The FEBS Journal

FEBS PRESS science publishing by scientists

Something special about CO-dependent CO_2 fixation

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

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

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

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

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

doi:10.1111/febs.14664

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

Introduction

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

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

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

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

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

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

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

Results and Discussion

CO_2 is everywhere in metabolism, CO is rare

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

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

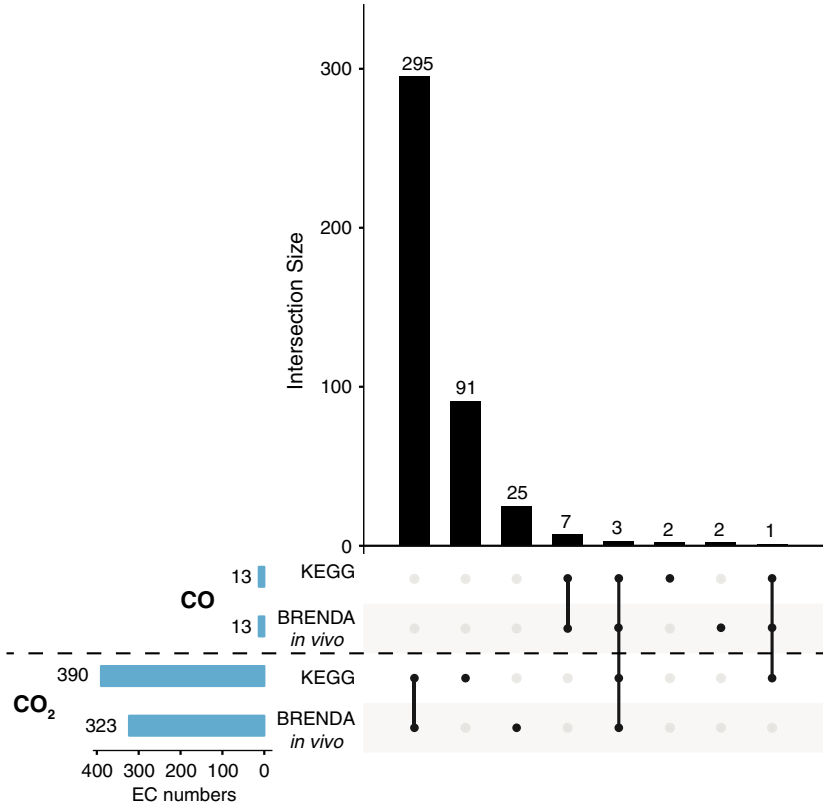

Fig. 1. Enzyme Commission (EC) numbers involving CO_2 and CO in KEGG and BRENDA (only *in vivo* reactions) and their overlaps. The horizontal bars display the total of EC numbers for each molecule in each database. The vertical bars display the size of the overlaps (intersections) between the databases.

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

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

CO_2 for all trades, CO only for CODH

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

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

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

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

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

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

Table 1. Enzyme commission numbers associated with carbon monoxide.

