

Article

High CO₂ levels drive the TCA cycle backwards towards autotrophy

<https://doi.org/10.1038/s41586-021-03456-9>

Received: 12 August 2020

Accepted: 15 March 2021

Published online: 21 April 2021

 Check for updates

Lydia Steffens^{1,6}, Eugenio Pettinato^{1,6}, Thomas M. Steiner^{2,6}, Achim Mall^{3,4}, Simone König⁵, Wolfgang Eisenreich² & Ivan A. Berg¹ 

It has recently been shown that in anaerobic microorganisms the tricarboxylic acid (TCA) cycle, including the seemingly irreversible citrate synthase reaction, can be reversed and used for autotrophic fixation of carbon^{1,2}. This reversed oxidative TCA cycle requires ferredoxin-dependent 2-oxoglutarate synthase instead of the NAD-dependent dehydrogenase as well as extremely high levels of citrate synthase (more than 7% of the proteins in the cell). In this pathway, citrate synthase replaces ATP-citrate lyase of the reductive TCA cycle, which leads to the spending of one ATP-equivalent less per one turn of the cycle. Here we show, using the thermophilic sulfur-reducing delta-proteobacterium *Hippea maritima*, that this route is driven by high partial pressures of CO₂. These high partial pressures are especially important for the removal of the product acetyl coenzyme A (acetyl-CoA) through reductive carboxylation to pyruvate, which is catalysed by pyruvate synthase. The reversed oxidative TCA cycle may have been functioning in autotrophic CO₂ fixation in a primordial atmosphere that is assumed to have been rich in CO₂.

Various natural autotrophic CO₂ fixation pathways have been identified to date^{3–5}. These pathways are adapted to organisms that live in specific ecological niches. In addition to the variety in autotrophic CO₂ fixation mechanisms, variants of these pathways exist that strongly differ from each other in their properties^{4,6}. A recently identified reversed oxidative tricarboxylic acid (roTCA) cycle^{1,2} is a version of the autotrophic reductive tricarboxylic acid (rTCA) cycle. The rTCA cycle differs from the oxidative TCA cycle mainly by replacing reactions that are irreversible under physiological conditions with reversible ones—that is, using ferredoxin-dependent 2-oxoglutarate synthase instead of the NAD-dependent dehydrogenase and using ATP-citrate lyase (ACL) (citrate + ATP + CoA → oxaloacetate + acetyl-CoA + ADP + P_i) instead of citrate synthase (acetyl-CoA + oxaloacetate + H₂O → citrate + CoA). By contrast, the roTCA cycle uses citrate synthase rather than ACL^{1,2} (Fig. 1a). The usage of citrate synthase for citrate cleavage is thermodynamically unfavourable (free-energy change ΔG' of more than 35 kJ mol⁻¹)⁷ but enables the spending of one ATP molecule less per acetyl-CoA synthesized from two molecules of CO₂, making the roTCA cycle a highly efficient autotrophic carbon fixation pathway. This pathway was shown in two obligate anaerobic bacteria that exhibit hydrogen oxidation coupled with sulfur reduction, the delta-proteobacterium *Desulfurella acetivorans*² and *Thermosulfidibacter takaii*, a representative of the Aquificae phylum¹. These bacteria grow equally well under autotrophic and heterotrophic conditions and have generation times of 5–7 h^{1,2}, which suggests that they are not affected kinetically by the use of this pathway. Therefore, the question arises which factors limit the distribution of the roTCA cycle—or, in other words, why the roTCA cycle did not outcompete the apparently less efficient ACL-containing rTCA cycle, considering the almost universal distribution of citrate synthase in autotrophic bacteria.

An important feature of the roTCA cycle is that it lacks unique enzymes, making it cryptic for bioinformatics analyses and possibly leading to the underestimation of its occurrence in the microbial world. The only characteristic feature is an extremely high activity of the citrate synthase reaction in cell extracts (more than 10 μmoles min⁻¹ mg⁻¹ protein)^{1,2}. Therefore, we decided to study the enzyme that catalyses this reaction in *D. acetivorans* in more detail. The genome contains three citrate synthase homologues, *desace_08345*, *desace_09325* and *desace_06860* (Extended Data Fig. 1). In order to determine which of the corresponding proteins is responsible for the citrate synthase reaction *in vivo*, we heterologously produced all three proteins in *Escherichia coli* (Extended Data Fig. 2) and used them as standards for protein quantification with target high-resolution mass spectrometry. *Desace_08345* was shown to be the main citrate synthase in *D. acetivorans* and one of the most abundant proteins in the lysates of autotrophically grown cells, contributing 7.2% to the total protein content, whereas *Desace_09325* and *Desace_06860* were much less abundant (0.2 and 0.1%, respectively) (Extended Data Fig. 3). Furthermore, comparison of expression in the proteomes of the autotrophically and heterotrophically grown cells hardly revealed any regulation in the synthesis of the enzymes of the roTCA cycle (Supplementary Tables 1, 2). Thus, *Desace_08345* is responsible for both the cleavage and synthesis of citrate, depending on the growth conditions.

Desace_08345 had the kinetic properties of a typical citrate synthase from heterotrophic organisms, being highly active in the direction of citrate synthesis but having only low activity in the reaction of citrate cleavage (V_{max} of 157 and 1 μmoles min⁻¹ mg⁻¹ protein for acetyl-CoA and citrate, respectively) (Extended Data Table 1). This is in line with the Haldane equation⁸, which relates the catalytic efficiencies (V_{max}/K_m) of the backward and forward reactions to the equilibrium constant

¹Institute for Molecular Microbiology and Biotechnology, University of Münster, Münster, Germany. ²Bavarian NMR Center—Structural Membrane Biochemistry, Department of Chemistry, Technische Universität München, Garching, Germany. ³K. G. Jebsen Centre for Deep Sea Research, University of Bergen, Bergen, Norway. ⁴Department of Biological Sciences, University of Bergen, Bergen, Norway. ⁵Core Unit Proteomics, Interdisciplinary Center for Clinical Research, Medical Faculty, University of Münster, Münster, Germany. ⁶These authors contributed equally: Lydia Steffens, Eugenio Pettinato, Thomas M. Steiner.  e-mail: wolfgang.eisenreich@mytum.de; ivan.berg@uni-muenster.de

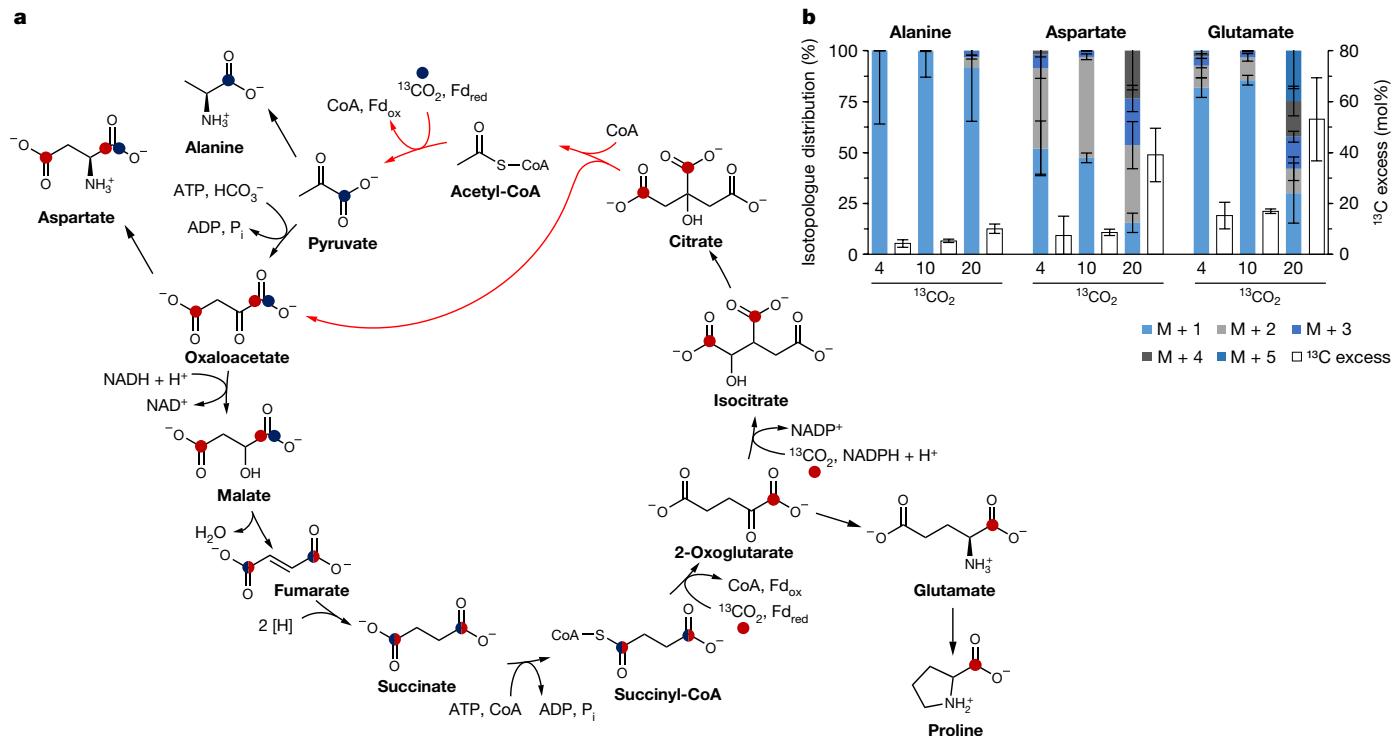


Fig. 1 | The roTCA cycle in *H. maritima*. **a**, The cycle and incorporation of $^{13}\text{CO}_2$ through pyruvate synthase and pyruvate carboxylase (blue) or 2-oxoglutarate synthase and isocitrate dehydrogenase (red). Half-circles indicate a statistical distribution of the ^{13}C label between both positions. **b**, ^{13}C -isotopologue fractions (normalized to 100%) and ^{13}C excess in Ala, Asp and Glu after growth

with H_2 , S^0 , 0.2 g l^{-1} yeast extract, and 4%, 10% or 20% $^{13}\text{CO}_2$ (Supplementary Tables 4–6). $M+X$ refers to isotopologues carrying $X^{13}\text{C}$ atoms, where X is 1–5. Data are mean \pm s.e.m. of six biological replicates for 4% $^{13}\text{CO}_2$ and of three biological replicates for 10% and 20% $^{13}\text{CO}_2$.

(2.24×10^6 at pH 7)⁷. Indeed, the Haldane equation predicts low V_{\max} and high K_m for an enzyme that catalyses the reversal of a reaction with a high equilibrium constant. Therefore, large amounts of the enzyme are required to maintain the required flux for citrate cleavage. This was indeed observed in *D. acetivorans* cells and may again indicate that the high abundance of citrate synthase is a characteristic feature of the roTCA cycle, due to the high expression of the corresponding genes. Notably, the kinetic properties of *D. acetivorans* citrate synthase differ from the properties of the enzyme from *T. takaii* (K_m values of 0.24 and 11 mM (Extended Data Table 1) versus 0.017 and 0.009 mM for CoA and citrate¹, respectively). The driving force for the evolution of the high specificity of the citrate synthase of *T. takaii* to its substrates is not clear, as high intracellular concentrations of citrate and CoA are necessary to enable the citrate cleavage reaction (1.4 and 1.1 mM in *D. acetivorans*, respectively)². The *T. takaii* enzyme is closely related to the citrate synthase of Deltaproteobacteria and eukaryotes (Extended Data Fig. 1), which do not share its high affinities for CoA and citrate².

Codon usage pattern of the genes can be used for the prediction of highly expressed genes⁹. We performed a codon usage analysis for various autotrophic bacteria using the software Interactive Codon Analysis (INCA)¹⁰ to evaluate whether this method can be used to predict the presence of the roTCA cycle based on the exceptionally high expression of the gene that encodes citrate synthase. In this analysis, the codon usage of ribosomal genes (which are supposed to be highly expressed) was compared to the codon usage of the remaining genes in the genome, and those genes that belonged to the top 400 expressed genes were regarded as highly expressed. The analysis was done for *D. acetivorans* and other Desulfurellaceae with fully sequenced genomes as well as for *T. takaii* and other sequenced Aquificae. As a control, a corresponding analysis was performed using 11 genomes of bacteria that belonged to different phylogenetic groups (Chlorobi, Nitrospirae, Alphaproteobacteria, Gammaproteobacteria

and Deltaproteobacteria) that use the ACL version of the cycle for autotrophic growth but nevertheless also possess a citrate synthase gene in their genomes. As expected, their citrate synthase genes were predicted to be only modestly expressed (Supplementary Table 3).

Among 13 studied Aquificae genomes, three possessed a citrate synthase gene: *T. takaii*, *Persephonella marina* and *Sulfurihydrogenibium azorensense*. The genomes of *P. marina* and *S. azorensense* also had a gene that encodes ACL. As expected, the citrate synthase gene of *T. takaii* was predicted to be highly expressed (Supplementary Table 3). Low expression of the citrate synthase gene and high expression of the ACL gene were predicted for *P. marina*. Indeed, this bacterium uses the classical rTCA cycle¹. However, genes encoding both ACL and citrate synthase are predicted to be highly expressed in *S. azorensense* (Supplementary Table 3), and only biochemical analysis could confirm the functioning of the ACL variant of the rTCA cycle in this bacterium (Extended Data Table 2). Therefore, although in silico analysis could predict the roTCA cycle in Aquificae, controversial cases need to be validated by biochemical assays.

