evolutionären Historie. Es stellt sich heraus, dass archaeelle und bakterielle Sequenzen der Enolase, Phosphoglyceratkinase und Triosephosphatisomerase eine klare Trennung in den Genbäumen aufweisen. Ein ähnliches Bild zeigt die Cofaktor-unabhängige Phosphoglyceratmutase, in deren Phylogenie Archaeen und Bakterien, bis auf wenige LGTs, zwei separate Gruppen bilden. Diese vier Phylogenien lassen darauf schließen, dass die Energiegewinnungsphase der Glykolyse anzestrale Enzyme umfasst, die ihren Ursprung in den frühesten Phasen biochemischer Evolution finden.

Neben dem starken Einfluss von lateralem Gentransfer zwischen Prokaryoten deuten die restlichen Phylogenien die Entwicklung domänenspezifischer Enzyme an, die im Laufe der Evolution vereinzelt horizontal auf Spezies der anderen Domäne übertragen wurden. Beispielhaft anzuführen sind dabei die archaeelle, bifunktionelle Fructose-1,6-bisphosphat-Aldolase/-Phosphatase (Say und Fuchs, 2010) oder die bakterielle Cofaktor-abhängige Phosphoglyceratmutase. Den archaeellen Enzymen reiht sich unter anderem eine anzestrale Form der Glycerinaldehyd-3-phosphat Dehydrogenase (GAPDH) ein, die ausschließlich gluconeogenetisch agiert und phylogenetisch eindeutig von der bakteriellen, bidirektionalen GAPDH zu trennen ist (Tästensen *et al*., 2018). Die dadurch auftretende Lücke in der

archaeellen Glykolyse wird durch die katabolen nicht-phosphorylierende GAPDH, GAPN, (Lorentzen *et al*., 2004; Reher und Schönheit, 2006) und Ferredoxin-abhängige Glycerinaldehyd-3-phosphat-Oxidoreduktase GAPOR (Adams 1994) geschlossen.

Ein abschließendes, äußerst nennenswertes Resultat der in dieser Publikation vorliegenden phylogenetischen Analyse ist die Position halophiler Archaeen in zahlreichen Genbäumen. Haloarchaeen sind aerobe, heterotrophe von autotrophen Methanogenen abstammende Organismen, deren Evolution stark durch lateralen Gentransfer von Bakterien beeinflusst wurde (Kelly *et al*., 2011; Nelson-Sathi *et al*., 2012; Mei *et al*., 2023). Phylogenetische Rekonstruktionen in dieser Arbeit ergeben, dass halophile Archaeen den Großteil ihres glykolytischen Stoffwechsels horizontal von Bakterien erhalten haben. In insgesamt acht Phylogenien ordnen sich die haloarchaeellen Sequenzen innerhalb bakterieller Gruppen ein, was auf einen bakteriellen Ursprung der betroffenen Enzyme schließen lässt. Keine andere prokaryotische Gruppe fällt in dieser Analyse in jenes Muster, somit stellt diese Beobachtung ein einzigartiges Merkmal der Haloarchaeen dar. Die Resultate unterstützen die Erkenntnisse von Nelson-Sathi *et al*. (2012), dass zahlreiche LGTs von Bakterien anaerobe, methanogene Archaeen in heterotrophe, aerobe Organismen umgewandelt haben, die unter anderem eine bakterielle Glykolyse sowie weitere bakterielle Enzyme des zentralen Kohlenstoffmetabolismus besitzen (Pickl *et al*., 2012; Reinhardt *et al*., 2019; Tästensen *et al*., 2020).

Die in **Publikation II** beschriebenen Ergebnisse lassen zusammenfassend darauf deuten, dass die Glykolyse einen autotrophen, gluconeogenetischen Ursprung aufweist. Eine universelle Verteilung von Enzymen der Energiegewinnungsphase sowie die nahezu vollständige archaeelle und bakterielle Monophylie in den damit verbundenen Phylogenien unterstützen, dass diese hoch konserviert sind und ursprünglich der Umwandlung von Pyruvat, ein Produkt des anzestralen Acetyl-CoA-Wegs, in Triosephosphate in einem autotrophen Szenario dienten. Dieser Vorgang stellte C3-Verbindungen für die Synthese von Aminosäuren, Basen und Cofaktoren zur Verfügung (Wimmer *et al*., 2021). In dieser Annahme entwickelten sich der Embden-Meyerhof Weg sowie weitere glykolytische Routen zum Abbau von Glucose und zur Mobilisierung von anzestralen intrazellulären Kohlenstoff- und Energiereserven in Form von Glykogen, welches bis zum letzten gemeinsamen Vorfahren der Archaeen und Bakterien zurückreicht. Dieses fungierte möglicherweise als eins der ersten heterotrophen Wachstumssubstrate in der Evolution des zentralen prokaryotischen Kohlenstoffmetabolismus. Die weite Verbreitung von Glykogen-Synthasen sowie glykolytischer Enzyme in methanogenen Archaeen, die nicht auf extrazellulär Glucose zurückgreifen können, unterstützt