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

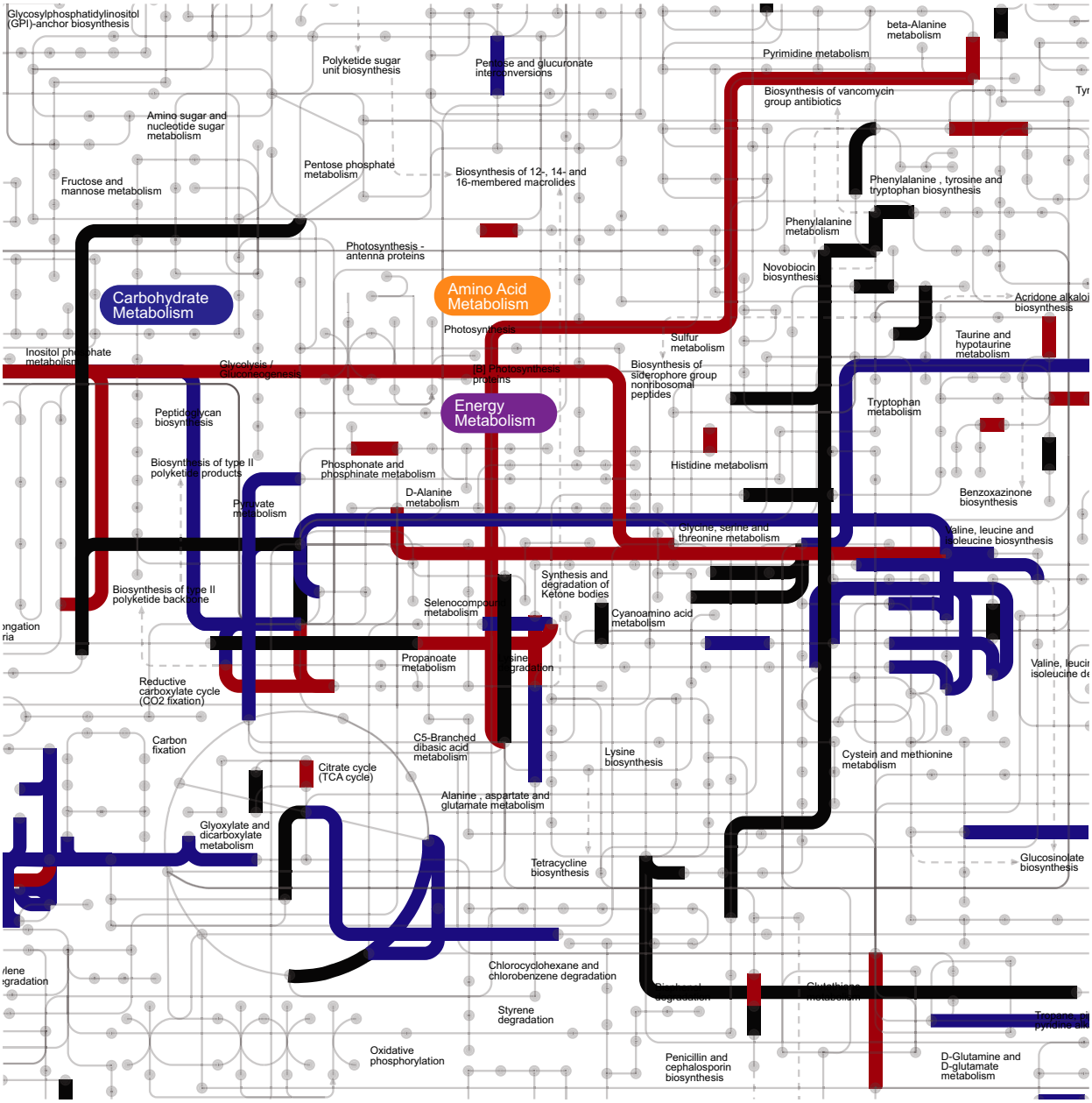

Fig. 3. CO_2 in a section of a global metabolic map (full map provided as Fig. S1). A portion of the KEGG map '01100 – metabolic pathways' with reactions involving CO_2 highlighted, portraying different directionality and reversibility assignments in BRENDA. In black, reactions not in BRENDA or where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