Citrate synthase was predicted to be highly expressed in all sequenced Desulfurellaceae (7 genomes), with *D. amilsii* being the only exception (Supplementary Table 3). Notably, a high level of expression of the citrate synthase gene was predicted for *H. maritima*, which is not capable of autotrophic growth and requires yeast extract for growth¹¹. In our study, *H. maritima* grew not only with H_2 and CO_2 up to around 1.2×10^8 cells per ml by means of sulfur reduction (Extended Data Fig. 4a), but also only in the presence of yeast extract (Extended Data Fig. 5a)—that is, under mixotrophic conditions. A slight decrease in amino acid concentrations in the medium after growth and the absence of fermentation products suggest that yeast extract is used for anabolism (Extended Data Fig. 4b). Similar to *D. acetivorans*², *H. maritima* could grow under heterotrophic conditions with acetate in the presence of CO_2 and yeast extract, both with and without H_2 (Extended

Table 1 | Enzymes of carbon metabolism in *H. maritima*

Enzyme	Specific activity ($\mu\text{mol min}^{-1} \text{mg}^{-1}$ protein)		Candidate gene(s)
	Acetate + $\text{H}_2 + \text{CO}_2$	$\text{CO}_2 + \text{H}_2$	
Citrate synthase (F) ^a	15.9 ± 0.7 (n = 3)	16.2 ± 3.3 (n = 3)	AEA33516.1
Citrate synthase (R) (NADH oxidation) ^b	0.71 ± 0.11 (n = 3)	0.50 ± 0.04 (n = 3)	AEA33516.1
Citrate synthase (R) (acetyl-CoA formation) ^c	0.53 ± 0.04 (n = 2)	0.81 ± 0.05 (n = 3)	AEA33516.1
ATP-citrate lyase (F)	<0.001 (n = 2)	<0.001 (n = 3)	–
Citrate lyase (F)	<0.003 (n = 2)	<0.003 (n = 3)	–
Aconitase (F)	0.44 ± 0.20 (n = 3)	0.78 ± 0.20 (n = 2)	AEA33195.1
Isocitrate dehydrogenase (NADP) (F)	19.5 ± 1.7 (n = 3)	19.4 ± 1.4 (n = 3)	AEA34490.1
2-Oxoglutarate synthase (BV) (R)	2.07 ± 0.36 (n = 3)	1.83 ± 0.09 (n = 2)	AEA34486.1, AEA34487.1 and AEA34488.1
Succinyl-CoA synthetase (F)	0.23 ± 0.01 (n = 3)	0.14 ± 0.04 (n = 3)	AEA34365.1 and AEA34366.1
Fumarase (F)	3.93 ± 0.01 (n = 2)	3.46 ± 0.71 (n = 3)	AEA33115.1 and AEA33116.1
Malate dehydrogenase (NADH) (R)	38.9 ± 4.4 (n = 3)	57.5 ± 5.0 (n = 3)	AEA33114.1
Pyruvate synthase (MV) (R)	0.25 ± 0.06 (n = 3)	0.19 ± 0.04 (n = 3)	AEA33766.1, AEA33767.1 and AEA33768.1; AEA33558.1, AEA33559.1 and AEA33560.1
Pyruvate carboxylase (F)	0.06 ± 0.01 (n = 2)	0.05 ± 0.01 (n = 3)	AEA33111.1
PEP carboxylase (F)	<0.003 (n = 3)	<0.003 (n = 3)	–
PEP carboxykinase (F)	<0.003 (n = 3)	<0.003 (n = 3)	–
Malate dehydrogenase (decarboxylating) (NADP) (F)	0.1 ± 0.03 (n = 3)	0.19 ± 0.01 (n = 3)	AEA33079.1
PEP synthase (R)	0.11 ± 0.02 (n = 3)	0.10 ± 0.01 (n = 3)	AEA33056.1

Specific activities were measured at 55 °C; data are mean ± s.d. and the number of biological repetitions (n) is shown. For each biological replication, at least two technical replications were carried out. Forward (F) and reverse (R) indicate in which direction the enzyme was measured (with respect to enzyme name). GenBank accession for candidate genes are included. –, no gene in the genome; BV, benzyl viologen; MV, methyl viologen.

^aCitrate synthase was measured here as oxaloacetate-dependent CoA formation from acetyl-CoA.

^bCitrate synthase was measured here after oxaloacetate formation from citrate in the presence of CoA (with malate dehydrogenase).

^cAcetyl-CoA formation from citrate and CoA was measured with ultrahigh-performance liquid chromatography.

Data Fig. 5b). Nevertheless, *H. maritima* cell extracts possessed high activities of citrate synthase and other enzymes of the roTCA cycle, while ACL activity was absent (Table 1), and genes for the characteristic enzymes of other autotrophic pathways were missing. Experiments of ^{13}C incorporation with *H. maritima* during growth on 0.2 g l⁻¹ yeast extract in the presence of 4%, 10% or 20% of $^{13}\text{CO}_2$ showed that only a minor part of the carbon in proteins was derived from $^{13}\text{CO}_2$, decreasing from 16.1% to 4.4% with a decrease in the CO_2 concentration from 20% to 4%, respectively (Fig. 1b and Supplementary Tables 4–7). Moreover, heavier isotopologues that carried multiple ^{13}C labels of, for example, Asp, Glu and related amino acids were formed in particular at higher $^{13}\text{CO}_2$ concentrations. This suggests that the amount of CO_2 fixed into cell material was determined by the concentration of CO_2 (Fig. 1b and Supplementary Tables 4–6).

We therefore wondered whether the roTCA cycle is involved in mixotrophic metabolism of *H. maritima*. The high amount of fully labelled (M + 5) glutamate (22%) (Fig. 1b and Supplementary Table 6) in the cells grown at 20% $^{13}\text{CO}_2$ proves the operation of the roTCA cycle in this bacterium. This isotopologue arises through multiple rounds of the roTCA cycle with continuous $^{13}\text{CO}_2$ incorporation. Although ^{13}C incorporation was low when the bacteria grew with 4% $^{13}\text{CO}_2$ (Fig. 1b and Supplementary Table 4), the roTCA cycle also functioned under these growth conditions, as was revealed by the labelling patterns in amino acids after growth with 14 mg l⁻¹ of [1- ^{13}C]glutamate, unlabelled 4% CO_2 and 0.2 g l⁻¹ yeast extract. In general, [1- ^{13}C]glutamate can either be oxidized in the oxidative TCA cycle with the irrevocable loss of the label from the C1 of glutamate, or be converted into [1- ^{13}C]citrate and further into acetyl-CoA and [4- ^{13}C]oxaloacetate (and thus [4- ^{13}C]aspartate) through the reactions of the roTCA cycle (Fig. 2a). In *H. maritima*, the observed ^{13}C excess in multiple amino acids (Fig. 2b) strongly supports the latter hypothesis (Fig. 2b and Supplementary Table 8). As a signature for the Si-specific citrate synthase that is forming the (pro-S)-carboxymethyl moiety in citrate from acetyl-CoA, the

^{13}C contents in the mass fragments carrying C1–C4 (Asp418) and C2–C4 (Asp390) were identical, which suggests that mainly the C4 but not the C1 of aspartate carried the ^{13}C label (Fig. 2c). The fact that glutamate also acquired a ^{13}C label at positions 2–5, as indicated by the labelling of Glu404 (which carries C2–C5), further supported the functionality of the roTCA cycle.

The strong dependence of the metabolism of *H. maritima* on the concentration of CO_2 was further confirmed in growth experiments under mixotrophic conditions with 0.2 g l⁻¹ yeast extract. Whereas the culture grew well at 20% and 40% CO_2 in the gas mixture (generation time of 17 ± 1.1 h and 17.2 ± 1.4 h, respectively) and moderately at 5% CO_2 (generation time of 29.3 ± 1.9 h), no growth was detected with 1% or 2% CO_2 (Fig. 3a). The same dependence (under strictly autotrophic conditions) was observed for the other studied Desulfurellaceae species (*D. acetivorans*, *Desulfurella multipotens* and *Desulfurella propionica*) (Fig. 3b–d). By contrast, the sulfate-reducing autotrophic delta-proteobacterium *Desulfovobacter hydrogenophilus*, which reportedly uses the ACL variant of the rTCA cycle¹², grew equally well with 1%, 2%, 5%, 20% and 40% CO_2 in the gas phase (Fig. 3e). Note that all studied CO_2 concentrations are much higher than the atmospheric concentration of CO_2 (0.04%, around 40 Pa).

We next investigated how the dependence of the roTCA cycle on the CO_2 concentration could be explained. Isotopologue profiling at different $^{13}\text{CO}_2$ concentrations (Fig. 1) revealed a sharp difference in the ^{13}C enrichments of amino acids of, for example, Ala (which is derived from pyruvate) and Glu (which is derived from 2-oxoglutarate) (Fig. 1b). These differences became especially evident in the carboxylic atoms (Fig. 1a) that originate from CO_2 in very similar ferredoxin-dependent carboxylation reactions, which are catalysed by pyruvate and 2-oxoglutarate synthases. More specifically, the ^{13}C incorporation into C1 of Ala decreased with the decrease in CO_2 concentration faster than the incorporation into C1 of Glu (Extended Data Fig. 6) despite the fact that further metabolism of 2-oxoglutarate in isocitrate dehydrogenase

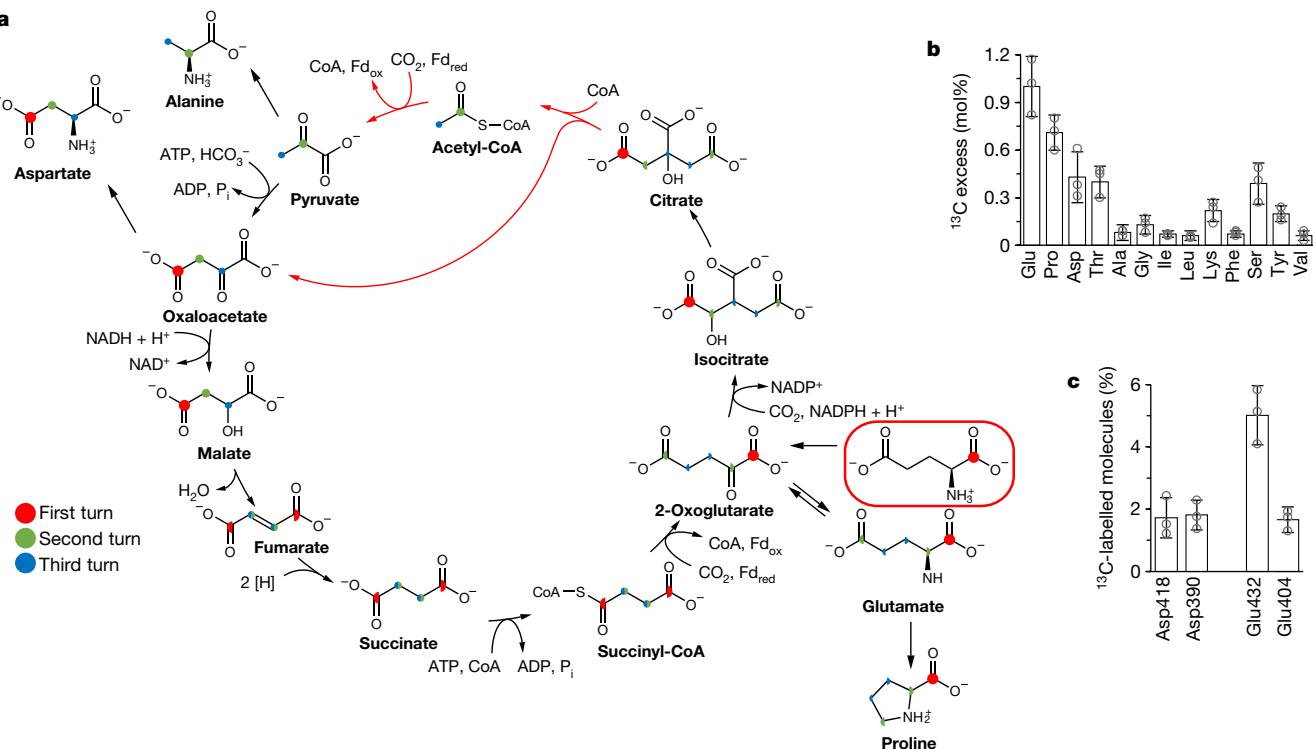


Fig. 2 | Incorporation of the $[1-^{13}\text{C}]$ glutamate in *H. maritima* via the roTCA cycle. **a**, Incorporation of exogenous $[1-^{13}\text{C}]$ glutamate (red box) after growth with 4% CO_2 , H_2 , S^0 , 0.2 g l⁻¹ yeast extract and 14 mg l⁻¹ $[1-^{13}\text{C}]$ glutamate via the roTCA cycle. Half-circles indicate a statistical distribution of the ^{13}C label

between both positions. **b**, Detected ^{13}C excess in amino acids. **c**, Percentage of ^{13}C -labelled molecules in different fragments of Asp (Asp418 = C1–4, Asp390 = C2–4) and Glu (Glu432 = C1–5, Glu404 = C2–5). Data are mean \pm s.e.m. of three biological replicates.

reaction is endergonic, whereas pyruvate conversion in either pyruvate carboxylase or phosphoenolpyruvate (PEP) synthase reactions is exergonic, accelerating pyruvate formation³. It can be concluded that *H. maritima* preferentially uses Ala from yeast extract, whereas Glu is preferentially synthesized from CO_2 , even though it is the most abundant amino acid in yeast extract (Extended Data Fig. 4). High consumption of Glu from the medium can probably account for the channelling of exogenous Glu into nitrogen metabolism. This observation suggests that pyruvate synthase is not functioning efficiently at low CO_2 concentrations. Indeed, the backward reaction of citrate synthase requires a high CoA:acetyl-CoA ratio (90 for *D. acetivorans* compared with around 2 for *E. coli*, in which the TCA cycle functions in the oxidative direction)^{2,13}. Acetyl-CoA carboxylation to pyruvate, with a standard

redox potential of -500 mV, is thermodynamically challenging³, and an unfavourable CoA:acetyl-CoA ratio places a further burden on this reaction. We propose that this is compensated for by an increased CO_2 concentration, enabling the assimilation of acetyl-CoA, the product of the roTCA cycle. In fact, the unfavourable CoA:acetyl-CoA ratio could be compensated for by an increase of around two orders of magnitude in inorganic carbon concentration, as predicted by our calculation of the bioenergetic efficiency of the 2-oxoglutarate synthase and pyruvate synthase reactions (Extended Data Fig. 7), whereas a 10^4 -fold increase in the concentration of inorganic carbon is necessary to compensate for the difference in free energy between the ACL and citrate synthase variants of the rTCA cycle (Extended Data Fig. 7). Notably, the *H. maritima* genome does not contain genes responsible for the synthesis of proteins involved in the active transport of inorganic carbon. Their usage would dissipate the energy rescued in the citrate cleavage reaction. A gene encoding a carbonic anhydrase is also missing from the genome.