die Theorie, dass diese Organismen reversible glykolytische Enzyme ursprünglich für den Aufbau von Kohlenstoffreserven nutzten (König *et al*., 1985). Ein Abbau dieser Reserven stellte möglicherweise Ausgangssubstrate für die Methanogenese zur Verfügung (Santiago-Martinez *et al*., 2016; Richter *et al*., 2024).

Die Resultate dieser Publikation stimmen mit früheren Theorien zum autotrophen, gluconeogenetischen Ursprung überein (Fuchs 1986; Romano und Conway, 1996; Ronimus und Morgan, 2003; Say und Fuchs, 2010), erweitern diese mit der potenziellen Rolle, die Glykogen im Kontext der frühen Evolution eingenommen haben könnte und bieten ein ganzheitliches Bild zur phylogenetischen Einordnung klassischer prokaryotischer Enzyme der Glykolyse, Gluconeogenese und des Glykogenstoffwechsels.

7 Literaturverzeichnis

Abbanat, D. R. und Ferry, J. G. (1991). Resolution of component proteins in an enzyme complex from *Methanosarcina thermophila* catalyzing the synthesis or cleavage of acetyl-CoA. *Proc. Natl. Acad. Sci. U.S.A.* 88, 8: 3272-6.

Adam, P. S. *et al*. (2018). Evolutionary history of carbon monoxide dehydrogenase/acetyl-CoA synthase, one of the oldest enzymatic complexes. *Proc. Natl. Acad. Sci. U.S.A.* 115, 6: E1166-E1173.

Adams, M. W. W. (1994). Biochemical diversity among sulfur-dependent, hyperthermophilic microorganisms. *FEMS Microbiol. Rev*. 15: 261-267.

Albers, S. V. und Meyer, B. H. (2011). The archaeal cell envelope. *Nat. Rev. Microbiol*. 9: 414-426.

Allen, J. F. *et al*. (2019). Nitrogenase inhibition limited oxygenation of Earth's proterozoic atmosphere. *Trends Plant Sci.* 24, 11: 1022-103.

Arnold, B. J. *et al*. (2022). Horizontal gene transfer and adaptive evolution in bacteria. *Nat. Rev. Microbiol.* 20, 4: 206-218.

Banerjee, R. V. und Matthews, R. G. (1990). Cobalamin-dependent methionine synthase. *FASEB J.* 4, 5: 1450-9.

Banerjee, R. und Ragsdale, S. W. (2003). The many faces of vitamin B12: catalysis by cobalamin-dependent enzymes. *Ann. Rev. Biochem.* 72: 209-47.

Belthle, K. S. und Tüysüz, H. (2023). Linking catalysis in biochemical and geochemical CO_2 fixation at the emergence of life. *ChemCatChem* 15, e202201462.

Beyazay, T. *et al*. (2023). Influence of composition of Nickel-Iron nanoparticles for abiotic CO_2 conversion to early prebiotic organics. *Angew. Chem. (International ed. in English)* 62, 22: e202218189.

Bräsen, C. *et al*. (2014). Carbohydrate metabolism in Archaea: current insights into unusual enzymes and pathways and their regulation. *MMBR* 78, 1: 89-175.

Brindley, A. A. *et al*. (2003). A story of chelatase evolution: identification and characterization of a small 13-15-kDa "ancestral" cobaltochelatase (CbiXS) in the archaea. *J. Biol. Chem.* 278, 25: 22388-95.

Broderick, J. B. *et al*. (2014). Radical S-adenosylmethionine enzymes. *Chem. Rev.* 114, 8: 4229-317.

Bryant, D. A. *et al*. (2020). Biosynthesis of the modified tetrapyrroles-the pigments of life. *J. Biol. Chem.* 295, 20: 6888-6925.

Budgen, N. und Danson, M. J. (1986). Metabolism of glucose via a modified Entner-Doudoroff pathway in the thermoacidophilic archaebacterium, Thermoplasma acidophilum. *FEBS Lett.* 196: 207–210.

Cech, T. R. (1986). A model for the RNA-catalyzed replication of RNA. *Proc. Natl Acad. Sci. USA* 83, 4360–4363.

Chandra, G. *et al*. (2011). Unexpected and widespread connections between bacterial glycogen and trehalose metabolism. *Microbiology*. 157: 1565–1572.