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

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

Metals and cofactors

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

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

CODH/ACS: Archaea and bacteria, but not aerobes

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

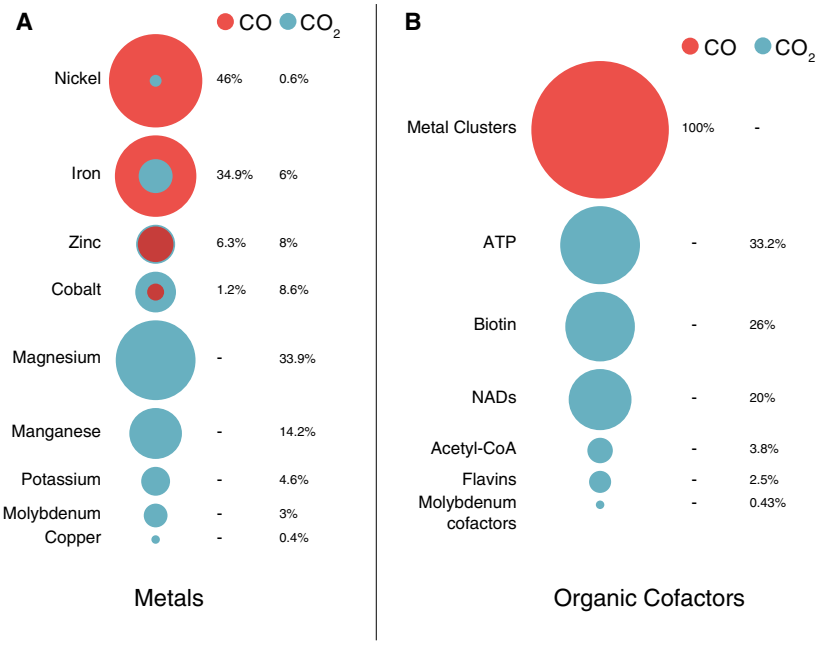

Fig. 4. Metals and organic cofactors in reactions that consume CO or CO_2. Percentage of entries of experimental evidence in BRENDA demonstrating the participation of different (A) metals and (B) cofactors in the catalytic activity of enzymes that use CO (red) or CO_2 (blue) as substrates.