In this study, we show that an increase in the CO_2 partial pressure allows bacteria to use the roTCA cycle. As citrate synthase in *D. acetivorans* does not have any specific adaptations to work backwards (Extended Data Table 1), this cycle can evolve easily and may be widespread among organisms that live in anoxic conditions at high CO_2 concentrations such as submarine hot vents and hydrothermal springs and sediments, where Desulfurellaceae are abundant^{14,15}. The concentration of CO_2 in the gas from the hot vents where *D. acetivorans* and *H. maritima* come from can be 90% and even higher^{16,17}. The roTCA cycle has recently been shown in *Geobacter sulfurreducens* (Deltaproteobacteria) during the growth on formate with Fe(III)citrate as the electron acceptor¹⁸. This species was previously considered to be strictly heterotrophic. In addition, genomic analyses suggested that the cycle is functional in ‘*Candidatus Nitrotheca patiens*’ (Chloroflexi)¹⁹ and *Deferribacter autotrophicus* (Deferribacteres)²⁰. In all of these cases, bacteria grew autotrophically in medium with elevated concentrations of inorganic carbon. Our codon usage analysis predicted that

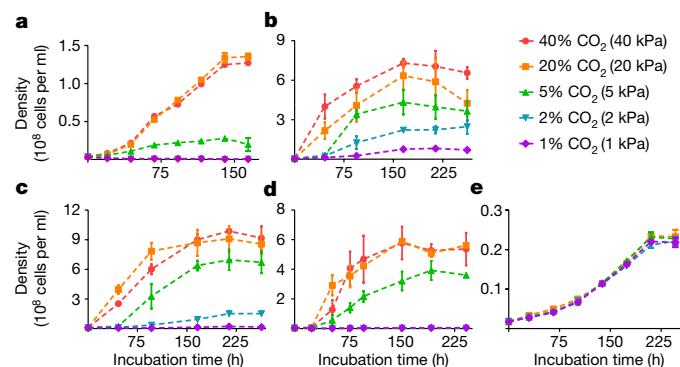


Fig. 3 | Growth of different bacteria depending on the concentration of CO_2 . **a–e**, Growth of *H. maritima* (**a**), *D. acetivorans* (**b**), *D. multipotens* (**c**), *D. propionica* (**d**) and *D. hydrogenophilus* (**e**) in H_2 (80 kPa) at 1% (1 kPa) CO_2 , 2% (2 kPa) CO_2 , 5% (5 kPa) CO_2 , 20% (20 kPa) CO_2 and 40% (40 kPa) CO_2 . For *H. maritima*, yeast extract (0.2 g l⁻¹) was added to the medium. Data are mean \pm s.e.m. of three biological replicates.

the citrate synthase gene of *G. sulfurreducens* and *D. autotrophicus* is highly expressed (Supplementary Table 3). This is in line with the proposed functioning of the roTCA cycle in these bacteria. The potential functioning of the roTCA cycle in uncultivated, hyperthermophilic members of the Aigarchaeota has also been discussed²¹, however a codon bias analysis for the closely related '*Candidatus Caldarchaeum subterraneum*' did not show a strong expression of citrate synthase (Supplementary Table 3), and it is uncertain whether these organisms can grow autotrophically. Nevertheless, the roTCA cycle could also contribute to mixotrophic growth, as we have shown for *H. maritima*. In fact, autotrophic growth using the rTCA (or roTCA) cycle has not been demonstrated in Archaea yet. The early proposals that the cycle functions in some Crenarchaeota were later rejected²².

The direction of the roTCA cycle is probably determined by the CO₂ concentration and/or the presence of acetate in the medium. Although some of the enzymes of the roTCA cycle were slightly more abundant in *D. acetivorans* under heterotrophic growth conditions (Supplementary Tables 1, 2), the regulation was not marked and reflected higher activities of the corresponding enzymes in cell extracts under these conditions². These unusually high activities of citrate synthase and malate dehydrogenase are necessary for citrate cleavage in the roTCA cycle but not for citrate synthesis during acetate oxidation. Thus, their high activity under both growth conditions² (Table 1) suggests that the cells are always prepared for autotrophic (or mixotrophic) CO₂ fixation through the roTCA cycle once environmental conditions enable its operation. The comparison of proteomes of *H. maritima* cells grown at 4% and 20% CO₂ hardly revealed any regulation in the synthesis of the roTCA cycle enzymes (Supplementary Tables 2, 9), which suggests that the roTCA cycle may be directly controlled by environmental conditions that permit its functioning (or do not permit it, which results in growth arrest).

The early atmosphere of Earth probably contained high amounts²³ of CO₂, creating ideal conditions for establishing the roTCA cycle as an ancient CO₂ fixation pathway. Previously, the ancestry of the roTCA cycle has been proposed based on phylometabolic evaluations¹. High concentrations of CO₂ still prevail in many contemporary ecological niches. Hydrothermal vent fluids are highly enriched in CO₂, with concentrations that are often between tens and hundreds of millimolars^{24,25}, and even a deep-sea hydrothermal vent that emits almost pure liquid CO₂ (2.7 moles CO₂ kg⁻¹ in hot vent fluid) has been described²⁶. In natural CO₂ reservoirs at moderate depths, the pressure of CO₂ can reach 40–50 MPa, compared with 40 Pa at ambient pressure in air²⁷. The eruption of the aquifers to the surface results in CO₂-driven cold-water geysers that bring to the surface groundwater that contains a complex microbial community, including autotrophic microorganisms, which use the reductive acetyl-CoA pathway and the rTCA (and possibly the roTCA) cycle, as revealed by coupled analyses using lipidomics and metagenomics²⁸.

An increased energetic efficiency at high CO₂ concentrations can also be envisaged for the reductive acetyl-CoA pathway (Wood-Ljungdahl pathway), arguably the most ancient metabolic pathway²⁹ that functions in acetogenic bacteria and methanogenic archaea. Although acetogenesis and methanogenesis are usually associated with a change in free energy ($\Delta G'$) that is insufficient to drive the synthesis of 1 mol ATP per mol of product under physiological conditions^{30,31}, the Gibbs free energy difference in this pathway may become sufficient for substrate phosphorylation at high partial pressures of CO₂ (50 kJ mol⁻¹ at 1 MPa CO₂, 10 mM acetate and 100 Pa H₂ partial pressure) (Extended Data Fig. 8 and Supplementary Discussion).

Biotechnological use of autotrophic organisms with high-CO₂-dependent, energetically efficient pathways can be advantageous in, for example, bioreactors under increased pressure, which results in a more efficient conversion of the substrate to a product. Even the existence of previously unknown pathways that are thermodynamically impossible under normal conditions could be envisaged,

thus offering a new perspective on the study and design of autotrophic CO₂ fixation pathways.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-021-03456-9>.

- Nunoura, T. et al. A primordial and reversible TCA cycle in a facultatively chemolithoautotrophic thermophile. *Science* **359**, 559–563 (2018).
- Mall, A. et al. Reversibility of citrate synthase allows autotrophic growth of a thermophilic bacterium. *Science* **359**, 563–567 (2018).
- Fuchs, G. Alternative pathways of carbon dioxide fixation: insights into the early evolution of life? *Annu. Rev. Microbiol.* **65**, 631–658 (2011).
- Berg, I. A. Ecological aspects of the distribution of different autotrophic CO₂ fixation pathways. *Appl. Environ. Microbiol.* **77**, 1925–1936 (2011).
- Song, Y. et al. Functional cooperation of the glycine synthase-reductase and Wood-Ljungdahl pathways for autotrophic growth of *Clostridium drakei*. *Proc. Natl. Acad. Sci. USA* **117**, 7516–7523 (2020).
- Könneke, M. et al. Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO₂ fixation. *Proc. Natl. Acad. Sci. USA* **111**, 8239–8244 (2014).
- Guynn, R. W., Gelberg, H. J. & Veech, R. L. Equilibrium constants of the malate dehydrogenase, citrate synthase, citrate lyase, and acetyl coenzyme A hydrolysis reactions under physiological conditions. *J. Biol. Chem.* **248**, 6957–6965 (1973).
- Bisswanger, H. *Enzyme Kinetics: Principles and Methods* (Wiley, 2008).
- Supek, F. & Vlahovicek, K. Comparison of codon usage measures and their applicability in prediction of microbial gene expressivity. *BMC Bioinformatics* **6**, 182 (2005).
- Supek, F. & Vlahovicek, K. INCA: synonymous codon usage analysis and clustering by means of self-organizing map. *Bioinformatics* **20**, 2329–2330 (2004).
- Miroshnichenko, M. L., Rainey, F. A., Rhode, M. & Bonch-Osmolovskaya, E. A. *Hippea maritima* gen. nov., sp. nov., a new genus of thermophilic, sulfur-reducing bacterium from submarine hot vents. *Int. J. Syst. Bacteriol.* **49**, 1033–1038 (1999).
- Schauder, R., Widdel, F. & Fuchs, G. Carbon assimilation pathways in sulfate-reducing bacteria. II. Enzymes of a reductive citric acid cycle in the autotrophic *Desulfobacter hydrogenophilus*. *Arch. Microbiol.* **148**, 218–225 (1987).
- Bennett, B. D. et al. Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat. Chem. Biol.* **5**, 593–599 (2009).
- Greene, A. C. in *The Prokaryotes* (eds Rosenberg, E. et al.) 135–142 (Springer, 2014).
- Dahle, H. et al. Energy landscapes in hydrothermal chimneys shape distributions of primary producers. *Front. Microbiol.* **9**, 1570 (2018).
- Zhao, W. *Diversity and Potential Geochemical Functions of Prokaryotes in Hot Springs of the Uzon Caldera, Kamchatka*. PhD thesis, Univ. Georgia, Athens (2008).
- Obzhirov, A. I. Chemistry of free and dissolved gases of Matupit Bay, Rabaul Caldera, Papua New Guinea. *Geo-Mar. Lett.* **12**, 54–59 (1992).
- Zhang, T., Shi, X.-C., Ding, R., Xu, K. & Tremblay, P.-L. The hidden chemolithoautotrophic metabolism of *Geobacter sulfurreducens* uncovered by adaptation to formate. *ISME J.* **14**, 2078–2089 (2020).
- Speck, E. et al. Extremophilic nitrite-oxidizing *Chloroflexi* from Yellowstone hot springs. *ISME J.* **14**, 364–379 (2020).
- Slobodkin, A. et al. Genomic insights into the carbon and energy metabolism of a thermophilic deep-sea bacterium *Deferribacter autotrophicus* revealed new metabolic traits in the phylum Deferribacterae. *Genes* **10**, 849 (2019).
- Hua, Z. S. et al. Genomic inference of the metabolism and evolution of the archaeal phylum Aigarchaeota. *Nat. Commun.* **9**, 2832 (2018).
- Berg, I. A. et al. Autotrophic carbon fixation in archaea. *Nat. Rev. Microbiol.* **8**, 447–460 (2010).
- Zahnle, K. et al. Emergence of a habitable planet. *Space Sci. Rev.* **129**, 35–78 (2007).
- German, C. R. & Von Damm, K. L. in *Treatise on Geochemistry* (eds Turekian, K. K. & Holland, H. D.) 18–222 (Elsevier, 2006).
- Pedersen, R. B. et al. Discovery of a black smoker vent field and vent fauna at the Arctic Mid-Ocean Ridge. *Nat. Commun.* **1**, 126 (2010).
- Lupton, J. et al. Submarine venting of liquid carbon dioxide on a Mariana Arc volcano. *Geochem. Geophys. Geosyst.* **7**, Q08007 (2006).
- Ajo-Franklin, J., Voltolina, M., Molins, S. & Yang, L. in *Geological Carbon Storage: Subsurface Seals and Caprock Integrity* (eds Vialle, S. et al.) 187–206 (Wiley, 2019).
- Probst, A. J. et al. Lipid analysis of CO₂-rich subsurface aquifers suggests an autotrophy-based deep biosphere with lysolipids enriched in CPR bacteria. *ISME J.* **14**, 1547–1560 (2020).
- Martin, W. F. Older than genes: the acetyl CoA pathway and origins. *Front. Microbiol.* **11**, 817 (2020).
- Thauer, R. K., Kaster, A. K., Seedorf, H., Buckel, W. & Hedderich, R. Methanogenic archaea: ecologically relevant differences in energy conservation. *Nat. Rev. Microbiol.* **6**, 579–591 (2008).
- Schuchmann, K. & Müller, V. Autotrophy at the thermodynamic limit of life: a model for energy conservation in acetogenic bacteria. *Nat. Rev. Microbiol.* **12**, 809–821 (2014).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2021

Article

Methods

Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Materials

Chemicals and gases were obtained from Sigma-Aldrich, Merck, Serva, Roth or VWR. Biochemicals were from Roche Diagnostics and Appli-Chem. ¹³C-labelled chemicals were obtained from Sigma-Aldrich and 99.0 atom% ¹³CO₂ gas was from Eurisotop. Materials for molecular biology were purchased from New England BioLabs, Qiagen, Merck and Sartorius. Materials and equipment for protein purification were obtained from GE Healthcare, Macherey-Nagel or Millipore. Lead acetate paper was obtained from Macherey-Nagel. Primers were synthesized by Sigma-Aldrich.