Cheong, C. G. *et al*. (2001). Structural investigation of the biosynthesis of alternative lower ligands for cobamides by nicotinate mononucleotide: 5,6-dimethylbenzimidazole phosphoribosyltransferase from Salmonella enterica. *J. Biol. Chem.* 276, 40: 37612-20.

Chistoserdova, L. *et al*. (1998). C1 transfer enzymes and coenzymes linking methylotrophic bacteria and methanogenic Archaea. *Science (New York, N.Y.)* 281, 5373: 99-102.

Cifuente, J. O. *et al*. (2019). Structural basis of glycogen metabolism in bacteria. *Biochem. J.* 476, 14: 2059-2092.

Cifuente, J. O. *et al*. (2024). Architecture, function, regulation, and evolution of α-glucans metabolic enzymes in prokaryotes. *Chem. Rev.* 124, 8: 4863-4934.

Colman, D. R. *et al*. (2022). Deep-branching acetogens in serpentinized subsurface fluids of Oman. *Proc. Natl. Acad. Sci. U.S.A.* 119, 42: e2206845119.

Crofts, T. S. *et al*. (2013). Cobamide structure depends on both lower ligand availability and CobT substrate specificity. *Chem. Biol.* 20, 10: 1265-74.

Darwin, C. (1860). Über die Entstehung der Arten im Thier- und Pflanzen-Reich durch natürliche Züchtung. Stuttgart: Schweizerbart.

De Duve, C. (1991). Blueprint for a cell: the nature and origin of life, *Neil Patterson Publishers* Burlington, North Carolina.

Debussche, L. *et al*. (1992). Assay, purification, and characterization of cobaltochelatase, a unique complex enzyme catalyzing cobalt insertion in hydrogenobyrinic acid a,c-diamide during coenzyme B12 biosynthesis in Pseudomonas denitrificans. *J Bacteriol.* 174, 22: 7445-51.

Decker, K. *et al*. (1970). Energy production in anaerobic organisms. *Angew. Chem. (International ed. in English)* 9, 2: 138-58.

Dmitrijeva, M. *et al*. (2024). A global survey of prokaryotic genomes reveals the eco-evolutionary pressures driving horizontal gene transfer. *Nat. Ecol. Evol.* 8, 5: 986-998.

Doolittle, W. F. (1999). Phylogenetic classification and the universal tree. *Science (New York, N.Y.)* 284, 5423: 2124-9.

Dubey, G. P. und Ben-Yehuda, S. (2011). Intercellular nanotubes mediate bacterial communication. *Cell.* 144, 4: 590-600.

Eigen, M. (1971). Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* 58: 465-523.

Eschenmoser, A. (1988). Vitamin B_{12}: experiments concerning the origin of its molecular structure. *Angew. Chem. Int. Ed.* 27: 5-39.

Fang, H. *et al*. (2017). Microbial production of vitamin B_{12}: a review and future perspectives. *Microb. Cell Fact.* 16, 1: 15.

Fothergill-Gilmore, L. A. und Michels, P. A. M. (1993). Evolution of glycolysis. *Prog. Biophys. Mol. Biol*. 59 :105–235.

Frank, S. *et al*. (2005). Anaerobic synthesis of vitamin B12: characterization of the early steps in the pathway. *Biochem. Soc. Trans.* 33, 4: 811-4.

Fuchs, G. (1986). CO_2 fixation in acetogenic bacteria: Variations on a theme. *FEMS Microbiol. Rev.* 2, 3: 181–213.

Fuchs, G. (2011). Alternative pathways of carbon dioxide fixation: insights into the early evolution of life? *Ann. Rev. Microbiol.* 65: 631-58.

Fuchs G. und Stupperich E. (1985). Evolution of autotrophic CO_2 fixation. *Evolution of Prokaryotes*, eds Schleifer K. H., Stackebrandt E. (London: Academic Press;): 235–225.

Furdui, C. und Ragsdale, S. W. (2000). The role of pyruvate ferredoxin oxidoreductase in pyruvate synthesis during autotrophic growth by the Wood-Ljungdahl pathway. *J. Biol. Chem.* 275, 37: 28494-9.

Gilbert, W. (1986). The RNA world. *Nature* 319: 618.

Gonzalez-Ordenes, F. *et al*. (2024). Glycogen metabolism in methanogens: a key pathway for metabolic response to nutrient availability. *Biochem. Biophys. Res. Commun.* 739: 150978.

Gough, S. P. *et al*. (2000). Anaerobic chlorophyll isocyclic ring formation in *Rhodobacter capsulatus* requires a cobalamin cofactor. *Proc. Natl. Acad. Sci. U.S.A.* 97, 12: 6908-13.