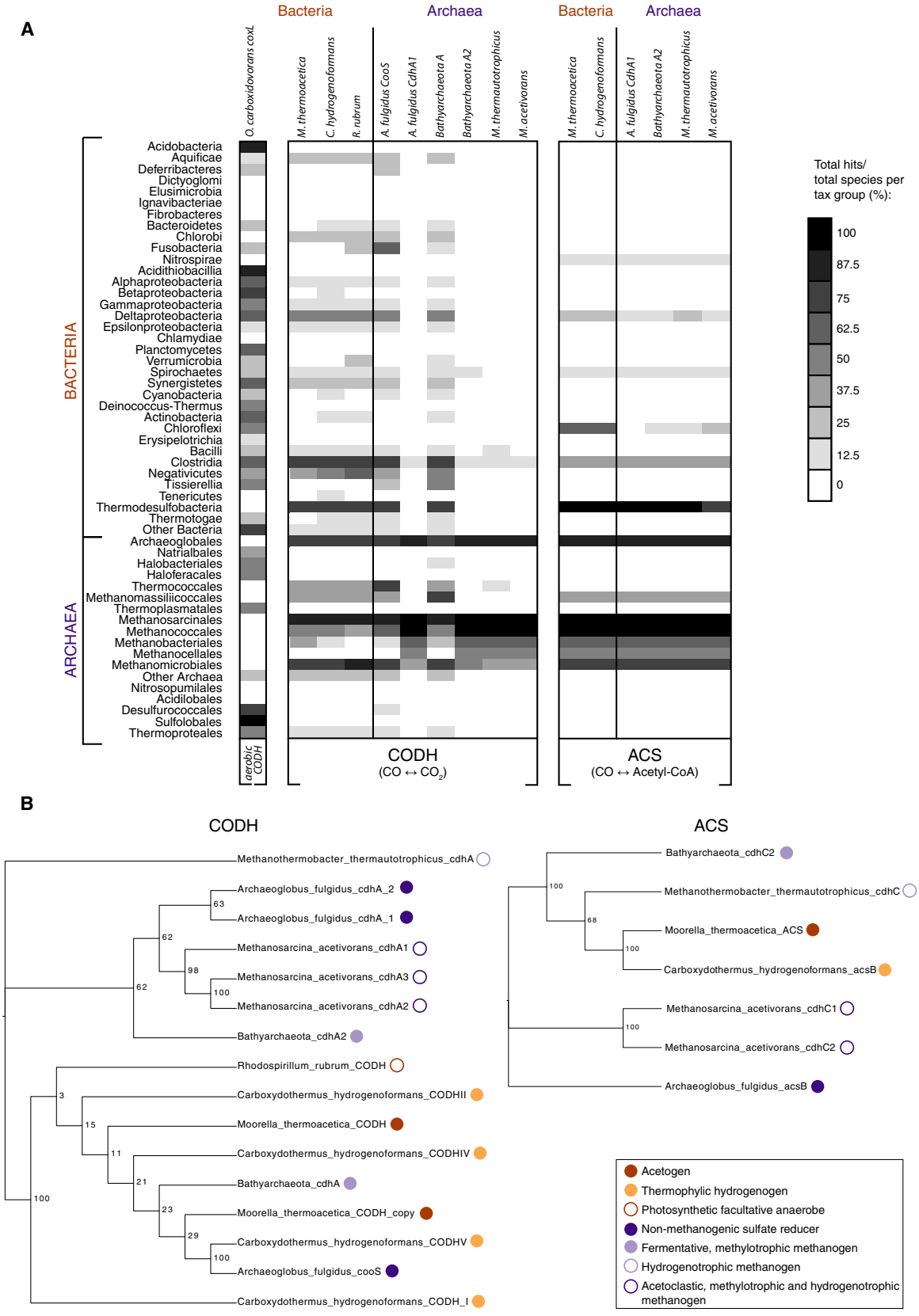
A
Bacteria
Archaea
Bacteria
Archaea
O. carboxidovorans coxL
M. thermoacetica
C. hydrogenoformans
R. rubrum
A. fulgidus CooS
A. fulgidus CdhA1
Bathyarchaeota A
Bathyarchaeota A2
M. thermautotrophicus
M. acetivorans
M. thermoacetica
C. hydrogenoformans
A. fulgidus CdhA1
Bathyarchaeota A2
M. thermautotrophicus
M. acetivorans
BACTERIA
ARCHAEA
Acidobacteria
Aquificae
Deferribacteres
Dictyoglomi
Elusimicrobia
Ignavibacteriae
Fibrobacteres
Bacteroidetes
Chlorobi
Fusobacteria
Nitrospirae
Acidithiobacillia
Alphaproteobacteria
Betaproteobacteria
Gammaproteobacteria
Deltaproteobacteria
Epsilonproteobacteria
Chlamydiae
Planctomycetes
Verrumicrobia
Spirochaetes
Synergistetes
Cyanobacteria
Deinococcus-Thermus
Actinobacteria
Chloroflexi
Erysipelotrichia
Bacilli
Clostridia
Negativicutes
Tissierellia
Tenericutes
Thermodesulfobacteria
Thermotogae
Other Bacteria
Archaeoglobales
Natrialbales
Halobacteriales
Haloferacales
Thermococcales
Methanomassiliicoccales
Thermoplasmatales
Methanosarcinales
Methanococcales
Methanobacteriales
Methanocellales
Methanomicrobiales
Other Archaea
Nitrosopumilales
Acidilobales
Desulfurococcales
Sulfolobales
Thermoproteales