Synthesis of CoA esters

Acetyl-CoA was synthesized from the corresponding anhydrides and CoA according to a previously described method³².

Strains and growth conditions

All bacterial strains with the exception of *E. coli* were obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ). Their growth was determined using Neubauer counting chambers. In addition, the growth of sulfur and sulfate reducers was monitored through H₂S production using lead acetate paper.

H. maritima MH₂^T (DSM10411) was grown mixotrophically in medium (DSMZ medium 854) containing 6.2 mM NH₄Cl, 2.2 mM CaCl₂·2H₂O, 1.6 mM MgCl₂·6H₂O, 4.4 mM KCl, 2.4 mM KH₂PO₄, 427.8 mM NaCl, 14.3 mM MOPS buffer, 200 mg l⁻¹ yeast extract, 1 mg l⁻¹ resazurin, 10 g l⁻¹ sulfur powder, 10 ml l⁻¹ trace element solution out of DSMZ medium 141 and 1 ml l⁻¹ Wolfe's vitamin solution (20 mg l⁻¹ biotin, 20 mg l⁻¹ folic acid, 100 mg l⁻¹ pyridoxamine dihydrochloride, 50 mg l⁻¹ thiamine dihydrochloride, 50 mg l⁻¹ riboflavin, 50 mg l⁻¹ nicotinic acid, 50 mg l⁻¹ DL-Ca-pantothenate, 1 mg l⁻¹ cyanocobalamin, 50 mg l⁻¹ 4-aminobenzoic acid and 50 mg l⁻¹ lipoic acid). The medium was prepared without sulfur and vitamin solution, made anaerobic by bubbling with N₂ (100%) and reduced by the addition of Na₂S·9H₂O to a final concentration of 0.05% (w/v). The pH of the medium was adjusted to 5.5–6.0. The medium was dispensed anaerobically into serum bottles containing sulfur powder; the bottles were sealed with butyl-rubber stoppers and aluminium caps and autoclaved for 40 min at 110 °C. Before inoculation, the vitamin solution was added and the gas phase was replaced with the required quantities of H₂:CO₂ (60:40, 80:20, 95:5, 98:2 or 99:1) at 1 bar overpressure. For heterotrophic cultures, the medium was additionally supplemented with 5 g l⁻¹ sodium acetate and the gas phase was replaced with N₂:CO₂ (80:20) at 1 bar overpressure. In some cases, the cells were cultured without yeast extract or with N₂, H₂:N₂ (80:20) or N₂:CO₂ (80:20) as a gas phase. Cultures were incubated at 55 °C while shaking at 130 rpm. For labelling experiments, either 14 mg l⁻¹[¹⁻¹³C]glutamate or 99.0 atom% ¹³CO₂ gas (4, 10 or 20%) was added to the medium before inoculation. Cells were collected during mid-exponential growth for activity measurements and during late-exponential growth for gas chromatography–mass spectrometry (GC–MS) analysis.

D. acetivorans A63 (DSM 5264), *D. multipotens* RH-8 (DSM 8415) and *D. propionica* U-8 (DSM 10410) were grown autotrophically in medium containing 6.2 mM NH₄Cl, 2.2 mM CaCl₂·2H₂O, 1.6 mM MgCl₂·6H₂O, 4.4 mM KCl, 2.4 mM KH₂PO₄, 14.3 mM MOPS, 10 g l⁻¹ sulfur powder, 1 mg l⁻¹ resazurin, 1 ml l⁻¹ SL-10 trace element solution (3 g l⁻¹ FeCl₂·4H₂O, 70 mg l⁻¹ ZnCl₂, 100 mg l⁻¹ MnCl₂·4H₂O, 4 mg l⁻¹ CuCl₂·2H₂O, 24 mg l⁻¹ NiCl₂·6H₂O, 36 mg l⁻¹ Na₂MoO₄·2H₂O, 30 mg l⁻¹ H₃BO₃, 224 mg l⁻¹ CoCl₂·6H₂O) and 1 ml l⁻¹ Wolfe's vitamin solution. The medium was prepared without sulfur and vitamin solution, made anaerobic by

bubbling with N₂ (100%) and reduced by the addition of Na₂S·9H₂O to a final concentration of 0.05% (w/v). The pH of the medium was adjusted to 7.3. The medium was dispensed anaerobically into serum bottles containing sulfur powder; the bottles were sealed with butyl rubber stoppers and aluminium caps and autoclaved for 40 min at 110 °C. Before inoculation, the vitamin solution was added and the gas phase was replaced with the required quantities of H₂:CO₂ (60:40, 80:20, 95:5, 98:2 or 99:1) at 1 bar overpressure. After the gas phase exchange, the pH was adjusted according to the optimum of each strain: pH 6.5–7.0 for *D. acetivorans*; pH 6.4–6.8 for *D. multipotens*; and pH 6.9–7.2 for *D. propionica*. Cultures were incubated at 55 °C while shaking at 130 rpm.

D. hydrogenophilus AcRS1 (DSM 3380) was grown autotrophically in DSMZ medium 195 containing 21.1 mM Na₂SO₄, 1.4 mM KH₂PO₄, 5.6 mM NH₄Cl, 359.3 mM NaCl, 15.3 mM MgCl₂·6H₂O, 6.7 mM KCl, 1.0 mM CaCl₂·2H₂O, 1 ml l⁻¹ selenite-tungstate solution out of DSMZ medium 385, 1 mg l⁻¹ resazurin, 1 ml l⁻¹ trace element solution SL-10 and 1 ml l⁻¹ Wolfe's vitamin solution. The medium was prepared without vitamin solution and made anaerobic by bubbling with N₂:CO₂ (80:20) to reach a pH below 6.0. The medium was dispensed anaerobically into serum bottles; the bottles were sealed with butyl-rubber stoppers and aluminium caps and autoclaved for 20 min at 121 °C. Before inoculation, the medium was reduced by the addition of Na₂S·9H₂O to a final concentration of 0.05% (w/v) and the vitamin solution was supplemented, resulting in a final pH of 7.1–7.4. The gas phase was replaced with the required quantities of H₂:CO₂ (60:40, 80:20, 95:5, 98:2 or 99:1) at 1 bar overpressure. After the gas phase exchange, the pH was adjusted to 7.1–7.4. Cultures were incubated at 30 °C while shaking at 130 rpm.

S. azorense AZ-Fu1 (DSM 15241) was grown autotrophically in DSMZ medium 1003 of the following composition: 28.4 mM MgSO₄·7H₂O, 12.6 mM Na₂S₂O₃, 2.7 mM CaCl₂·2H₂O, 6.4 mM KCl, 8.2 mM MgCl₂, 10 mM MES, 2 ml l⁻¹ solution A (1.6 M NH₄Cl, 0.9 M MgCl₂·6H₂O and 0.3 M CaCl₂·2H₂O), 1.5 ml l⁻¹ solution B (0.4 M K₂HPO₄), 10 ml l⁻¹ trace element solution (1.5 mM Na-EDTA·2H₂O, 0.6 mM CoCl₂·6H₂O, 0.5 mM MnCl₂·4H₂O, 0.4 mM FeSO₄·7H₂O, 0.7 mM ZnCl₂, 0.2 mM AlCl₃·6H₂O, 0.1 mM Na₂WO₄, 0.2 mM CuCl₂·2H₂O, 0.05 mM Ni₂SO₄·6H₂O, 0.07 mM Na₂SeO₃, 0.17 mM H₃BO₃ and 0.05 mM Na₂MoO₄·2H₂O). The pH was adjusted to 6.0. The medium was dispensed aerobically into serum bottles; the bottles were sealed with butyl-rubber stoppers and aluminium caps and sparged with CO₂ (100%) for 20 min, then autoclaved for 20 min at 121 °C. For the microaerophilic growth condition, before inoculation, the gas phase was replaced with 0.25 ml of O₂ per ml medium and H₂ at 1.38 bar overpressure. Cultures were incubated at 68 °C while shaking at 130 rpm.

E. coli strains (Top10, DH5α, Rosetta 2 (DE3) and C41) were grown at 37 °C in lysogeny broth medium. Antibiotics were added to the cultures to a final concentration of 100 µg ml⁻¹ ampicillin and 34 µg ml⁻¹ chloramphenicol.

Analysis of medium from *H. maritima*

To quantify the concentrations of amino acids in the medium during *H. maritima* growth, 250 µl of the medium at the respective time point was centrifuged (14,000 rpm, 4 °C, 15 min) and the supernatant was transferred to remove the elemental sulfur and bacterial cells. Samples were lyophilized after addition of 25 µl Norvaline solution (5 mM, aqueous) as an internal standard. The residue was hydrolysed in 6 M HCl (105 °C, 15 h) to account for free as well as peptide- and protein-bound amino acids. After hydrolysis, the sample was dried under a stream of nitrogen at 70 °C. The dried residue was derivatized with 50 µl anhydrous acetonitrile (ACN; Sigma-Aldrich) and 50 µl *N*-(*tert*-butyldimethylsilyl)-*N*-methyltrifluoroacetamide (MTBSTFA) with 1% *N*-*tert*-butyldimethylsilylchloride (TBDMS; Sigma-Aldrich) at 70 °C for 1 h. The TBDMS derivatives of amino acids were then analysed by GC–MS.

To quantify formate, acetate, lactate and succinate in the medium, 750 µl of the centrifuged medium at the respective time point was

lyophilized after addition of 25 µl Norvaline solution (5 mM, aqueous) as an internal standard. Then, 1 ml of methanol (VWR, HPLC-grade) was added to the dried residue and the mixture was treated for 15 min in an ultrasonic bath. After centrifugation (2 min, 7,000 rpm), the supernatant was dried under a stream of nitrogen at room temperature. The residue was derivatized with 50 µl anhydrous ACN (Sigma-Aldrich) and 50 µl MTBSTFA (Sigma-Aldrich) at 70 °C for 1 h. The TBDMS derivatives of acids were then analysed by GC-MS.

Preparation of cell extracts from *H. maritima*

H. maritima cultures (200 ml) were filtrated (Whatman Schleicher & Schuell, diameter 240 mm) under aerobic conditions to remove the elemental sulfur before centrifugation. The cells were collected during mid-exponential growth by centrifugation (8,000 rpm, 4 °C, 30 min). The supernatant was carefully transferred, and the cell pellet was resuspended in 20 ml of fresh medium. The resuspended cell suspension was centrifuged again (4,000 rpm, 4 °C, 30 min) and the supernatant was removed gently with the pipette. The cell pellet was washed with 0.9% saline and the cell solution was centrifuged again (4,000 rpm, 4 °C, 30 min). The cells were frozen in liquid nitrogen and stored at -80 °C. The cells were resuspended in 20 mM Tris-HCl pH 8, 5 mM dithioerythritol (DTE) and lysed by sonification aerobically with SONOPULS mini20 or anaerobically in the glove box with UW2070 (Bandelin Electronic) on ice (60% amplitude, 4 min, 1-s pulse, 2-s breaks; total energy input 2,000 kJ). Insoluble material was removed by centrifugation (14,000 rpm, 30 min, 4 °C). The protein concentration was determined using the Bradford method³³ with bovine serum albumin as a standard.

Protein hydrolysis

Isolation of protein bound amino acids was done as described previously³⁴. About 2 mg of bacterial sample (lyophilized cell pellet) was suspended in 500 µl of 6 M hydrochloric acid and hydrolysed overnight at 105 °C. The reaction mixture was dried under a stream of nitrogen. The residue was suspended in 200 µl of 50% acetic acid. Amino acids were isolated using a small column of Dowex 50W X8 (7 × 10 mm; 200–400 mesh, 34–74 µm, H⁺-form). The column was first washed with 2 ml H₂O, then amino acids were eluted with 1 ml 4 M aqueous ammonia solution. The ammonia eluate was dried under a stream of nitrogen at 70 °C. The dried residue was treated with 50 µl of MTBSTFA containing 1% TBDMS and 50 µl of anhydrous ACN at 70 °C for 30 min. The TBDMS derivatives of amino acids were then analysed by GC-MS. Furthermore, acid hydrolysis leads to the conversion of glutamine and asparagine to glutamate and aspartate, respectively. Therefore, results given for aspartate and glutamate correspond to asparagine or aspartate and glutamine or glutamate, respectively.

For the analysis of the composition of the amino acids, samples were treated as described above. To account for different derivatization and ionization efficiency of each amino acid, an equimolar amino acid mixture (2.5 µM in 0.1 M HCl, Sigma Aldrich) was used to determine the response factor for each amino acid. Therefore, 200 µl of the amino acid mixture was dried under a stream of nitrogen at 70 °C. The dried residue was treated with 50 µl MTBSTFA containing 1% TBDMS and 50 µl anhydrous ACN at 70 °C for 30 min.

GC-MS analysis

GC-MS analysis was performed with a QP2010 Plus gas chromatograph-mass spectrometer (Shimadzu) equipped with a fused silica capillary column (Equity TM-5; 30 m × 0.25 mm, 0.25-µm film thickness; SUPELCO) and a quadrupole detector working with electron impact ionization at 70 eV. An aliquot (0.1–6 µl) of the TBDMS-derivatized samples was injected in 1:5 split mode at an interface temperature of 260 °C and a helium inlet pressure of 70 kPa. For the analysis of ¹³C excess and isotopologue composition of bacterial amino acids, selected ion monitoring was used with a sampling rate of 0.5 s and

LABSOLUTION software (Shimadzu) was used for data collection and analysis. Isotopologue calculations were performed for m/z [M – 57]⁺ or m/z [M – 85]⁺. For analysis of the relative amino acid composition in *H. maritima* protein as well as the medium composition, measurements were performed in scan mode in a mass range from 45 m/z to 700 m/z with an injection volume of 0.1 µl.