Gruber, K. *et al*. (2011). Vitamin B12-derivatives-enzyme cofactors and ligands of proteins and nucleic acids. *Chem. Soc. Rev.* 40, 8: 4346-63.

Haeckel, E. (1866b). Generelle Morphologie der Organismen Zweiter Band: Allgemeine Entwicklungsgeschichte der Organismen. *Berlin: Verlag von Georg Reimer*.

Haldane, J. B. S. (1929). The origin of life. *The Rationalist Annual* 148: 3–10.

Hansen, T. *et al*. (2005). Cupin-type phosphoglucose isomerases (Cupin-PGIs) constitute a novel metal-dependent PGI family representing a convergent line of PGI evolution. *J Bacteriol*. 187, 5: 1621-1631.

Hazra, A. B. *et al*. (2015). Anaerobic biosynthesis of the lower ligand of vitamin B12. *Proc. Natl. Acad. Sci. U.S.A.* 112, 34: 10792-7.

Henriques Pereira, D. P. *et al*. (2022). Role of geochemical protoenzymes (geozymes) in primordial metabolism: specific abiotic hydride transfer by metals to the biological redox cofactor NAD. *FEBS J.* 289, 11: 3148-3162.

Hensel, R. *et al*. (1989). Sequence comparison of glyceraldehyde-3-phosphate dehydrogenases from the three urkingdoms: evolutionary implication. *Can. J. Microbiol.* 35, 1: 81-5.

Huang, D. D. *et al*. (1984). delta-Aminolevulinic acid-synthesizing enzymes need an RNA moiety for activity. *Science (New York, N.Y.)* 225, 4669: 1482-4.

Iddar, A. *et al*. (2005). Widespread occurrence of non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase among gram-positive bacteria. *Int. Microbiol*. 8: 251–258.

Jones, D. und Sneath, P. H. (1970). Genetic transfer and bacterial taxonomy. *Bacteriol. Rev.* 34, 1: 40-81.

Jumper, J. *et al*. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 7873: 583-589.

Kaur, H. *et al*. (2024). A prebiotic Krebs cycle analog generates amino acids with H_2 and NH_3 over nickel. *Chem* 10, 5: 1528-1540.

Keller, M. A. *et al*. (2014). Non-enzymatic glycolysis and pentose phosphate pathway-like reactions in a plausible Archean ocean. *Mol. Sys. Biol.* 10, 4 725. 25.

Keller, A. *et al*. (2019). Energy-conserving enzyme systems active during syntrophic acetate oxidation in the thermophilic bacterium *Thermacetogenium phaeum*. *Front. Microbiol.* 10: 2785. 29.

Kelley, D. S. *et al*. (2001). An off-axis hydrothermal vent field near the Mid-Atlantic Ridge at 30 degrees N. *Nature* 412, 6843: 145-9.

Kelley, D. S. *et al*. (2005). A serpentinite-hosted ecosystem: the Lost City hydrothermal field. *Science (New York, N.Y.)* 307, 5714: 1428-34.

Kelly S. *et al*. (2011). Archaeal phylogenomics provides evidence in support of a methanogenic origin of the Archaea and a thaumarchaeal origin for the eukaryotes. *Proc. Biol. Sci.* 278, 1708: 1009-18.

Kolhouse, J. F. und Allen, R. H. (1977). Recognition of two intracellular cobalamin binding proteins and their identification as methylmalonyl-CoA mutase and methionine synthetase. *Proc. Natl. Acad. Sci. U.S.A.* 74, 3: 921-5.

König, H. *et al*. (1982). Glycogen in thermoacidophilic Acaebacteria of the genera *Sulfolobus*, *Thermoproteus*, *Desulfurococcus* and *Thermococcus*. *Arch. Microbiol*. 132: 297–303.

König, H. *et al*. (1985). Glycogen in *Methanolobus* and *Methanococcus*. *FEMS Microbiol. Lett.* 28, 3: 265-269.

Kuprat, T. *et al*. (2021). Glucose metabolism and acetate switch in Archaea: the enzymes in *Haloferax volcanii*. *J. Bacteriol.* 203, 8: e00690-20. 23.

Lee, A. *et al*. (2021). Identification of the genes related to the glycogen metabolism in hyperthermophilic archaeon, *Sulfolobus acidocaldarius*. *Front. Microbiol.* 12: 661053.

Lipmann, F. (1941). Metabolic generation and utilization of phosphate bond energy. In *Adv. Enzymol. Relat. Areas Mol. Biol.* (eds F.F. Nord and C.H. Werkman).

Lorentzen, E. *et al*. (2004). Structural basis of allosteric regulation and substrate specificity of the non-phosphorylating glyceraldehyde 3-6 phosphate dehydrogenase from *Thermoproteus tenax*. *J. Mol. Biol*. 13; 341, 3: 815-28.

Maitra, P. K. *et al*. (2001). Metabolite and enzyme profiles of glycogen metabolism in Methanococcoides methylutens. *FEMS Microbiol. Lett.* 198, 1: 23-9.

Martin, W. F. und 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. *Phil. Trans. R. Soc. Lond. B, Biol. Sci.* 358, 1429: 59-83; discussion 83-5.

Martin, W. F. und Russell, M. J. (2007). On the origin of biochemistry at an alkaline hydrothermal vent. *Phil. Trans. R. Soc. Lond. B, Biol. Sci.* 362, 1486: 1887-925.

Martin, W. F. *et al.* (2008). Hydrothermal vents and the origin of life. *Nat. Rev. Microbiol.* 6: 805–814.

Martin, W. F. und Sousa, F. L. (2015). Early microbial evolution: the age of anaerobes. *Cold Spring Harb. Perspect. Biol.* 8, 2: a018127.

Martin, W. F. (2020). Older than genes: the acetyl CoA pathway and origins. *Front. Microbiol.* 11: 817.

Matthews, R. G. (2009). Cobalamin- and corrinoid-dependent enzymes. *Met. Ions Life Sci.* 6: 53-114.

Mei, R. *et al*. (2023). The origin and evolution of methanogenesis and Archaea are intertwined. *PNAS nexus* 2, 2: pgad023.

Mereschkowsky, C. (1905). Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biol. Zentralbl*. 25: 593–604.

Mereschkowsky C. (1910). Theorie der zwei Plasmaarten als Grundlage der Symbiogenesis, einer neuen Lehre von der Entstehung der Organismen. *Biol. Centralbl.* 30: 278–288, 289,–303, 353–367.

Miller, S. L. (1953). A production of amino acids under possible primitive Earth conditions. *Science* 117: 528–9.

Modjewski, L. D. *et al.* (2024). Evidence for corrin biosynthesis in the last universal common ancestor. *FEBS J.* 292, 4: 827-850.

Morawska, L. P. und Kuipers, O. P. (2023). Cell-to-cell non-conjugative plasmid transfer between Bacillus subtilis and lactic acid bacteria. *Microb. Biotechnol.* 16, 4: 784-798.

Mrnjavac, N. *et al*. (2024a). Chemical antiquity in metabolism. *Acc. Chem. Res.* 57, 16: 2267-2278

Mrnjavac, N. *et al*. (2024b). The radical impact of oxygen on prokaryotic evolution-enzyme inhibition first, uninhibited essential biosyntheses second, aerobic respiration third. *FEBS Lett.* 598, 14: 1692-1714.

Muchowska, K. B. *et al*. (2017). Metals promote sequences of the reverse Krebs cycle. *Nat. Ecol. Evol.* 1, 11: 1716-1721.

Murakami, T. *et al*. (2006). A novel branching enzyme of the GH-57 family in the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. *J. Bacteriol.* 188, 16: 5915-24.

Murray, P. A. und Zinder, St. H. (1987). Polysaccharide reserve material in the acetotrophic methanogen, *Methanosarcina thermophila* strain TM-1: accumulation and mobilization. *Arch. Microbiol*. 147: 109–116.

Nelson-Sathi, S. *et al*. (2012). Acquisition of 1,000 eubacterial genes physiologically transformed a methanogen at the origin of Haloarchaea. *Proc. Natl. Acad. Sci. U.S.A.* 109, 50: 20537-42.

Nelson-Sathi, S. *et al.* (2015). Origins of major archaeal clades correspond to gene acquisitions from bacteria. *Nature* 517, 7532: 77-80.

Ochman, H. *et al*. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* 405, 6784: 299-304.

Oparin, A. I. (1938). The origin of life. Translation with annotations by Sergius Morgulis. *New York: The Macmillan Company*.

Orgel, L. E. (2004). Prebiotic chemistry and the origin of the RNA world. *Crit. Rev. Biochem. Mol. Biol.* 39, 2: 99-123.

Pickl, A. *et al*. (2012). Fructose degradation in the haloarchaeon *Haloferax volcanii* involves a bacterial type phosphoenolpyruvate-dependent phosphotransferase system, fructose-1-phosphate kinase, and class II fructose-1,6-bisphosphate aldolase. *J. Bacteriol.* 194, 12: 3088-97.

Preiner, M *et al*. (2020). A hydrogen-dependent geochemical analogue of primordial carbon and energy metabolism. *Nat. Ecol. Evol.* 4, 4: 534-542.

Preiss, J. (2014). Glycogen: Biosynthesis and Regulation. *EcoSal Plus* 6, 1.

Ragsdale, S. W. (2008). Enzymology of the wood-Ljungdahl pathway of acetogenesis. *Ann. N. Y. Acad. Sci.* 1125: 129-36.

Ragsdale S. W. und Pierce E. (2008). Acetogenesis and the Wood-Ljungdahl pathway of CO_2 fixation. *Biochim. Biophys. Acta* 1784: 1873–1898.

Ragsdale, S. W. *et al*. (1987). Mössbauer, EPR, and optical studies of the corrinoid/iron-sulfur protein involved in the synthesis of acetyl coenzyme A by *Clostridium thermoaceticum*. *J. Biol. Chem.* 262, 29: 14289-97.

Raux, E. *et al*. (1997). A role for *Salmonella typhimurium* cbiK in cobalamin (vitamin B12) and siroheme biosynthesis. *J. Bacteriol.* 179, 10: 3202-12.

Raymann, K. *et al*. (2015). The two-domain tree of life is linked to a new root for the Archaea. *Proc. Natl. Acad. Sci. U.S.A.* 112, 21: 6670-5.

Reher, M. und Schönheit, P. (2006). Glyceraldehyde dehydrogenases from the thermoacidophilic euryarchaeota *Picrophilus torridus* and *Thermoplasma acidophilum*, key enzymes of the non-phosphorylative Entner-Doudoroff pathway, constitute a novel enzyme family within the aldehyde dehydrogenase superfamily. *FEBS Lett.* 580, 5: 1198-204.

Reher, M. *et al*. (2007). Glyceraldehyde-3-phosphate ferredoxin oxidoreductase (GAPOR) and nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase (GAPN), key

enzymes of the respective modified Embden-Meyerhof pathways in the hyperthermophilic crenarchaeota *Pyrobaculum aerophilum* and *Aeropyrum pernix*. *FEMS Microbiol. Lett.* 273, 2: 196-205.

Reinhardt, A. *et al*. (2019). L-Rhamnose catabolism in archaea. *Mol. Microbiol.* 111, 4: 1093-1108.

Richter, M. *et al*. (2024). Pyruvate-dependent growth of *Methanosarcina acetivorans*. *J. Bacteriol.* 206, 2: e0036323.

Rodionov, D. A. *et al*. (2003). Comparative genomics of the vitamin B12 metabolism and regulation in prokaryotes. *J. Biol. Chem.* 278, 42: 41148-59.

Romano, A. H. und Conway, T. (1996). Evolution of carbohydrate metabolic pathways. *Res. Microbiol*. 147: 448–455.

Ronimus, R. S. und Morgan, H. W. (2003). Distribution and phylogenies of enzymes of the Embden-Meyerhof-Parnas pathway from archaea and hyperthermophilic bacteria support a gluconeogenic origin of metabolism. *Archaea* 1: 199–221.

Sakuraba, H. und Ohshima, T. (2002). Novel energy metabolism in anaerobic hyperthermophilic archaea: a modified Embden-Meyerhof pathway. *J. Biosci. Bioeng.* 93: 441–448.

Santiago-Martínez, M. G. *et al*. (2016). The nutritional status of *Methanosarcina acetivorans* regulates glycogen metabolism and gluconeogenesis and glycolysis fluxes. *FEBS J.* 283, 10: 1979-99.

Sato, T. und Atomi, H. (2011). Novel metabolic pathways in archaea. *Curr. Opin. Microbiol*. 14: 307–314.

Sattler, C. *et al*. (2024). Physiological barriers for glucose utilization in *Methanosarcina acetivorans*. *biorxiv*.

Say, R. F. und Fuchs, G. (2010). Fructose 1,6-bisphosphate aldolase/phosphatase may be an ancestral gluconeogenic enzyme. *Nature* 464: 1077–1081.

Schlikker, M. L. *et al*. (2024). Conversion of pyridoxal to pyridoxamine with NH_3 and H_2 on nickel generates a protometabolic nitrogen shuttle under serpentinizing conditions. *FEBS J.* 10.1111/febs.17357.

Schmerk, C. L. *et al*. (2015). Elucidation of the *Burkholderia cenocepacia* hopanoid biosynthesis pathway uncovers functions for conserved proteins in hopanoid-producing bacteria. *Environ. Microbiol.* 17, 3: 735-50.

Schönheit, P. und Schäfer, T. (1995). Metabolism of hyperthermophiles. *World J. Microbiol. Biotechnol.* 11, 1: 26-57.

Schwander, L. *et al*. (2023). Serpentinization as the source of energy, electrons, organics, catalysts, nutrients and pH gradients for the origin of LUCA and life. *Front. Microbiol.* 14: 1257597.

Scossa, F. und Fernie A. R. (2020). The evolution of metabolism: how to test evolutionary hypotheses at the genomic level. *Comput. Struct. Biotechnol. J.* 18: 482-500.

Scott, A. I. *et al*. (1993). Biosynthesis of vitamin B12. Discovery of the enzymes for oxidative ring contraction and insertion of the fourth methyl group. *FEBS Lett.* 331, 1-2: 105-8.

Selig, M. *et al*. (1997). Comparative analysis of Embden-Meyerhof and Entner-Doudoroff glycolytic pathways in hyperthermophilic archaea and the bacterium *Thermotoga*. *Arch. Microbiol*. 167: 217–232.

Shemin, D. und Rittenberg, D. (1946). The biological utilization of glycine for the synthesis of the protoporphyrin of hemoglobin. *J. Biol. Chem.* 166, 2: 621-5.

Shieh, J. S. und Whitman, W. B. (1987). Pathway of acetate assimilation in autotrophic and heterotrophic methanococci. *J. Bacteriol.* 169, 11: 5327-9.

Siebers, B. und Hensel, R. (1993). Glucose catabolism of the hyperthermophilic archaeum *Thermoproteus tenax*. *FEMS Microbiol. Lett*. 111: 1–7.

Siebers, B. und Schönheit, P. (2005). Unusual pathways and enzymes of central carbohydrate metabolism in Archaea. *Curr. Opin. Microbiol.* 8, 6: 695-705.

Slotznick, S. P. *et al*. (2022). Reexamination of 2.5-Ga "whiff" of oxygen interval points to anoxic ocean before GOE. *Sci. Adv.* 8, 1: eabj7190.

Song, Y. *et al*. (2024). Effect of alkali- and alkaline-Earth-metal promoters on silica-supported Co-Fe alloy for autocatalytic CO_2 fixation. *Angew. Chem. (International ed. in English)* 63, 8: e202316110.

Sousa, F. L. und Martin, W. F. (2014). Biochemical fossils of the ancient transition from geoenergetics to bioenergetics in prokaryotic one carbon compound metabolism. *Biochim. Biophys. Acta* 1837, 7: 964-81.

Stupperich, E. und Kräutler, B. (1988). 5'-Methylbenzimidazolyl-cobamides are the corrinoid from some sulfate-reducing and sulfur-metabolizing bacteria. *Eur. J. Biochem.* 176: 461-469.

Stupperich, E. *et al*. (1988). Diversity of corrinoids in acetogenic bacteria. P-cresolylcobamide from *Sporomusa ovata*, 5-methoxy-6-methylbenzimidazolylcobamide from *Clostridium formicoaceticum* and vitamin B12 from *Acetobacterium woodii*. *Eur. J. Biochem.* 172, 2: 459-64.

Suzuki, E. und Suzuki, R. (2016). Distribution of glucan-branching enzymes among prokaryotes. *CMLS* 73, 14: 2643-60.

Svetlitchnaia, T. *et al*. (2006). Structural insights into methyltransfer reactions of a corrinoid iron-sulfur protein involved in acetyl-CoA synthesis. *Proc. Natl. Acad. Sci. U.S.A.* 103, 39: 14331-6.

Tästensen, J. B. und Schönheit, P. (2018). Two distinct glyceraldehyde-3-phosphate dehydrogenases in glycolysis and gluconeogenesis in the archaeon *Haloferax volcanii*. *FEBS Lett.* 592, 9: 1524-1534.

Tästensen, J. B. *et al*. (2020). D-galactose catabolism in archaea: operation of the DeLey-Doudoroff pathway in *Haloferax volcanii*. *FEMS Microbiol. Lett.* 367, 1: fnaa029.

Ueno, Y. *et al*. (2006). Evidence from fluid inclusions for microbial methanogenesis in the early Archaean era. *Nature* 440, 7083: 516-9.

Urey, H. C. (1952). On the early chemical history of the Earth and the origin of life. *Proc. Natl. Acad. Sci. U.S.A*. 38: 351–63.

Varma, S. J. *et al*. (2018). Native iron reduces CO_2 to intermediates and end-products of the acetyl-CoA pathway. *Nat. Ecol. Evol.* 2, 1019–1024.

Verhees, C. H. *et al*. (2001). The phosphoglucose isomerase from the hyperthermophilic archaeon *Pyrococcus furiosus* is a unique glycolytic enzyme that belongs to the cupin superfamily. *J. Biol. Chem.* 276, 44: 40926-32.

Vitols, E. *et al*. (1966). Enzymatic conversion of vitamin B-12s to a cobamide coenzyme, alpha-(5,6-dimethylbenzimidazolyl)deoxyadenosylcobamide (adenosyl-B-12). *J. Biol. Chem.* 241, 7: 1455-61.

Wächtershäuser, G. (1988). Before enzymes and templates: theory of surface metabolism. *Microbiol. Rev.* 52, 4: 452-84.

Wald, G. (1964). The origins of life. *Proc. Nat. Acad. Sci. U.S.A.* 52: 595-611.

Wallin, I.E. (1927). Symbionticism and the origin of species. *London: Bailliere, Tindall and Cox*; 171.

Wang, L. *et al*. (2019). Bioinformatics analysis of metabolism pathways of archaeal energy reserves. *Sci. Rep*.; 9: 1034.

Warren, M. J. *et al*. (2002). The biosynthesis of adenosylcobalamin (vitamin B12). *Nat. Prod. Rep.* 19, 4: 390-412.

Wei, Y. *et al*. (2017). Structural and functional insights into corrinoid iron-sulfur protein from human pathogen *Clostridium difficile*. *J. Inorg. Biochem.* 170: 26-33.

Weiss, M. C. *et al*. (2016). The physiology and habitat of the last universal common ancestor. *Nat. Microbiol.* 1, 9: 16116.

Wilkinson, J. F. (1959). The problem of energy-storage compounds in bacteria. *Experiment. Cell. Res*. 7: 111–130.

Wimmer, J. L. E. *et al*. (2021). The autotrophic core: an ancient network of 404 reactions converts H_2, CO_2, and NH_3 into amino acids, bases, and cofactors. *Microorganisms* 9, 2: 458.

Woese, C. R. und Fox, G. E. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. U.S.A.* 74, 11: 5088-90.

Woese, C. (1998). The universal ancestor, *Proc. Natl. Acad. Sci. U.S.A.* 95, 12: 6854-6859.

Ycas, M. (1955). A note on the origin of life. *Proc. Natl. Acad. Sci. U.S.A.* 41, 10: 714-6.

Yu, J. P. *et al*. (1994). Pathway of glycogen metabolism in *Methanococcus maripaludis*. *J. Bacteriol.* 176, 2: 325-332.

Zhang, Y. *et al*. (2009). Comparative genomic analyses of nickel, cobalt and vitamin B12 utilization. *BMC Genomics.* 10: 78.

Zimorski, V. *et al*. (2014). Endosymbiotic theory for organelle origins. *Curr. Opin. Microbiol.* 22: 38-48.

Danksagung

Ich möchte mich herzlich bei einer Reihe von Menschen bedanken, die mich in den letzten Jahren begleitet und unterstützt haben.

Allen voran möchte ich mich bei Prof. Dr. William F. Martin bedanken, der mir die Möglichkeit bot, meine Bachelor- und Masterarbeit im Institut für Molekulare Evolution anzufertigen und mir mit seinem Vertrauen und seiner Unterstützung diese Dissertation ermöglichte. Ich bedanke mich außerdem herzlich bei meinem Mentor und Zweitgutachter Prof. Dr. Martin Lercher.

Darüber hinaus bedanke ich mich bei Prof. Dr. Peter Schönheit und Assoz. Prof. Dr. Filipa Sousa für herausragende Kooperationen und einen wertvollen wissenschaftlichen Austausch.

Ich bedanke mich herzlich bei dem gesamten MolEvol-Team, allen aktuellen sowie ehemaligen Kolleginnen und Kollegen, aber insbesondere dem Spaßbüro. Ihr habt diese Zeit unvergesslich gemacht, wir haben zusammen viele Erinnerungen geschaffen und es herrschte stets eine freundschaftliche Atmosphäre, die ein angenehmes Arbeitsumfeld geschaffen hat.

Ein großer Dank gilt allen Korrekturleserinnen und -lesern, die mir ihre Zeit und Hilfe zur Verfügung gestellt haben.

Besonders bedanken möchte ich mich bei meiner Frau Franzisca. Du hast mich immer bedingungslos unterstützt und mir, besonders in Phasen des Zweifels, die Motivation gegeben, die ich in den letzten Jahren gebraucht habe. Ich bin unendlich dankbar für deinen Rückhalt und dafür, dich an meiner Seite zu haben.

Abschließend möchte ich meiner Familie, insbesondere meinen Eltern, danken. Ihr habt mir stets euren Rückhalt geboten, mich ebenfalls ausnahmslos unterstützt und mir immer ein gutes Gefühl gegeben, was meine Entscheidungen betrifft. Ich weiß, dass ich immer auf euch zählen konnte.