aerobic CODH
CODH
(CO ↔ CO2)
ACS
(CO ↔ Acetyl-CoA)
Total hits/
total species per
tax group (%):
100
87.5
75
62.5
50
37.5
25
12.5
0
B
CODH
ACS
Methanothermobacter_thermautotrophicus_cdhA
Archaeoglobus_fulgidus_cdhA_2
Archaeoglobus_fulgidus_cdhA_1
Methanosarcina_acetivorans_cdhA1
Methanosarcina_acetivorans_cdhA3
Methanosarcina_acetivorans_cdhA2
Bathyarchaeota_cdhA2
Rhodospirillum_rubrum_CODH
Carboxydothermus_hydrogenoformans_CODHII
Moorella_thermoacetica_CODH
Carboxydothermus_hydrogenoformans_CODHIV
Bathyarchaeota_cdhA
Moorella_thermoacetica_CODH_copy
Carboxydothermus_hydrogenoformans_CODHV
Archaeoglobus_fulgidus_cooS
Carboxydothermus_hydrogenoformans_CODH_I
63
62
98
100
62
3
15
11
21
23
29
100
100
Bathyarchaeota_cdhC2
Methanothermobacter_thermautotrophicus_cdhC
Moorella_thermoacetica_ACS
Carboxydothermus_hydrogenoformans_acsB
Methanosarcina_acetivorans_cdhC1
Methanosarcina_acetivorans_cdhC2
Archaeoglobus_fulgidus_acsB
100
68
100
100
Acetogen
Thermophylic hydrogenogen
Photosynthetic facultative anaerobe
Non-methanogenic sulfate reducer
Fermentative, methylotrophic methanogen
Hydrogenotrophic methanogen
Acetoclastic, methylotrophic and hydrogenotrophic methanogen