For amino acids, the column was heated to 150 °C and kept at 150 °C for 3 min, after which was heated to 280 °C (7 °C per min) and held at that temperature for 3 min. For analysis of the TBDMS derivatives of formate, acetate, lactate and succinate, the column was heated to 60 °C and kept at 60 °C for 6 min. Then, the column was heated to 280 °C with a gradient of 10 °C min⁻¹ and kept at that temperature for another 3 min. The injector temperature was 260 °C. Measurements were performed in SCAN mode with a scan interval of 0.5 s and a mass range of 50–600 m/z . All samples were measured three times for technical replicates.

The calculation of ¹³C excess was done as described previously³⁴ and comprises (1) the detection of GC-MS spectra of unlabelled derivatized metabolites; (2) the determination of the absolute mass of isotopologue enrichments and distributions of labelled metabolites of the experiment; and (3) the correction of the absolute ¹³C incorporation by subtracting the contributions of the heavy isotopologues due to the natural abundances in the derivatized metabolites to calculate the enrichments and distributions of the isotopologues. For the ¹³C-labelling data analysis, Isotopo-4 software was used³⁵.

Cloning of *D. acetivorans* citrate synthase genes in *E. coli*

Standard protocols were used for the purification, cloning, transformation and amplification of DNA³⁶.

The gene encoding Desace_08345 was amplified by Gibson assembly³⁷ with Q5 High-Fidelity DNA Polymerase, using a forward primer (5'-GCCATCATCATCATCACAGCAGCG GCATGTCCTTT TTAAAGGAAAAATTAG-3') and a reverse primer (5'GCTTGTAG CAGCCGGATCTTACTTAATTCTGCCATTTC-3'). PCR conditions for the amplification of the gene of interest were as follows: 35 cycles of 10 s denaturation at 98 °C, 30 s primer annealing at 57.5 °C and 3 min elongation at 72 °C. The expression vector pET28b was amplified by PCR with Q5 High-Fidelity DNA Polymerase using a forward primer (5'-GATCCGGCTGCTAACAAAG-3') and a reverse primer (5'-GCCGCTGCTGTGATGATG-3'). PCR conditions were as follows: 35 cycles of 10 s denaturation at 98 °C, 30 s primer annealing at 64 °C and 3 min elongation at 72 °C. The reaction mixture contained 100 ng of the insert and 50 ng of the vector and it was incubated in a thermocycler at 50 °C for 15 min. The resulting pET28b vector containing the gene of interest with a N-terminal His₆-tag was transformed into *E. coli* Top10 for amplification, followed by purification and sequencing.

The gene encoding Desace_09325 was amplified by PCR with Q5 High-Fidelity DNA Polymerase using a forward primer (5'-GCGTAGAATTCCA TGAAGCTTAAAGAAAAC-3') introducing an EcoRI site (bold) and a reverse primer (5'-ATTATA**GTCGACT**AAAACCACACTTGTAGCT-3') introducing a Sall site (bold). PCR conditions were as follows: 35 cycles of 10 s denaturation at 98 °C, 30 s primer annealing at 66 °C and 3 min elongation at 72 °C. The PCR product was isolated and ligated into the expression vector pET28b containing a N-terminal His₆-Tag and transformed into *E. coli* TOP10 for amplification, followed by purification and sequencing.

The gene encoding Desace_06860 was amplified by PCR with Q5 High-Fidelity DNA Polymerase using a forward primer (5'-CGTAGCAT**ATG**TCAGATTGTTG-3') introducing a NdeI site (bold) and a reverse primer (5'-CAGAA**AG**CTTAACTCCTAGTTGTAG-3') introducing a HindIII site (bold). PCR conditions were as follows: 30 cycles of 10 s denaturation at 98 °C, 30 s primer annealing at 61 °C and 40 s elongation at 72 °C. The PCR product was isolated and cloned into pJET1.2 and transformed into *E. coli* Top10. After plasmid amplification, purification and sequencing, the PCR product was digested

Article

and ligated into the expression vector pET23b containing a C-terminal His₆-tag.

Heterologous expression of *D. acetivorans* citrate synthase genes in *E. coli*

The recombinant enzymes (Desace_08345, Desace_09325 and Desace_06860) were produced in *E. coli* Rosetta 2 (DE3) or *E. coli* C41 that had been transformed with the corresponding plasmids. The cells were grown at 37 °C in lysogeny broth medium supplemented with ampicillin and chloramphenicol. Expression was induced at an optical density (OD_{600 nm}) of 0.6–0.8 with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and the temperature was lowered to 20 °C. The cells were collected after overnight growth and stored at –20 °C until use.

Preparation of cell extracts

Frozen cells were suspended in triple volume of 20 mM Tris-HCl (pH 7.8), containing 500 mM NaCl, 20 mM imidazole and 0.1 mg ml^{−1} DNase I. The cell suspensions were lysed by a threefold passage through a chilled French pressure cell (137 MPa), and the cell lysates were centrifuged for 1 h (100,000g; 4 °C). The supernatant (cell extract) was immediately used for protein purification or enzyme assays.

Purification of recombinant proteins

The heterologously produced His-tagged citrate synthase Desace_08345, Desace_09325, Desace_06860 enzymes were purified using affinity chromatography. The corresponding cell extracts were applied at a flow rate of 0.5 ml min^{−1} to a 1-ml Protino Ni-NTA column (Macherey-Nagel) that had been equilibrated with 20 mM Tris-HCl (pH 7.8) containing 500 mM NaCl and 20 mM imidazole. The column was washed at a flow rate of 0.5 ml min^{−1} to elute unwanted protein with the same buffer containing 50 mM imidazole for Desace_08345; in two steps with the same buffer containing 35 mM and 70 mM imidazole for Desace_09325 and 50 mM and 70 mM imidazole for Desace_06860. The recombinant enzymes were eluted with the same buffer containing 500 mM imidazole. The enzymes were concentrated and the imidazole washed out using 10K Vivaspin Turbo 4 and stored at –20 °C with 50% glycerol. The identity of the purified recombinant proteins was confirmed using in-gel digestion by trypsin followed by liquid chromatography (LC)–MS/MS on a Synapt G2 Si instrument coupled to M-Class reversed-phase nanoUPLC (Waters).

Enzyme assays

Spectrophotometric enzyme assays (0.3 ml assay mixture) were performed in a glass (for visible light) cuvette at 55 °C for *H. maritima* cell extracts and *D. acetivorans* citrate synthase (Desace_08345) and at 42 °C or 55 °C for *S. azorensis* cell extracts, depending on the stability of the helping enzymes and substrates used in an assay. For anaerobic assays, cuvettes were sealed with rubber plugs, made anaerobic by gassing with N₂, and the anaerobic reaction mixture and substrates were added with Hamilton syringes. For LC analysis, reactions were stopped in an equal amount of stop solution (1 M HCl, 10% ACN). Proteins and insoluble particles were removed with three centrifugation steps (16,000g, 15 min, 4 °C). Samples were analysed with the Agilent 1290 Infinity II UHPLC system using a reversed-phase C18 column (Agilent InfinityLab Poroshell 120 EC-C18, 1.9 μm, 2.1 mm × 50 mm column). An ACN gradient from 2% to 8% in 10 mM potassium phosphate buffer (pH 7) with a flow rate of 0.55 ml min^{−1} was used. The retention times of succinyl-CoA, CoA and acetyl-CoA were 0.7, 0.8 and 1.6 min, respectively. The identification of the CoA esters was based on co-chromatography with standards and analysis of the ultraviolet-light spectra of the products.

Citrate synthase was measured spectrophotometrically at 412 nm as the oxaloacetate-dependent formation of free CoA from acetyl-CoA. The reaction mixture contained 100 mM Tris-HCl (pH 7.5), 5 mM oxaloacetate, 0.5 mM acetyl-CoA, 1 mM DTNB ($\epsilon_{412} = 14.2 \text{ mM}^{-1} \text{ cm}^{-1}$)³⁸ and cell extract. For citrate synthase characterization in *D. acetivorans*

(Desace_08345), Tris-HCl (pH 7.5) was substituted with MOPS (pH 7). The backward reaction was measured spectrophotometrically at 365 nm in combination with malate dehydrogenase as the citrate- and CoA-dependent oxidation of NADH ($\epsilon_{365} = 3.4 \text{ mM}^{-1} \text{ cm}^{-1}$)³⁹ and via UHPLC as the citrate-, CoA- and NADH-dependent formation of acetyl-CoA. For spectrophotometric assays, the assay mixture contained 100 mM Tris-HCl (pH 7.8), 5 mM DTE, 5 mM MgCl₂, 0.5 mM CoA, 20 mM citrate, 0.5 mM NADH and cell extract. For the characterization of Desace_08345, 20 U ml^{−1} porcine malate dehydrogenase (Sigma M1567) was added. For UHPLC assays, 1 mM CoA and 5 mM of NADH were used instead and 20 U ml^{−1} porcine malate dehydrogenase (Sigma M1567) was added. To stop the reaction, samples were mixed at each time point with an equal amount of 1 M HCl and 10% ACN. Proteins and insoluble particles were removed with three centrifugation steps (16,000g, 15 min, 4 °C).

ATP-citrate lyase and citrate lyase were measured at 42 °C spectrophotometrically at 365 nm by coupling to malate dehydrogenase, which reduces the produced oxaloacetate with NADH. The reaction mixture for ATP-citrate lyase was based on a previously published study⁴⁰ and contained 100 mM Tris-HCl (pH 7.5), 5 mM MgCl₂, 0.5 mM ATP, 0.5 mM CoA, 0.4 mM NADH, 20 U ml^{−1} porcine malate dehydrogenase (Sigma M1567) and 20 mM citrate. The reaction was started by the addition of ATP. The reaction mixture for citrate lyase contained 100 mM Tris-HCl (pH 7.5), 5 mM dithiothreitol (DTT), 5 mM MgCl₂, 0.5 mM NADH, 20 U ml^{−1} porcine malate dehydrogenase (Sigma M1567) and 20 mM citrate.

Isocitrate dehydrogenase was measured spectrophotometrically at 365 nm as the isocitrate-dependent reduction of NADP ($\epsilon_{365} = 3.5 \text{ mM}^{-1} \text{ cm}^{-1}$)³⁹. The assay mixture contained 100 mM Tris-HCl (pH 7.5), 5 mM DTE, 5 mM MgCl₂, 1 mM NADP, 10 mM DL-isocitrate and cell extract.

Aconitase was measured spectrophotometrically at 365 nm under anaerobic conditions in combination with endogenous isocitrate dehydrogenase as the citrate-dependent reduction of NADP. The reaction mixture contained 100 mM Tris-HCl (pH 8), 5 mM DTE, 5 mM MgCl₂, 1 mM NADP, 5 mM citrate and cell extract.

Malate dehydrogenase was measured spectrophotometrically at 365 nm as the oxaloacetate-dependent oxidation of NADH. The assay mixtures contained 100 mM Tris-HCl (pH 8), 5 mM DTE, 5 mM MgCl₂, 0.5 mM NADH, 2.5 mM oxaloacetate and cell extract.

Malic enzyme (malate dehydrogenase, decarboxylating) was measured spectrophotometrically through the malate-dependent reduction of NADP. The reaction mixture contained 100 mM Tris-HCl (pH 7.5), 5 mM MgCl₂, 5 mM DTE, 1 mM NADP, 30 mM malate and cell extract.

Fumarase was measured spectrophotometrically under anaerobic conditions at 240 nm as the cell-extract-dependent formation of malate from fumarate ($\epsilon_{240} = 2.4 \text{ mM}^{-1} \text{ cm}^{-1}$)⁴¹. The assay mixture contained 100 mM Tris-HCl (pH 7.5), 5 mM MgCl₂, 5 mM DTT, 0.4 mM fumarate and cell extract.

Pyruvate synthase and 2-oxoglutarate synthase activities were measured spectrophotometrically at 578 nm by following the pyruvate- or 2-oxoglutarate-dependent reduction of methyl viologen ($\epsilon_{578} = 9.7 \text{ mM}^{-1} \text{ cm}^{-1}$)⁴² in the case of pyruvate synthase and of benzyl viologen ($\epsilon_{578} = 8.65 \text{ mM}^{-1} \text{ cm}^{-1}$)⁴² in the case of 2-oxoglutarate synthase. The assay mixture contained 50 mM Tris-HCl (pH 7.5), 5 mM MgCl₂, 2.5 mM DTE, 0.5 mM CoA, 1 mM methyl or benzyl viologen and 10 mM pyruvate or 2-oxoglutarate. Anaerobic cuvettes and reaction mixture were flushed with N₂ gas before the measurement. The reactions were started by the addition of the substrates.

PEP carboxylase activity was measured spectrophotometrically in a coupled assay with endogenous malate dehydrogenase as the PEP-dependent oxidation of NADH. The reaction mixture contained 100 mM MOPS-KOH (pH 7.2), 4 mM MnCl₂, 5 mM DTE, 40 mM NaHCO₃, 0.5 mM NADH, 5 mM PEP and cell extract.

PEP carboxykinase activity was measured spectrophotometrically, as described above for PEP carboxylase, but the reaction was started with 0.5 mM ADP.

Pyruvate carboxylase activity was measured spectrophotometrically in a coupled assay with endogenous malate dehydrogenase as the pyruvate- and ATP-dependent oxidation of NADH. The reaction mixture contained 100 mM Tris-HCl (pH 7.5), 20 mM MgCl₂, 5 mM DTE, 15 mM NaHCO₃, 0.5 mM NADH, 0.2 mM acetyl-CoA, 40 mM pyruvate, 5 mM ATP (pH 7.0) and cell extract.

PEP synthase activity was measured spectrophotometrically backwards as phosphate-dependent formation of pyruvate from PEP and AMP by coupling it to the lactate dehydrogenase reaction. The reaction mixture (55 °C) contained 100 mM Na-K phosphate buffer (pH 7.0), 100 mM MgCl₂, 50 mM AMP (pH 6.8), 2 mM PEP, 0.5 mM NADH, 10 U lactate dehydrogenase (Sigma L2500) and cell extract. The reaction was started by adding the PEP and was AMP-dependent.

Succinyl-CoA synthetase was measured using UHPLC as the ATP-dependent formation of succinyl-CoA from succinate and CoA. The assay mixture contained 100 mM Tris-HCl (pH 7.5), 5 mM DTE, 5 mM MgCl₂, 10 mM succinate, 5 mM ATP, 1 mM CoA and cell extract.

Protein quantification using MS

Recombinantly expressed isoforms of citrate synthase (Desace_08345, Desace_09325, Desace_06860) were heterologously produced, characterized and used as standards for quantification and MS method development. To that end, gel bands were destained in 25 mM ammonium bicarbonate containing 50% methanol, washed with water and desolvated by adding ACN. The gel bands were dried in a speedvac before adding 100 mM DTT in 50 mM ammonium bicarbonate for 1 h reduction at room temperature. After removal of the solution, the gel piece was dried again and alkylated for 30 min in the dark by adding iodoacetamide. After the removal of the solution and drying of the gel piece, a DTT solution was added for quenching (15 min). The solution was removed, ACN added and the band was dried. For digestion, 30 µl trypsin (20 ng µl⁻¹) in 50 mM ammonium bicarbonate containing 10% ACN was added. Once the solution was taken up by the gel, the band was covered with the ammonium bicarbonate solution to prevent drying out. Digestion was performed at 37 °C overnight. Peptides were extracted with 1% formic acid solution containing 10%, 50% and 80% ACN. Extracts were pooled, dried and redissolved in 20 µl 0.1% formic acid containing 5% ACN. Pseudo-multiple reaction monitoring (MRM) experiments were set up for the isoforms individually on a Synapt G2 Si instrument coupled to a M-Class reversed-phase nanoUPLC (Waters; column 1.8 µm HSS T3: 75 µm × 150 mm; trap column M-class Trap Symmetry C18: 5 µm; 180 µm × 20 mm, 100 Å; 300 nL min⁻¹; solvents: A, 0.1% aqueous formic acid; B, 0.1% formic acid in ACN) using the most abundant unique tryptic peptides (as chosen by Skyline (Extended Data Fig. 3a)). Calibration curves were generated by injecting 0.1, 0.2, 0.3, 0.4 and 0.5 µl (Extended Data Fig. 3b).

To quantify the corresponding proteins in *D. acetivorans* cells, the proteins were separated using one-dimensional polyacrylamide gel electrophoresis (PAGE). Analyses were subsequently performed on the isoform bands that had been separated using gel electrophoresis in triplicate. In addition, the relative quantification of the cell lysate proteomes of *D. acetivorans* and *H. maritima* was achieved with data-independent label-free high-definition (HD)MS protein expression analysis on Synapt G2 Si⁴³ following filter-based tryptic digestion (1 µg on-column) using the UniProt *D. acetivorans* and *H. maritima* databases as previously described⁴⁴. In brief, cell lysates were reduced, alkylated and tryptically digested on 10 kDa centrifugal filter devices and prepared for LC-MS at 250 ng µl⁻¹ in 0.1% formic acid and 5% ACN. This method measures the entire proteomes of replicate samples in single experiments. It is not as accurate as the MRM method, but it still verifies the concentration range of the three isoforms (Supplementary Table 10). For statistical analyses, Progenesis QIP software was used (nonlinear diagnostics/Waters, fixed modification carbamidomethylation, variable modification methionine oxidation, 1 missed cleavage allowed).

Bioinformatics

Query sequences for the database searches were obtained from the National Center for Biotechnology Information (NCBI) database. The BLAST searches were performed using the NCBI BLAST server (<http://www.ncbi.nlm.nih.gov/BLAST/>).

INCA¹⁰ was used for codon usage analysis for various bacteria using each ribosomal proteins as a reference set. On the basis of a previously published study⁹, the method Measure Independent of Length and Composition (MILC) was chosen for the x (to the reference set) and y axes (to all genes) and its derivative method MILC-based Expression Level Predictor (MELP) was used as a filter to quantitatively predict gene expression levels from genomic data. The genes were listed from the highest MELP value to the lowest and genes shorter than 100 codons were excluded from the analysis.

Other methods

The concentration of hydrogen sulfide in *H. maritima* cultures was measured by a previously published colorimetric method⁴⁵. Samples were taken after inoculation and every 24 h until the stationary phase was reached. SDS-PAGE (12.5%) was performed as described by previously⁴⁶. An unstained protein marker (Pierce Unstained Protein MW Marker, 14.4–116 kDa, Thermo Scientific) was used as a molecular mass standard. Apparent *K_m* and *V_{max}* values and growth curves were calculated using GraphPad Prism5 software. Circles were drawn with ChemDraw Professional. DNA sequence determination was performed by Eurofins. Standard deviations and standard errors of mean were calculated using Microsoft Excel.

Reporting summary

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

Data availability

All data generated in this manuscript are included within the paper (and its Supplementary Information). The raw data are presented in the manuscript and/or available from the corresponding authors upon reasonable request. For any further inquiries about our work please contact the corresponding authors. Source data are provided with this paper.

32. Simon, E. J. & Shemin, D. The preparation of S-succinyl coenzyme A. *J. Am. Chem. Soc.* **75**, 2520 (1953).
33. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254 (1976).
34. Eyerlert, E. et al. Carbon metabolism of *Listeria monocytogenes* growing inside macrophages. *Mol. Microbiol.* **69**, 1008–1017 (2008).
35. Ahmed, Z. et al. 'Isotopo' a database application for facile analysis and management of mass isotopomer data. *Database* **2014**, bau077 (2014).
36. Ausubel, F. M. et al. *Current Protocols in Molecular Biology* (Wiley, 1987).
37. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
38. Riddles, P. W., Blakeley, R. L. & Zerner, B. Ellman's reagent: 5,5'-dithiobis(2-nitrobenzoic acid)—a reexamination. *Anal. Biochem.* **94**, 75–81 (1979).
39. Bergmeyer, H. U. Neue Werte für die molaren Extinktions-Koeffizienten von NADH und NADPH zum Gebrauch im Routine-Laboratorium. *Z. Klin. Chem. Klin. Biochem.* **13**, 507–508 (1975).
40. Hügler, M., Huber, H., Molyneaux, S. J., Vetriani, C. & Sievert, S. M. Autotrophic CO₂ fixation via the reductive tricarboxylic acid cycle in different lineages within the phylum Aquificae: evidence for two ways of citrate cleavage. *Environ. Microbiol.* **9**, 81–92 (2007).
41. Flint, D. H. Initial kinetic and mechanistic characterization of *Escherichia coli* fumarase A. *Arch. Biochem. Biophys.* **311**, 509–516 (1994).
42. Dawson, R. M. C., Elliott, D. C., Elliott, W. H. & Jones, K. M. *Data for Biochemical Research* (Clarendon, 1986).
43. Distler, U., Kuharev, J., Navarro, P. & Tenzer, S. Label-free quantification in ion mobility-enhanced data-independent acquisition proteomics. *Nat. Protoc.* **11**, 795–812 (2016).
44. Wildschütz, L. et al. Transcriptomic and proteomic analysis of iris tissue and aqueous humor in juvenile idiopathic arthritis-associated uveitis. *J. Autoimmun.* **100**, 75–83 (2019).

Article

45. Trüper, H. G. & Schlegel, H. G. Sulphur metabolism in Thiorhodaceae. I. Quantitative measurements on growing cells of *Chromatium okenii*. *Antonie van Leeuwenhoek* **30**, 225–238 (1964).
46. Laemmli, U. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685 (1970).
47. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
48. Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
49. Noor, E. et al. Pathway thermodynamics highlights kinetic obstacles in central metabolism. *PLOS Comput. Biol.* **10**, e1003483 (2014).
50. Witt, A., Pozzi, R., Diesch, S., Hädicke, O. & Grammel, H. New light on ancient enzymes – in vitro CO₂ fixation by pyruvate synthase of *Desulfovibrio africanus* and *Sulfolobus acidocaldarius*. *FEBS J.* **286**, 4494–4508 (2019).
51. Furdui, C. & Ragsdale, S. W. The role of pyruvate ferredoxin oxidoreductase in pyruvate synthesis during autotrophic growth by the Wood-Ljungdahl pathway. *J. Biol. Chem.* **275**, 28494–28499 (2000).
52. Flamholz, A., Noor, E., Bar-Even, A. & Milo, R. eQuilibrator—the biochemical thermodynamics calculator. *Nucleic Acids Res.* **40**, D770–D775 (2012).

Acknowledgements We thank G. Fuchs for discussions during this work and for critical reading of the manuscript, W. Schulz for his help with cloning of citrate synthase genes, A. M.

Berg and D. Ackermann for technical assistance and A. Probst for the help with bioinformatics analysis. This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project-ID BE 4822/5-1 to I.A.B. and Project-ID 364653263-TRR 235 to W.E. as well as the Hans-Fischer Gesellschaft (Munich) to W.E.

Author contributions L.S. and E.P. performed growth experiments, cloning, purification and characterization of enzymes and enzyme assays. L.S performed codon usage analysis and sulfide determination. T.M.S. performed isotopologue profiling experiments and GC-MS analysis of *H. maritima* medium. S.K. conducted proteomics analyses. A.M. performed phylogenetic analysis and bioenergetics calculations. I.A.B. and W.E. wrote the manuscript with input from other authors. L.S., E.P., T.M.S., A.M., S.K. and W.E. prepared figures. The manuscript was reviewed and approved by all coauthors.

Competing interests The authors declare no competing interests.

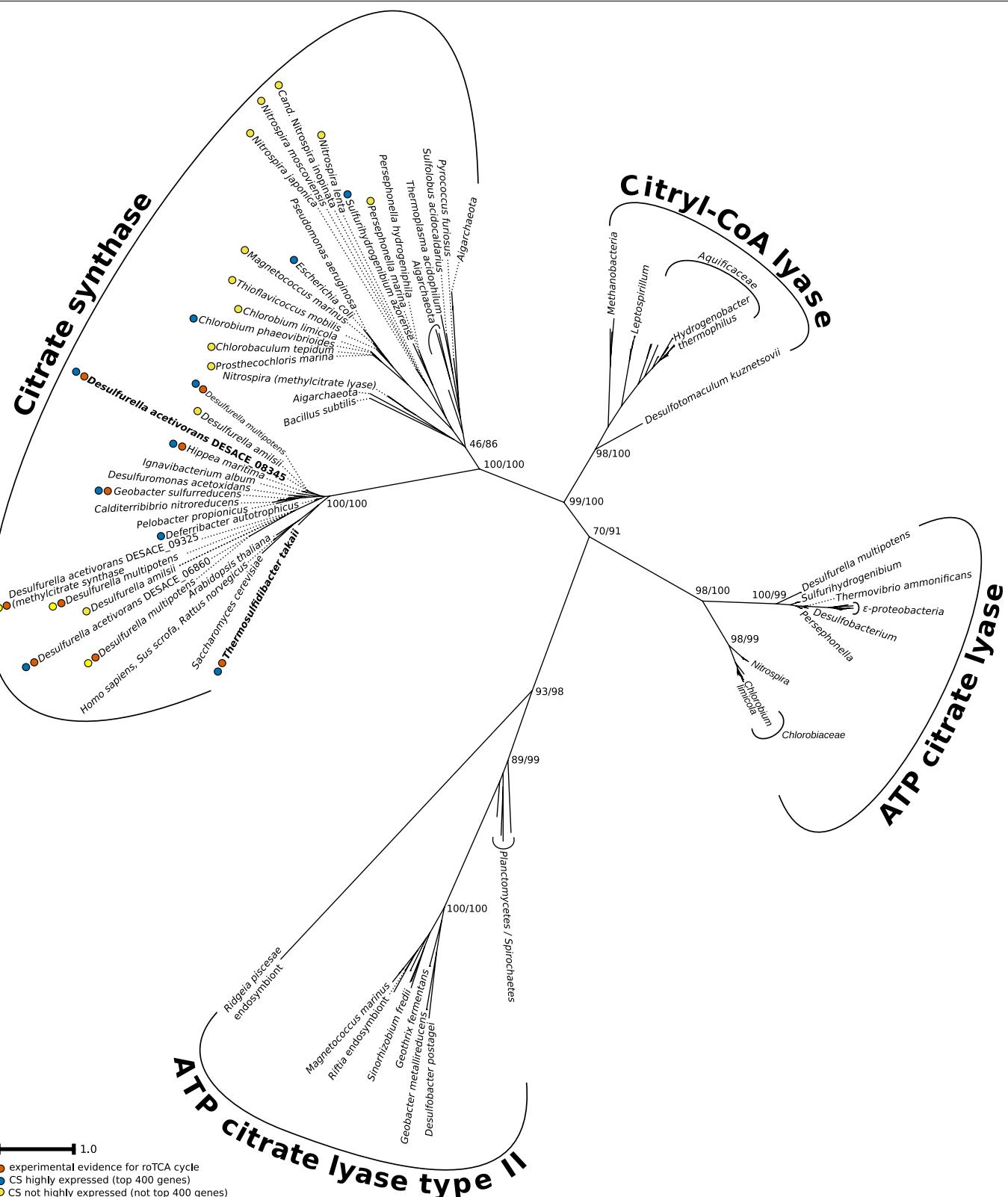
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-021-03456-9>.

Correspondence and requests for materials should be addressed to W.E. or I.A.B.

Peer review information *Nature* thanks William Martin and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at <http://www.nature.com/reprints>.