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

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

CO forms stronger bonds to metals than CO_2

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

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

Conclusion

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

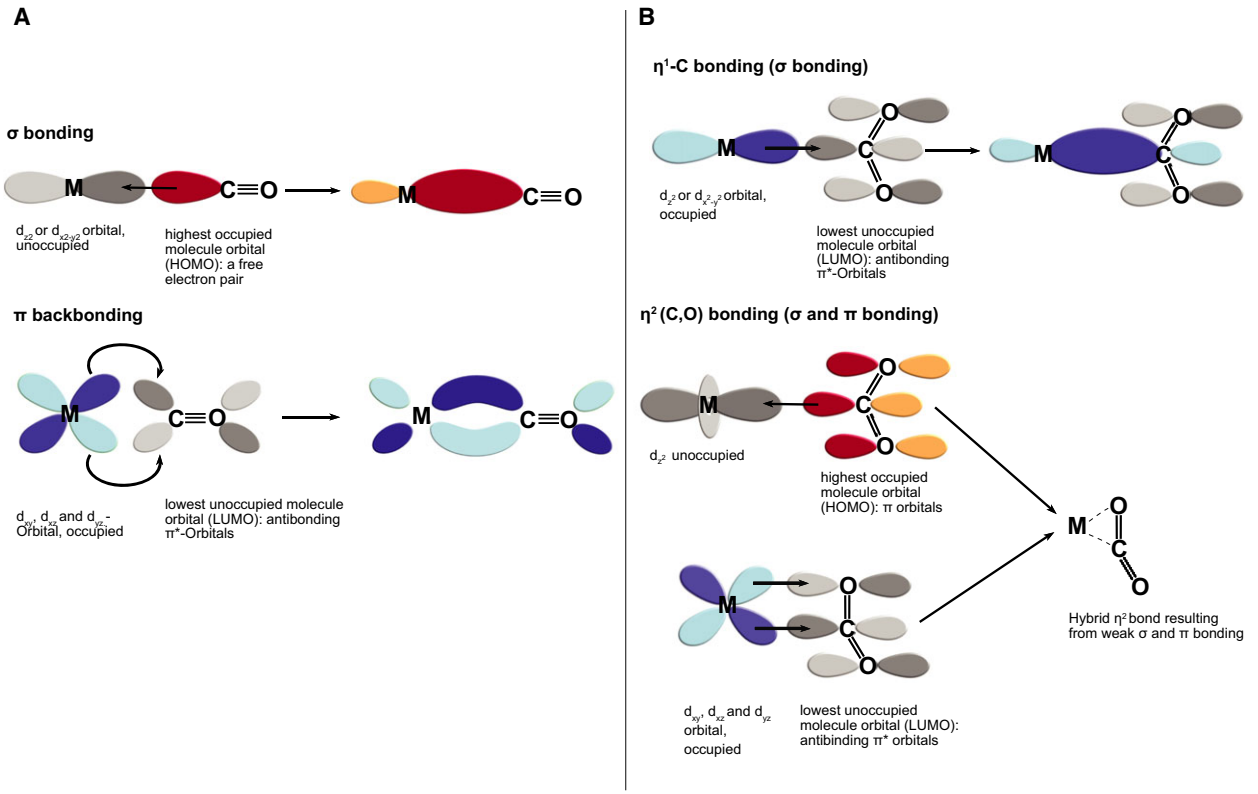

Fig. 6. CO (A) and CO_2 (B) bonding to transition metals. (A) CO binds to a transition metal (M) via the free electron pair of its carbon atom. The electron density in this orbital (red = positive phase, yellow = negative phase) can be placed into empty metal d orbitals forming a σ bond. Concurrently,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

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

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

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

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

$$2CO_2 + 4H_2 + \text{CoASH} \rightarrow CH_3\text{COSCoA} + 3H_2O \qquad \Delta G'_0 = -59\text{kJ} \cdot \text{mol}^{-1} \qquad (2)$$

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

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

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

Materials and methods

Data retrieval and integration

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

Taxonomy annotation

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

Statistical analysis and metabolic maps

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

Analysis of distributions of CO enzymes

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

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

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

Acknowledgements

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

Author contributions

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

References

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

23 Varma SJ, Muchowska KB, Chatelain P & Moran J (2018) Native iron reduces CO_2 to intermediates and end-products of the acetyl-CoA pathway. *Nat Ecol Evol* **2**, 1019–1024.
24 Sousa FL, Preiner M & Martin WF (2018) Native metals, electron bifurcation, and CO_2 reduction in early biochemical evolution. *Curr Opin Microbiol* **43**, 77–83.
25 Buckel W & Thauer RK (2018) Flavin-based electron bifurcation, ferredoxin, flavodoxin, and anaerobic respiration with protons (Ech) or NAD+(Rnf) as electron acceptors: a historical review. *Front Microbiol* **9**, https://doi.org/10.3389/fmicb.2018.00401
26 Adam PS, Borrel G & Gribaldo S (2018) Evolutionary history of carbon monoxide dehydrogenase/acetyl-CoA synthase, one of the oldest enzymatic complexes. *Proc Natl Acad Sci USA* **115**, E1166–E1173 (erratum appears in Proc Natl Acad Sci USA 115, E5836–E5837).
27 Techtmann SM, Lebedinsky AV, Colman AS, Sokolova TG, Woyke T, Goodwin L & Robb FT (2012) Evidence for horizontal gene transfer of anaerobic carbon monoxide dehydrogenases. *Front Microbio* **3**, https://doi.org/10.3389/fmicb.2012.00132
28 Vorholt J, Kunow J, Stetter KO & Thauer RK (1995) Enzymes and coenzymes of the carbon monoxide dehydrogenase pathway for autotrophic CO_2 fixation in *Archaeoglobus lithotrophicus* and the lack of carbon monoxide dehydrogenase in the heterotrophic *A. profundus*. *Arch Microbiol* **163**, 112–118.
29 Ragsdale SW (2008) Enzymology of the Wood-Ljungdahl pathway of acetogenesis. *Ann N Y Acad Sci* **1125**, 129–136.
30 Ragsdale SW (2004) Life with carbon monoxide. *Crit Rev Biochem Mol Biol* **39**, 165–195.
31 Sokolova TG, Henstra A-M, Sipma J, Parshina SN, Stams AJM & Lebedinsky AV (2009) Diversity and ecophysiological features of thermophilic carboxydotrophic anaerobes. *FEMS Microbiol Ecol* **68**, 131–141.
32 Diender M, Stams AJM & Sousa DZ (2015) Pathways and bioenergetics of anaerobic carbon monoxide fermentation. *Front Microbiol* **6**, https://doi.org/10.3389/fmicb.2015.01275
33 Borrel G, Adam PS & Gribaldo S (2016) Methanogenesis and the Wood-Ljungdahl pathway: an ancient, versatile, and fragile association. *Genome Biol Evol* **8**, 1706–1711.
34 Schuchmann K & Müller V (2014) Autotrophy at the thermodynamic limit of life: a model for energy conservation in acetogenic bacteria. *Nat Rev Microbiol* **12**, 809–821.
35 Kopke M, Held C, Hujer S, Liesegang H, Wiezer A, Wollherr A, Ehrenreich A, Liebl W, Gottschalk G & Durre P (2010) *Clostridium ljungdahlii* represents a microbial production platform based on syngas. *Proc Natl Acad Sci USA* **107**, 13087–13092.
36 Jansen K, Thauer RK, Widdel F & Fuchs G (1984) Carbon assimilation pathways in sulfate reducing bacteria. Formate, carbon dioxide, carbon monoxide, and acetate assimilation by *Desulfovibrio baarsii*. *Arch Microbiol* **138**, 257–262.
37 Oelgeschläger E & Rother M (2008) Carbon monoxide-dependent energy metabolism in anaerobic bacteria and archaea. *Arch Microbiol* **190**, 257–269.
38 Bender G, Pierce E, Hill JA, Darty JE & Ragsdale SW (2011) Metal centers in the anaerobic microbial metabolism of CO and CO_2. *Metallomics* **3**, 797–815.
39 Can M, Armstrong FA & Ragsdale SW (2014) Structure, function, and mechanism of the nickel metalloenzymes, CO dehydrogenase, and acetyl-CoA synthase. *Chem Rev* **114**, 4149–4174.
40 Gregg CM, Goetzl S, Jeoung JH & Dobbek H (2016) AcsF catalyzes the ATP-dependent insertion of Nickel into the Ni, Ni-[4Fe4S] cluster of Acetyl-CoA synthase. *J Biol Chem* **291**, 18129–18138.
41 Doukov TI, Iverson TM, Seravalli J, Ragsdale SW & Drennan CL (2002) A Ni-Fe-Cu center in a bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase. *Science* **298**, 567–572.
42 Maynard EL & Lindahl PA (1999) Evidence of a molecular tunnel connecting the active sites for CO_2 reduction and acetyl-CoA synthesis in acetyl-CoA synthase from *Clostridium thermoaceticum*. *J Am Chem Soc* **121**, 9221–9222.
43 Seravalli J & Ragsdale SW (2000) Channeling of carbon monoxide during anaerobic carbon dioxide fixation. *Biochemistry* **39**, 1274–1277.
44 Gong W, Hao B, Wei Z, Ferguson DJ, Tallant T, Krzycki JA & Chan MK (2008) Structure of the alpha2epsilon2 Ni-dependent CO dehydrogenase component of the *Methanosarcina barkeri* acetyl-CoA decarbonylase/synthase complex. *Proc Natl Acad Sci USA* **105**, 9558–9563.
45 King GM & Weber CF (2007) Distribution, diversity and ecology of aerobic CO-oxidizing bacteria. *Nat Rev Microbiol* **5**, 107–118.
46 Dobbek H, Gremer L, Meyer O & Huber R (1999) Crystal structure and mechanism of CO dehydrogenase, a molybdo iron-sulfur flavoprotein containing S-selanylcysteine. *Proc Natl Acad Sci USA* **96**, 8884–8889.
47 Wilcoxen J & Hille R (2013) The hydrogenase activity of the molybdenum/copper-containing carbon monoxide dehydrogenase of *Oligotropha carboxidovorans*. *J Biol Chem* **288**, 36052–36060.
48 Hille R, Dingwall S & Wilcoxen J (2015) The aerobic CO dehydrogenase from *Oligotropha carboxidovorans*. *J Biol Inorg Chem* **20**, 243–251.
49 Meyer O, Jacobitz S & Krüger B (1986) Biochemistry and physiology of aerobic carbon monoxide-oxidizing bacteria. *FEMS Microbiol Rev* **39**, 161–179.

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

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

Supporting information

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

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