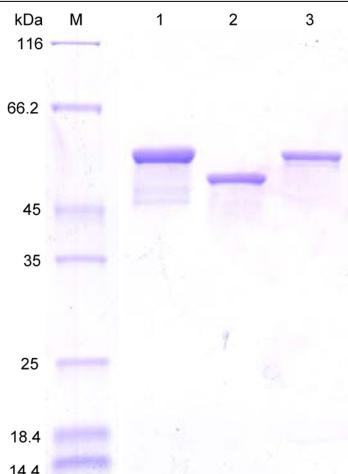


Extended Data Fig. 1 | Phylogenetic tree of citrate-cleaving enzymes.

Sequences were aligned using MUSCLE⁴⁷, the phylogeny was calculated with IQtree⁴⁸ using the maximum likelihood method (LG + R6 substitution model). The scale bar represents amino acid substitutions per site. Branch labels are

SH-aLRT support/ultrafast bootstrap support (%) values from 1,000 replications. The sequences used to build the tree are available in Supplementary Table 11.

Article



Extended Data Fig. 2 | SDS-PAGE (12.5%) of recombinant citrate synthases from *D. acetivorans* after purification with Ni-NTA column. M, molecular mass standard proteins; lane 1, Desace_08345 (2 µg); lane 2, Desace_06860 (2 µg); lane 3, Desace_09325 (2 µg). Their predicted molecular masses are 49 kDa (Desace_08345) and 50 kDa (Desace_06860; Desace_09325). Proteins were stained with Coomassie blue. For gel source data, see Supplementary Fig. 1. The SDS-PAGE analysis was conducted three times and all results were similar to the ones shown here.

a

Desace_06 | KIGEIKISQIINGMRGLK**VLFTDLSSVDPEKGIIWFGYTVDDEVLQKLKP**PPLSKMPYVEA
 Desace_08 | KIDEVTISQAIGGMRGIK**SLVTDISYLDPEEGIRFFGYTIPLEVLEKLKP**VPGAEMPYVEG
 Desace_09 | KVGDITIAQVIGGMRGLK**VLVTDISYLDPFEGR**YRGYTDEVLQKLKPKGAEAMPYDEA

Desace_06 | QFYLLITGDI PSETEVRE**EIIDIFNER**RKLPDYVNILKTMPKETDPNIMLSVAINSMQHE
 Desace_08 | HFYLLLTGDPVPTKEVK**EVAEEFKKRR****ALPEYVKDTLK**KAMPRDTHPMTMFAAGILAMQRE
 Desace_09 | QFYLLMTGDI PTEQEVIIDLKEKR**KVPNYYVK**VLD SMP TQARP DVMLAVAVD TMQQE

Desace_06 | SVFAKAYQKGKVTKYNAWEYMLVEDVLNLLPKISIISAFIYRLKYKNNMFIKE DDSDL DLGG
 Desace_08 | SKFAAYYNAGKFNKNTAWEPMFEDAMDLMAK**LPSLGLAYIYR**MKYKSDTHIPS NPD L D L G G
 Desace_09 | SVFAK**AYAENKITK**QNAWEYMLVEDVLNLLPKIPMIAAYIYRLKYKNNNQI PENP D L D F G G

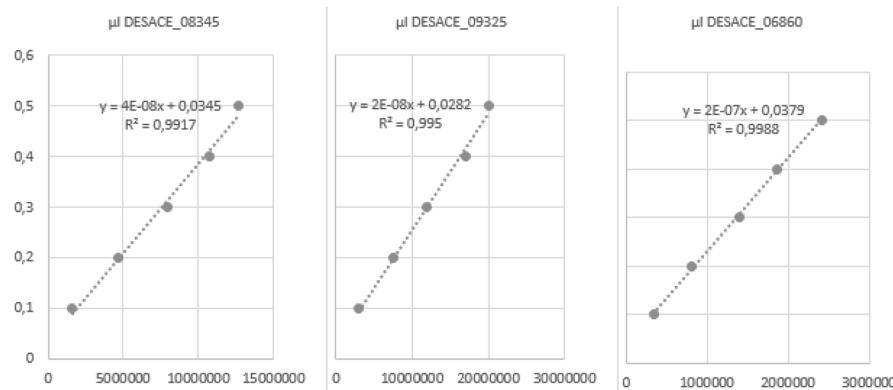
Desace_06 | NLAYMMGIKQPYDEFRLYMI L SDHENGNVSAHTTHLVASALADAYCSISAGINGLSP
 Desace_08 | DFANMMGIDKPYDDVARLYFILHDHESGNVSAHTAHLVASALS DAY YAYSAAMCGIAGP
 Desace_09 | NFAHMMGIDKPYDFSRLYFILLSDHESGNVSAHTHLVASAWADAYSLAAGINGLSP

Desace_06 | LHGGAIRESFNWMNLYQQK--KNQLTKKILEQICVDMNLNK**GLVIPGYGHAVLR**NKD PRF
 Desace_08 | LHGLANQEVLK**WQETIDKK**LGGKVPTKEELKKFVEETLSSGQVIPGYGHAVLRKTDPY
 Desace_09 | LHGGATQEALKWFQELLKKL--NKIPTKEELEQFCWDTLNSGQVIPGYGHAVLRKTDPF**E**

Desace_06 | LAQMKGEEYLKND E LFKIAR**LAFDTIPHILK**KHTKIKNPWPVN DNI SGI IHSHYGI DK--
 Desace_08 | VAQREFALKHMPDDIFQVVSM LYEVVPPILSSLGKVDPWPVN DAHSGC IQWHYGV VEY
 Desace_09 | **VAQLEFGKK**HLPDDKL FQQLVSLLYEVAPDVLT KHGAKNPWPNC DN ITGT IKAHYGVNQY

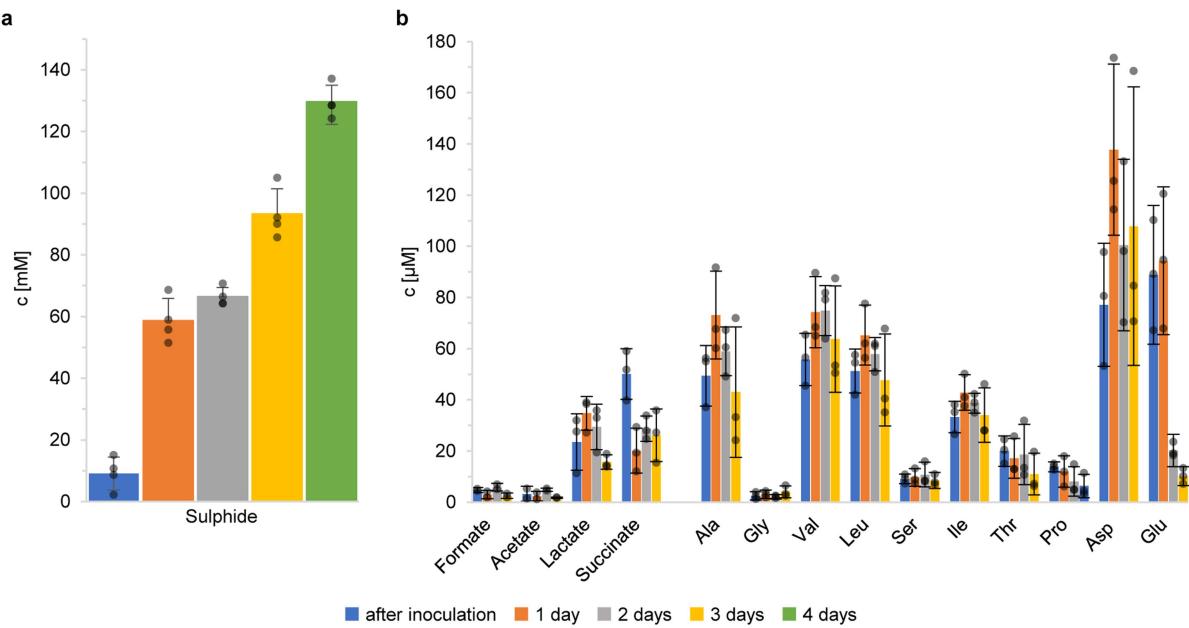
Desace_06 | NFFT VLF SNS IAVG SLAN IF WDR**A1NLPIERPK**SVT ISQL FEL--NKTRI KLA AALEHHH
 Desace_08 | DFYTVLFGIGR**ALGVLANLVWDRALGYAIE RPK**S VTT DMLEK-----WAGIK---
 Desace_09 | EFYPVLF DISRSMGVLSNITWDR**ALGYPIERPK**S VT LKM LESLVQKDTTA KVWF-----

b



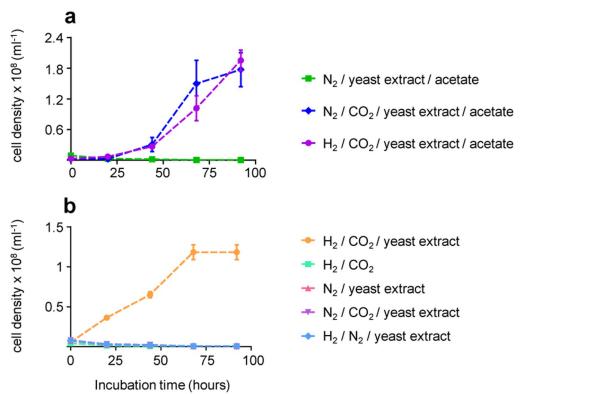
Extended Data Fig. 3 | Protein quantification using MS. a, CLUSTAL format alignment by MAFFT (v.7.452) of sequence parts of citrate synthase isoforms to illustrate the unique peptides chosen for quantification. Desace_06, Desace_06860; Desace_08, Desace_08345; Desace_09, Desace_09325.

b, Calibration curves generated for each citrate synthase using tryptic digests of the recombinant proteins isolated by one-dimensional PAGE. Injection volume of standard solution (μl , yaxis) versus MS peak area in MRM experiments (xaxis) as calculated by Skyline.



Extended Data Fig. 4 | Metabolites in *H. maritima* growth medium. **a**, Sulfide production during growth under mixotrophic conditions (CO_2 , H_2 , S^0 and 0.2 g l^{-1} yeast extract). **b**, Concentrations of potential fermentation products and some amino acids in the medium after inoculation and after 1, 2 and 3 days

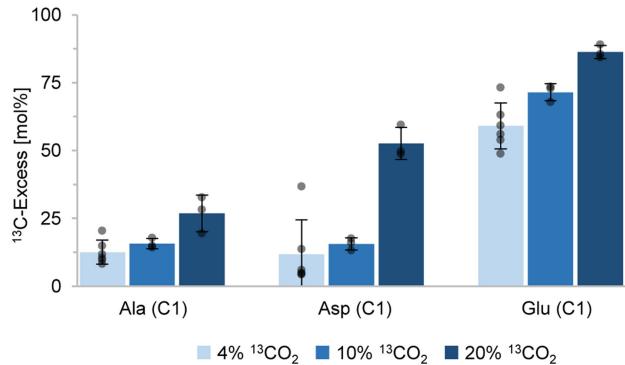
of cultivation. Data are mean \pm s.e.m. of four biological replicates for sulfide determination and of three biological and three technical replicates for the potential fermentation products and amino acids in the medium.



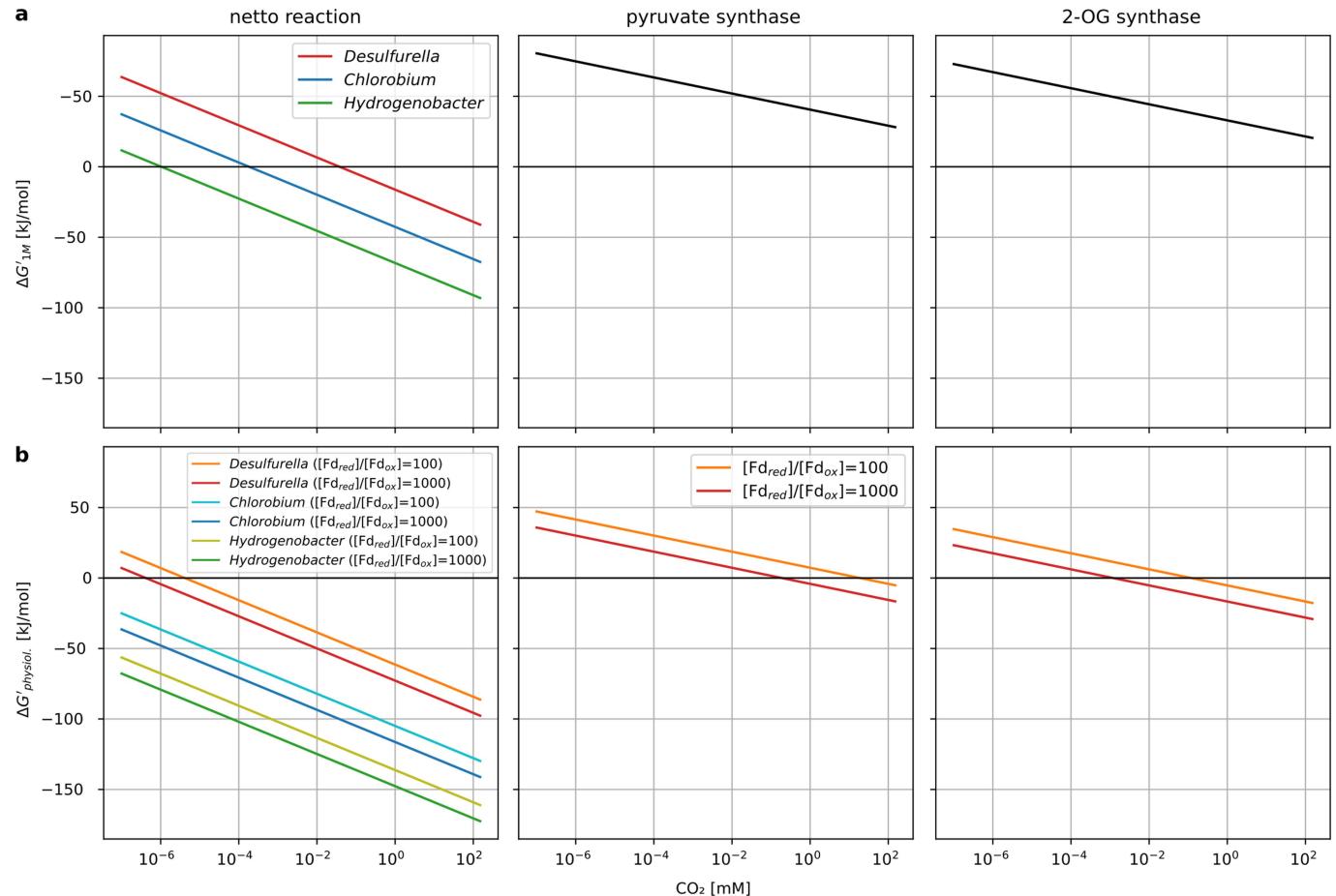
Extended Data Fig. 5 | Growth of *H. maritima* with different substrates.

a, Growth of *H. maritima* in 80% H_2 , 20% CO_2 , 0.2 g l⁻¹ yeast extract and 5 g l⁻¹ acetate, in 80% N_2 , 20% CO_2 , 0.2 g l⁻¹ yeast extract and 5 g l⁻¹ acetate, and in 100% N_2 , 0.2 g l⁻¹ yeast extract and 5 g l⁻¹ acetate. **b**, Growth of *H. maritima* in 80% H_2 , 20% N_2 and 0.2 g l⁻¹ yeast extract, in 80% N_2 , 20% CO_2 and 0.2 g l⁻¹ yeast extract, in 100% N_2 and 0.2 g l⁻¹ yeast extract, in 80% H_2 and 20% CO_2 , and in 80% H_2 , 20% CO_2 and 0.2 g l⁻¹ yeast extract. For all growth curves, data are presented as mean \pm s.e.m. of three biological replicates.

Article



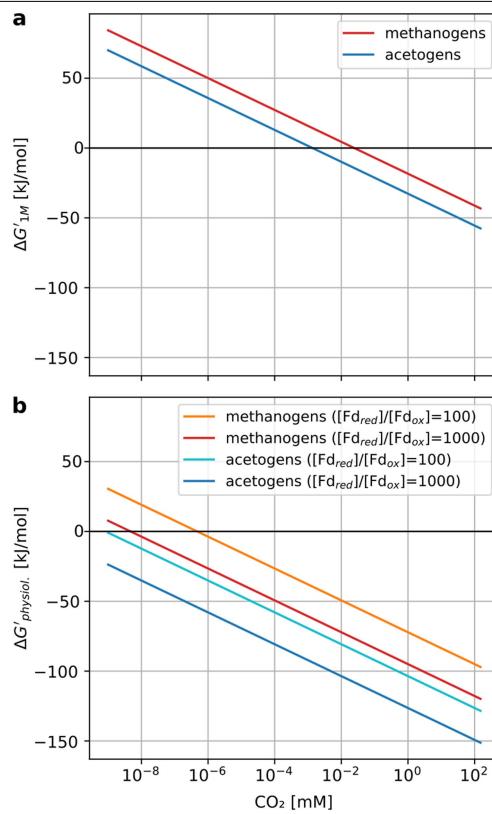
Extended Data Fig. 6 | The labelling in C1 of Ala, Asp and Glu after growth in H₂, S⁰, 0.2 g l⁻¹ yeast extract and 4, 10 or 20% ¹³CO₂. The labelling was calculated from the isotopologue composition of the corresponding amino acids and their fragments after the loss of the carboxylic atom (Supplementary Tables 4–6). Data are mean ± s.e.m. of six biological replicates for 4% ¹³CO₂ and of three biological replicates for 10% and 20% ¹³CO₂.



Extended Data Fig. 7 | Dependency of energetic efficiencies of the rTCA cycle variants, pyruvate synthase and 2-oxoglutarate synthase reactions on CO_2 concentration. Net equations are as follows². rTCA cycle (citrate synthase, as in Desulfurellaceae; red and orange lines): $2 \text{CO}_2(\text{total}) + \text{NADPH} + 2 \text{NADH} + 2 \text{Fd}_{\text{red}} + \text{CoA} + \text{ATP} \leftrightarrow \text{acetyl-CoA} + \text{NADP}^+ + 2 \text{NAD}^+ + 2 \text{Fd}_{\text{ox}} + \text{ADP} + \text{P}_i + 4 \text{H}_2\text{O}$. rTCA cycle in *Chlorobium* (ACL; dark and light blue lines): $2 \text{CO}_2(\text{total}) + \text{NADPH} + 2 \text{NADH} + 2 \text{Fd}_{\text{red}} + \text{CoA} + 2 \text{ATP} \leftrightarrow \text{acetyl-CoA} + \text{NADP}^+ + 2 \text{NAD}^+ + 2 \text{Fd}_{\text{ox}} + 2 \text{ADP} + 2 \text{P}_i + 3 \text{H}_2\text{O}$. rTCA cycle in *Hydrogenobacter thermophilus* (citryl-CoA synthase/citryl-CoA lyase and ATP-dependent 2-oxoglutarate carboxylase; green and olive lines): $2 \text{CO}_2(\text{total}) + 3 \text{NADH} + 2 \text{Fd}_{\text{red}} + \text{CoA} + 3 \text{ATP} \leftrightarrow \text{acetyl-CoA} + 3 \text{NAD}^+ + 2 \text{Fd}_{\text{ox}} + 3 \text{ADP} + 3 \text{P}_i + 2 \text{H}_2\text{O}$. Fumarate reductase is assumed to be NADH-dependent². For ferredoxin, a reduction potential E° of -418 mV,

corresponding to the hydrogen/proton couple, was assumed. Temperature is assumed 25 °C, pH 7, ionic strength 0.1 M. CO_2 is the sum of $\text{CO}_2(\text{aq})$ and its hydrated forms (H_2CO_3 , HCO_3^- and CO_3^{2-}). Under the *D. acetivorans* growth conditions (55 °C, pH 7, 2 bar pressure, 30.15 g l⁻¹ salinity), a CO_2 partial pressure of 40% corresponds to 147.6 mM CO_2 , a partial pressure of 1% to 3.7 mM.

a, Standard conditions (1 M concentrations for all reactants except CO_2). **b**, Assumed physiological metabolite concentrations (0.1 mM NADPH, 0.01 mM NADP⁺, 0.1 mM NADH, 1 mM NAD⁺, 5 mM ATP, 0.5 mM ADP, 10 mM phosphate⁴⁹, 1.07 mM CoA, 0.013 mM acetyl-CoA², 0.1 mM Fd_{red} , 0.001/0.0001 mM Fd_{ox} (refs. ^{3,50}), 0.23 mM succinyl-CoA, 0.44 mM 2-oxoglutarate¹³ and 0.18 mM pyruvate⁵¹). The calculations were done using eQuilibrator⁵².



Extended Data Fig. 8 | Dependency of energetic efficiency of the reductive acetyl-CoA pathway in methanogens and acetogens on CO₂ concentration.

Net equation of the reductive acetyl-CoA pathway in methanogens and acetogens are as follows². Methanogens (*Methanothermobacter marburgiensis*; red and orange lines): $2 \text{CO}_2(\text{total}) + 4 \text{Fd}_{\text{red}} + \text{CoA} + 2 \text{F}_{420}\text{H}_2 \leftrightarrow \text{acetyl-CoA} + 4 \text{Fd}_{\text{ox}} + 2 \text{F}_{420} + 5 \text{H}_2\text{O}$. Acetogens (*Acetobacterium woodii*; dark and light blue lines): $2 \text{CO}_2(\text{total}) + \text{NADPH} + \text{NADH} + 4 \text{Fd}_{\text{red}} + \text{CoA} + \text{ATP} \leftrightarrow \text{acetyl-CoA} + \text{NADP}^+ + \text{NAD}^+ + 4 \text{Fd}_{\text{ox}} + \text{ADP} + \text{P}_i + 4 \text{H}_2\text{O}$. A free energy change ($\Delta G'$) of acetogenesis and methanogenesis is usually not sufficient to drive the synthesis of 1 mol ATP per mol of product under physiological conditions, and energy generation requires chemiosmotic coupling and entails flavin-based electron bifurcation^{30,31}. However, a high CO₂ partial pressure may enable the functioning of a hybrid reductive acetyl-CoA pathway that combines CO₂ reduction with ferredoxin, as in methanogens, and acetate formation from acetyl-CoA coupled with substrate phosphorylation, as in acetogens. Calculations were done as in Extended Data Fig. 7. **a**, Standard conditions (1M concentrations for all reactants except CO₂). **b**, Assumed physiological metabolite concentrations (0.1 mM NADPH, 0.01 mM NADP⁺, 0.1 mM NADH, 1 mM NAD⁺, 5 mM ATP, 0.5 mM ADP, 0.1 mM reduced coenzyme F₄₂₀, 0.1 mM oxidized coenzyme F₄₂₀, 10 mM phosphate⁴⁹, 0.28 mM CoA, 0.0104 mM acetyl-CoA⁵¹, 0.1 mM Fd_{red}, 0.001/0.0001 mM Fd_{ox} (refs. ^{3,50}).

Extended Data Table 1 | Catalytic properties of *D. acetivorans* citrate synthase Desace_08345

Substrate	Desace_08345		
	V_{max} (U mg ⁻¹ protein)	K_m (μM)	K_{cat}/K_m (s ⁻¹ μM ⁻¹)
Acetyl-CoA	157 ± 4	28 ± 3	4.5
Oxaloacetate	131 ± 3	48 ± 7	2.2
Coenzyme A	0.7 ± 0.03	240 ± 44	2.4 · 10 ⁻³
Citrate	1 ± 0.14	11,060 ± 3,616	7.4 · 10 ⁻⁵

Data are mean ± s.d. of three technical replicates.

Article

Extended Data Table 2 | Enzymes of carbon metabolism in cell extracts of autotrophically grown *S. azorense*

Enzyme	Specific activity [$\mu\text{mol min}^{-1} \text{mg}^{-1}$ protein] \pm standard deviation		Candidate gene(s), GenBank Acc.
	Measured at *42/ †55 °C	Extrapolated to 68°C	
Citrate synthase (measured as citrate synthesis)	0.01 \pm 0.006 (n=3) [†]	0.03 \pm 0.02	WP_012674633.1
ATP-citrate lyase	0.04 \pm 0.01 (n=3)*	0.26 \pm 0.06	WP_012673967.1 (ACL), WP_012673604.1 (ACL), WP_012674203.1 (CCL)
Isocitrate dehydrogenase	0.03 \pm 0.01 (n=3) [†]	0.09 \pm 0.04	WP_012674929.1
Malate dehydrogenase	0.64 \pm 0.04 (n=3) [†]	1.76 \pm 0.11	WP_012674698.1

Activities were measured at 42°C or 55°C. The activities were extrapolated to 68°C based on the assumption that a 10°C increase in temperature doubles the reaction rate. The number of biological repetitions (n) is shown. For each biological replication, at least two technical replications were carried out. CCL, citryl-CoA lyase.

*Reaction measured at 42°C.

[†]Reaction measured at 55°C.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

All bacterial strains with exception of E. coli were obtained from Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ).

The following software was used for data collection:

- for spectrophotometric enzyme assays and sulphide quantification: Agilent Cary UV Workstation Version 1.0.1284;
- for enzyme assays with UHPLC: Agilent OpenLAB CDS ChemStation Edition C.01.07 SR3 [465];
- GC-MS: LabSolution 4.20 Software (Shimadzu).

Data analysis

Standard deviations and standard errors of mean were calculated using Microsoft Excel 2013, 2016. Apparent Km and Vmax values and growth curves were calculated using GraphPad Prism5 software. The software Interactive Codon Analysis INCA2.1 was used for codon usage analysis. For protein quantification, Excel 2016, Skyline 4.1.0.11796 and Progenesis QIP v. 4.2 software was used (nonlinear diagnostics/ Waters Corp., Manchester, UK). For GC/MS analysis, LABSOLUTION 4.20 software (Shimadzu) was used. 13C labelling data was analyzed using Isotopo-4 software (<http://spp1316.uni-wuerzburg.de/bioinformatics/isotopo/>). Sequences were aligned using MUSCLE v3.8.1551, the phylogeny was calculated with IQ-TREE multicore version 2.0.3 for Linus 64-bit built Apr 26 2020.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The following databases were used: NCBI database (<https://www.ncbi.nlm.nih.gov/>), Uniprot (<https://www.uniprot.org/>).

The following Figures have associated raw data: Figure 3, Extended Data Figures 4 and 5.

All data generated in this manuscript are included within the paper (and its Supplementary Information files). The raw data are presented in the manuscript and/or available from the corresponding authors upon reasonable request. For any further inquiries about our work please contact the corresponding authors.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes were chosen based on convention in the field and are sufficient given the robust signal changes measured in the experiments. All labelling studies in hydrolysates of <i>H. maritima</i> were performed thrice with biological replicates (and in six replication for the 4% CO ₂ culture). All enzyme assays in cell extracts were performed several times to test cell extract and substrate dependencies, and at least twice (in most cases thrice) in biological replicates, or more often when the deviation in the activities was >10%. Overall, the obtained replicates were in good agreement with each other, which is why we are confident that this sample size sufficiently accounts for biological variance as well as errors of measurement. Enzyme assays performed with purified enzymes as well as MRM-experiments were carried out in three technical replicates to account for errors of measurement.
Data exclusions	No data were excluded from the analysis.
Replication	See point 1. All attempts at replication of our findings were successful.
Randomization	Randomization was not necessary because the subjects of our experiments (proteins or bacteria) were studied in large numbers.
Blinding	Blinding was not performed as the subjects of our experiments were either proteins or bacteria that were studied in large numbers.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging