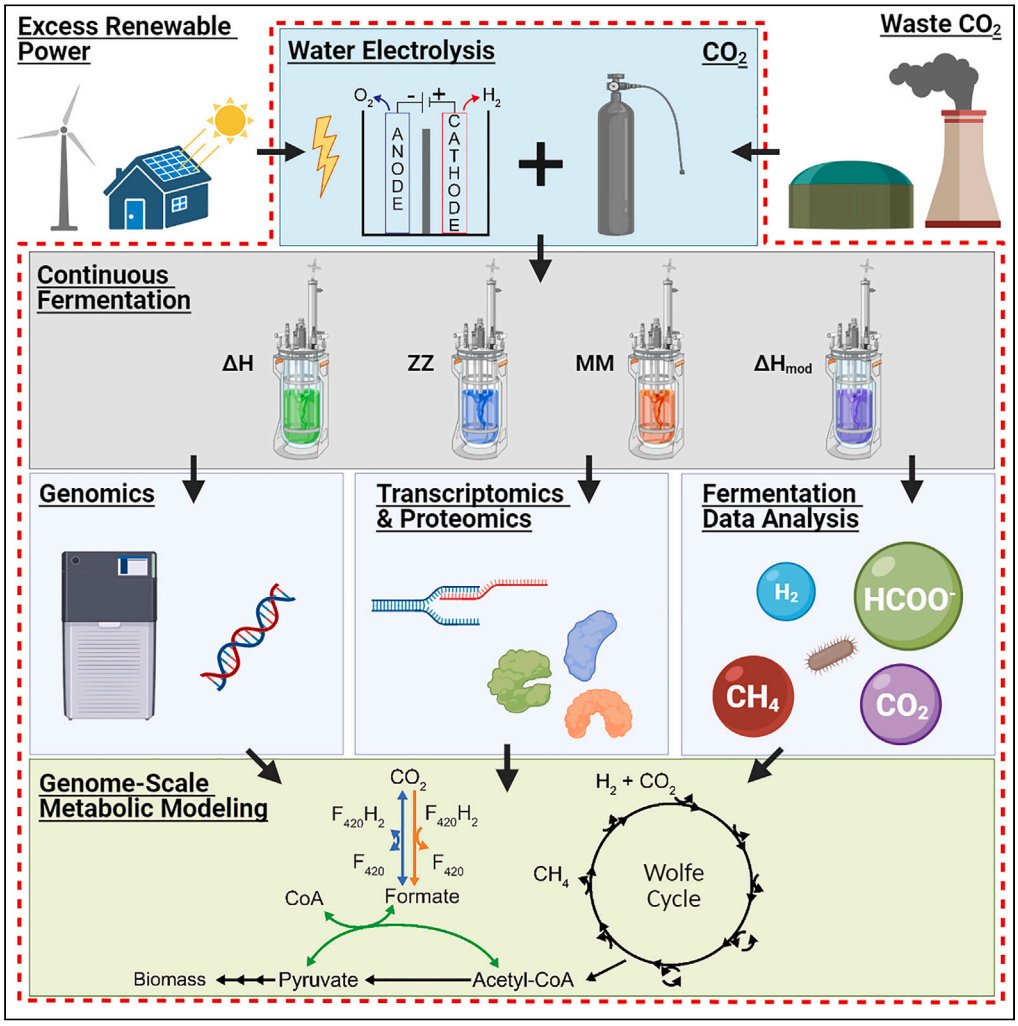
iScience

CellPress
OPEN ACCESS

Article

An integrated systems biology approach reveals differences in formate metabolism in the genus *Methanothermobacter*



Isabella Casini, Tim McCubbin, Sofia Esquivel-Elizondo, ..., Esteban Marcellin, Largus T. Angenent, Bastian Molitor

bastian.molitor@uni-tuebingen.de

Highlights

Curated genome-scale reconstructions for thermophilic methanogens are presented

Replicate chemostat bioreactors result in reproducible systems-biology datasets

Genetically engineered strain has stable performance in chemostat bioreactors

Formate anabolism is hypothesized to influence carbon distribution between products

Casini et al., iScience 26, 108016
October 20, 2023
https://doi.org/10.1016/j.isci.2023.108016

iScience

CellPress
OPEN ACCESS

Article

An integrated systems biology approach reveals differences in formate metabolism in the genus *Methanothermobacter*

Isabella Casini,[1] Tim McCubbin,[2,3,4] Sofia Esquivel-Elizondo,[5] Guillermo G. Luque,[5] Daria Evseeva,[6,7] Christian Fink,[1] Sebastian Beblawy,[1] Nicholas D. Youngblut,[5] Ludmilla Aristilde,[8] Daniel H. Huson,[6,7,9] Andreas Dräger,[6,7,9] Ruth E. Ley,[5,9] Esteban Marcellin,[2,3,4] Largus T. Angenent,[1,9,10,11,12] and Bastian Molitor[1,9,13,*]

SUMMARY

Methanogenesis allows methanogenic archaea to generate cellular energy for their growth while producing methane. Thermophilic hydrogenotrophic species of the genus *Methanothermobacter* have been recognized as robust biocatalysts for a circular carbon economy and are already applied in power-to-gas technology with biomethanation, which is a platform to store renewable energy and utilize captured carbon dioxide. Here, we generated curated genome-scale metabolic reconstructions for three *Methanothermobacter* strains and investigated differences in the growth performance of these same strains in chemostat bioreactor experiments with hydrogen and carbon dioxide or formate as substrates. Using an integrated systems biology approach, we identified differences in formate anabolism between the strains and revealed that formate anabolism influences the diversion of carbon between biomass and methane. This finding, together with the omics datasets and the metabolic models we generated, can be implemented for biotechnological applications of *Methanothermobacter* in power-to-gas technology, and as a perspective, for value-added chemical production.

INTRODUCTION

Solutions are needed to mitigate the devastating effects of greenhouse gas emissions, primarily carbon dioxide (CO_2), and defossilize the energy and industrial sectors. Societies must efficiently implement: (1) renewable electric power to replace fossil sources; and (2) the use of the emitted CO_2 as a feedstock for the production of commodities within a circular carbon economy. Power-to-gas technologies can convert excess renewable electric power into dioxygen (O_2) and molecular hydrogen (H_2) via water electrolysis. Thus, power-to-gas systems can provide H_2 to store renewable electric energy but the hydrogen storage and transportation infrastructure is not well-established.[1,2] Legal regulatory frameworks typically limit the injection of H_2 into the established natural gas grid to under 10% v/v due to the different physical properties of H_2 and natural gas, which result in limited infrastructure and material compatibility.[1,3,4] Instead, methane (CH_4), the main constituent of natural gas, can be injected into the existing natural gas grid infrastructure to replace fossil natural gas without limitations for storage, distribution, and consumption.[5,6]

In power-to-gas technology, CH_4 can be derived from an additional methanation step in which the H_2 is combined with CO_2 to produce CH_4 and water (Equation 1).

$$CO_2\ (g) + 4H_2\ (g) \rightarrow CH_4\ (g) + 2H_2O\ (l) \qquad \text{(Equation 1)}$$

[1]Environmental Biotechnology Group, Department of Geosciences, University of Tübingen, Schnarrenbergstraße 94-96, 72076 Tübingen, Germany
[2]Australian Institute for Bioengineering and Nanotechnology, The University of Queensland, Brisbane, QLD 4072, Australia
[3]Queensland Metabolomics and Proteomics (Q-MAP), The University of Queensland, Brisbane, QLD 4072, Australia
[4]ARC Centre of Excellence in Synthetic Biology (COESB), The University of Queensland, Brisbane, QLD 4072, Australia
[5]Department of Microbiome Science, Max Planck Institute for Biology Tübingen, Max-Planck-Ring 5, 72076 Tübingen, Germany
[6]Department of Computer Science, University of Tübingen, Sand 14, 72076 Tübingen, Germany
[7]Institute for Bioinformatics and Medical Informatics (IBMI), University of Tübingen, 72076 Tübingen, Germany
[8]Department of Civil and Environmental Engineering, Northwestern University, Evanston, IL 60208, USA
[9]Cluster of Excellence – Controlling Microbes to Fight Infections, University of Tübingen, Auf der Morgenstelle 28, 72076 Tübingen, Germany
[10]AG Angenent, Max Planck Institute for Biology Tübingen, Max-Planck-Ring 5, 72076 Tübingen, Germany
[11]Department of Biological and Chemical Engineering, Aarhus University, Gustav Wieds Vej 10D, 8000 Aarhus C, Denmark
[12]The Novo Nordisk Foundation CO_2 Research Center (CORC), Aarhus University, Gustav Wieds Vej 10C, 8000 Aarhus C, Denmark
[13]Lead contact
*Correspondence: bastian.molitor@uni-tuebingen.de
https://doi.org/10.1016/j.isci.2023.108016

The methanation step in power-to-gas can be performed thermochemically via the hydrocarbon-forming Sabatier process at high temperatures (>200°C) and pressures (>1 MPa) with metal catalysts (e.g., iron, nickel, cobalt, ruthenium), which are sensitive to gas impurities.[7–10] Alternatively, the methanation step can be performed biologically with microbes as biocatalysts. Different groups of hydrogenotrophic methanogenic archaea (methanogens) natively metabolize CO_2 and H_2, producing CH_4 (Equation 1),[6,9,11] and are therefore being investigated as biocatalysts for the methanation step in power-to-gas technology.[12,13] Many methanogens exhibit low growth rates and biomass yields, which limits their potential for large-scale industrial processes. An exception to this are thermophilic methanogenic species of the genus *Methanothermobacter*, which readily grow in minimal salt media without any complex compounds. Their thermophilic growth temperature coupled with the exothermic methanogenic process results in lower temperature-control demands at industrial scales,[14] explaining the success of *Methanothermobacter* as production strains in large-scale industrial processes.[13,15–17] As an example, the company Electrochaea GmbH is creating an industrial-scale biotechnology platform with *Methanothermobacter*.

A systematic understanding of the biocatalyst's metabolic capacities is required to advance this biotechnology platform further. Previous studies expanded knowledge on the core metabolism of the genus *Methanothermobacter*, which includes methanogenesis.[18] Exemplary is an extensive comparative genome study between *Methanothermobacter thermautotrophicus* ΔH (formerly *Methanobacterium thermoautotrophicum* ΔH) and *Methanothermobacter marburgensis* Marburg (formerly *Methanobacterium thermoautotrophicum* Marburg).[19] However, only limited studies have looked at *Methanothermobacter* strains systematically in chemostat bioreactors and at the transcriptomic or proteomic level.[20–26] In fact, more than 500 hypothetical proteins and many pathways have not yet been resolved.[27] Some strains, such as *Methanothermobacter thermautotrophicus* Z-245 (formerly *Methanobacterium thermoformicicum* Z-245[18]), can utilize formic acid (or the deprotonated form, formate) as a sole growth substrate (Equation 2) in addition to CO_2 and H_2 due to the presence of a catabolic formate dehydrogenase.[23]

$$4HCOOH\ (aq) \rightarrow CH_4\ (g) + 3CO_2\ (g) + 2H_2O\ (l) \qquad \text{(Equation 2)}$$

Formic acid is derived from the (electro)chemical reduction of CO_2 and has the advantage of being in the liquid phase. Therefore, in the context of a circular carbon economy, formic acid can serve as a potential intermediate storage molecule in converting CO_2 into valuable chemicals, with formatotrophic (i.e., formic acid-utilizing) microbes as vital biocatalysts.[28]

A promising approach to systematically predict phenotypes and identify potential bottlenecks in microbial metabolisms is through mathematical analyses of their potential metabolic networks.[29] Genome-scale network reconstructions represent the theoretical metabolic capacities of a microbe. A genome-scale metabolic model (GEM) is a mathematically constrained reconstruction. These mathematical models enable *in-silico* investigations, including investigating the effects of gene deletions and insertions for metabolic engineering purposes. Thousands of reconstructions and GEMs have been assembled across all three domains of life; however, archaea remain underrepresented. Only 127 of 6239 (2%) are archaeal GEMs.[30] Of those, only ten models have been manually curated, which is a procedure that leads to higher-quality models.[30] The applications of GEMs and related models have been extensively reviewed, covering methods that are based on GEMs from flux balance analysis to kinetic modeling and machine learning.[31–36]

To further optimize methanogens for large-scale applications, metabolic engineering of these biocatalysts and an in-depth understanding of their metabolism is required. However, well-developed genetic tools exist for only a couple of mesophilic methanogens, namely different species of the genus *Methanosarcina* and *Methanococcus maripaludis*.[37,38] In previous work, we generated a genetic system for *M. thermautotrophicus* ΔH,[39,40] which allows genetically modifying this strain. To use *Methanothermobacter* strains in more versatile biotechnological applications, such as chemical production, it is necessary to further expand the available systems biology datasets and methodologies. Here, we conduct a multi-disciplinary systematic approach to assess differences in three species of the genus *Methanothermobacter*: *M. thermautotrophicus* ΔH, *M. marburgensis* Marburg, and *M. thermautotrophicus* Z-245. With this approach, we reveal differences in the formate metabolism of closely related strains, including a genetically engineered strain.[40] We observed that these differences alter the diversion of carbon between biomass and CH_4 production in the three strains. Our observations could, in future work, be exploited to implement metabolic engineering strategies for chemical production with species of the genus *Methanothermobacter*.

RESULTS

Updated genome sequences provide the basis for high-quality genome-scale metabolic reconstructions

De-novo sequencing and assembly of the genomes of *M. thermautotrophicus* ΔH, *M. thermautotrophicus* Z-245, and *M. marburgensis* Marburg with long-read sequencing technology provided high-quality sequences (first release of the *M. thermautotrophicus* Z-245 genome sequence; Table 1; Data S1 - Table S1; STAR Methods). Our sequencing approach allowed us to deduce the specific methylation pattern of the three strains, which could be important for future genetic-engineering efforts (supplemental information - Note S1; Data S1 - Table S2). In our genome annotation, we assigned COG functional categories to 1501/1796, 1505/1804, and 1440/1730 genes for *M. thermautotrophicus* ΔH, *M. thermautotrophicus* Z-245, and *M. marburgensis* Marburg, respectively (Data S1 - Table S3). We also highlighted differences to the old annotations (supplemental information - Note S2; Data S1 - Table S4; STAR Methods).

Based on the updated genome sequences, we constructed a pan-genome-scale metabolic reconstruction (pan-model) that integrates the metabolic capabilities of all three strains (STAR Methods). The pan-model reflects 29.3% (526/1796, *M. thermautotrophicus* ΔH), 28.8% (520/1804, *M. thermautotrophicus* Z-245), and 29.4% (509/1730, *M. marburgensis* Marburg) of the protein-coding genes. Therefore, our pan-model is comparable to published methanogen reconstructions.[41] Based on the pan-model, we derived strain-specific GEMs for each of the three

Table 1. Summary of sequenced genomes

Strain	Plasmid	Length (bps)	GC %	Gene Count	CDS	Previous Sequence Accessions	Difference with Old Sequence	Similarities with Old Sequence
M. thermautotrophicus ΔH	N/A	1,751,429	49.56	1844	1796	NC_000916.1 Smith et al.[62]	448 SN[a] 247 IB[b] 91 O2N[c] 92 N2O[d]	1733 CDS
M. thermautotrophicus Z-245	pFZ1	1,758,798 11,015[e]	49.46 42.51[e]	1840 12[e]	1792 12[e]	GCA_013330715.1[f] Rinke et al.[63]	N/A	N/A
M. marburgensis Marburg	pM2001 (pMTBMA4)	1,634,705 4,441[e]	48.65 45.40[e]	1774 5[e]	1725 5[e]	NC_014408.1 Liesegang et al.[64]	68 SN 30 IB 25 O2N 66 N2O	1675 CDS

[a]SN, single nucleotides.
[b]IB, indel bases.
[c]O2N, old genes that do not map to any new genes, Data S1 - Table S4.
[d]N2O, new genes that do not map to any old genes, Data S1 - Table S4.
[e]Indicates values for the plasmid.
[f]Sequence was from a metagenome and the assembly only to scaffold level.

strains (*i*MTD22IC, *i*MTZ22IC, *i*MMM22IC), which we used for further modeling to explain observed differences between the three strains as described below (Data S3).

Chemostat bioreactors reveal differences in CH_4 and biomass production rates

We operated chemostat bioreactors with the three *Methanothermobacter* strains and analyzed the generated data to compare their metabolism under steady-state growth conditions (Figures 1 and S1A; STAR Methods). Compared to both *M. thermautotrophicus* Z-245 and *M. marburgensis* Marburg, *M. thermautotrophicus* ΔH had a significantly higher consumption rate for H_2 (1.59-fold and 1.72-fold, respectively) and CO_2 (1.69-fold and 1.72-fold, respectively), as well as a higher rate of CH_4 production (1.62-fold and 1.82-fold; Figures 1A and 1B; Table 2; Data S1 - Table S5; Data S3; STAR Methods). However, *M. marburgensis* Marburg had a significantly higher biomass concentration (1.36-fold) than the other two strains (Figures 1A and 1B; Table 2; Data S1 - Table S5; Data S3). We assured that this difference was not due to an air intrusion (that led to higher O_2 levels) into one of the *M. marburgensis* Marburg bioreactors (supplemental information - Note S3; Data S3; Figure S2). The divergence resulted in differences in the normalized distribution of products, with the highest CH_4-to-biomass ratio of 96.4 $\pm$ 0.57 for *M. thermautotrophicus* ΔH and the lowest of 93.5 $\pm$ 0.98 for *M. marburgensis* Marburg (Figure 1C).

Multi-omics analysis under steady-state bioreactor growth conditions reveals different gene-expression patterns

To clarify why the three closely related strains displayed significant differences in the growth behavior under identical growth conditions (Figure 1), we performed transcriptomic and proteomic analyses during steady state for all replicate bioreactors (Data S4; STAR Methods). For each strain, we obtained highly reproducible transcriptomic and proteomic data across replicates as well as a high coverage of the transcriptome (94.18–99.77%) and proteome (78.29–79.82%) (Table 3 Individual analyses).

To compare the omics datasets between strains, we merged the genome-derived transcriptome and proteome of the three strains into pan-transcriptome and pan-proteome databases (creating reference genes, termed "gene groups," Data S4; STAR Methods). Merging allowed us to perform differential expression analyses. We found many differentially expressed genes in the pan-transcriptome but much fewer differentially abundant proteins in the pan-proteome comparison, and in both omics sets, fewer differences between *M. thermautotrophicus* ΔH and *M. thermautotrophicus* Z-245 than to *M. marburgensis* Marburg, all of which is consistent with their phylogeny (Figures S3C–S3G; Table 3 – Differential analyses; supplemental information - Note S4; Data S1 - Table S6; Data S4).[18] Because the quality of both the transcriptomic and proteomic data was high (supplemental information - Note S4), we hypothesized that each strain may have evolved different transcriptional regulation patterns leading to a very similar outcome on the proteome level to maximize growth under thermodynamically limited conditions.

A closer look at the differential expression patterns of methanogenesis – the core metabolic feature of methanogens – did not provide any clear explanation for the observed differences in CH_4 production between the strains (supplemental information - Note S5). Next, we investigated reactions that branch off from methanogenesis toward biomass growth and further downstream in the anabolism but still did not find explanations for the observed differences in growth behavior (supplemental information - Note S6). Thus, we integrated the omics and fermentation data into the GEMs to better understand these differences.

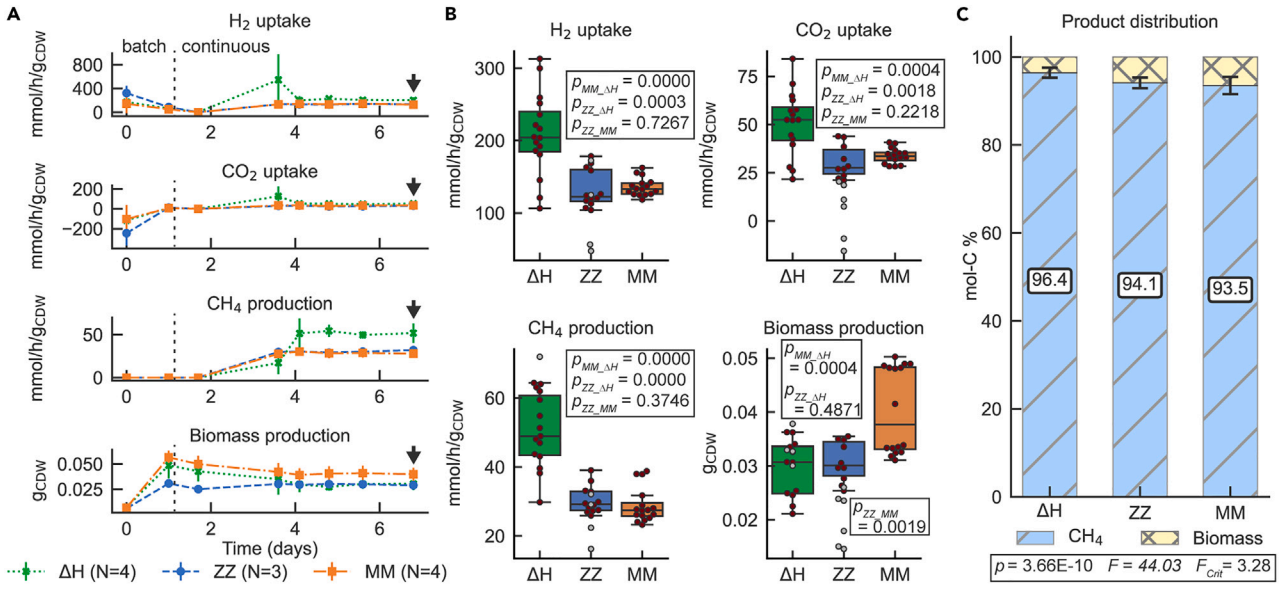


Figure 1. Fermentation data from the first experiment in chemostat bioreactors with *M. thermautotrophicus* ΔH (ΔH), *M. thermautotrophicus* Z-245 (ZZ), and *M. marburgensis* Marburg (MM)

(A) Gas consumption (H_2 and CO_2 uptake), and CH_4 and biomass production data from quadruplicate ($N = 4$, ΔH and MM) and triplicate ($N = 3$, ZZ) bioreactors for the fermentation period of 7 days (mean ±1 standard deviation). For ZZ, we discarded one replicate due to a pump malfunction in this bioreactor and wash-out of the cells. Data for further analyses (transcriptomics, proteomics) were taken on day seven, as indicated by arrows.

(B) Average gas consumption (H_2 and CO_2 uptake) and CH_4 and biomass production data during a steady-state period (days 4–7). For statistical analysis in pairwise comparisons with *t*-test, data points without suspected gross measurement error (red circles) were included, data points with suspected gross measurement error (gray circles) were excluded (STAR Methods).

(C) Average normalized product distribution (mean ±1 standard deviation), including statistical analysis by ANOVA (n = 11 (ΔH), n = 10 (ZZ), n = 16 (MM); STAR Methods). ΔH, *M. thermautotrophicus* ΔH; ZZ, *M. thermautotrophicus* Z-245; MM, *M. marburgensis* Marburg: *F*, *F* value; F_{crit}, *F* critical value.

Integration of experimental data into GEMs results in solutions for flux balance analyses

We constrained the GEMs in flux balance analysis simulations with the steady-state CO_2 and H_2 consumption rates and CH_4 and biomass production rates from the bioreactor experiment; these rates were corrected for experimental errors to obtain gross-measurement-error-adjusted values (Data S5 [Simulations 25–28, 37–40, and 49–52]; STAR Methods). In these simulations, the GEMs found solutions with a 0.1%, 0.1%, and 0.5% deviation of the reaction flux bounds from the calculated fluxes for *M. thermautotrophicus* ΔH, *M. thermautotrophicus* Z-245, and *M. marburgensis* Marburg, respectively (Data S5). The maximum fluxes of the non-growth-associated maintenance energy (reaction ATPM) varied between the three strains (Data S1 - Table S7).

We further constrained the GEMs with transcriptomics and proteomics using the GIMME algorithm,[42] which resulted in reduced GEMs. We performed a set of flux balance analysis simulations comparing the unconstrained and reduced GEMs (Data S5 [transcriptomics: Simulations 29–32, 41–44, and 53–56; proteomics: Simulations 33–36, 37–40, and 57–60]; supplemental information - Note S7; STAR Methods). This led us to detect differences in the formate anabolism pathway among the three strains, as detailed below.

Flux balance analyses reveal differences in formate anabolism

Among the three investigated *Methanothermobacter* strains, only *M. thermautotrophicus* Z-245 is able to use formate for methanogenesis, and thus as the sole growth substrate.[23] However, in all three *Methanothermobacter* strains, formate is a crucial metabolite for anabolism and has to be synthesized during growth with CO_2 and H_2.[43–45] This is because formate is an important intermediate for *de-novo* purine biosynthesis.[19,46]

Formate for biomass could result from the activity of a formate dehydrogenase (Fdh), which, in the case of methanogens, interconverts CO_2 into formate with concomitant reduction of cofactor F_{420}.[47] The use of formate as a growth substrate in *M. thermautotrophicus* Z-245 depends on the presence of a Fdh catalyzing the reverse reaction.[23] Although *M. thermautotrophicus* ΔH encodes an Fdh (MTH1140/ISG35_05415 and MTH1139/ISG35_05410, gene groups 967 and 968; Data S4), this strain is not able to utilize formate as a growth substrate.[18] Instead, the Fdh of gene groups 967 and 968 was previously hypothesized to be the relevant enzyme for the synthesis of formate for biomass growth.[19] In spite of this, reducing the GEM with proteomic data eliminated the corresponding FDHf420 reaction in the flux balance analysis simulation because Fdh was not detected via proteomics (Figure 2; Data S5).

In contrast, the model constrained with the proteomic data suggested a formate acetyltransferase/pyruvate-formate lyase (Pfl) (PFL reaction) as the preferred enzyme for the production of formate (Figure 2; Data S5) in *M. thermautotrophicus* ΔH (MTH346/ISG35_1600). This

Table 2. Summary of the gas fermentation data for the four strains used in the first and second chemostat bioreactor experiments

First Fermentation Experiment – Period 2 (H_2 and CO_2)[a]			
	ΔH[b]	ZZ[c]	MM[d]
Biomass (g_{CDW})	0.03 ± 0.01	0.03 ± 0.00	0.04 ± 0.01
H_2 Uptake Rates ($mmol/g_{CDW}/h$)	234.36 ± 40.76	147.24 ± 23.88	136.12 ± 12.00
CO_2 Uptake Rates ($mmol/g_{CDW}/h$)	58.18 ± 11.49	34.36 ± 6.98	33.80 ± 3.78
CH_4 Production Rates ($mmol/g_{CDW}/h$)	52.50 ± 8.53	32.34 ± 3.92	28.83 ± 4.93
Second Fermentation Experiment (H_2 and CO_2 or Formate)[e]			
	ΔH_H[f]	ΔH_F[g]	ZZ_F[h]
Biomass (g_{CDW})	0.04 ± 0.01	0.04 ± 0.01	0.04 ± 0.02
H_2 Uptake Rates ($mmol/g_{CDW}/h$)	127.31 ± 28.18	N/A	N/A
H_2 Production Rates ($mmol/g_{CDW}/h$)	N/A	0.84 ± 0.34	0.55 ± 0.28
CO_2 Uptake Rates ($mmol/g_{CDW}/h$)	29.16 ± 3.91	N/A	N/A
CO_2 Production Rates ($mmol/g_{CDW}/h$)	N/A	31.87 ± 11.12	34.29 ± 9.77
CH_4 Production Rates ($mmol/g_{CDW}/h$)	30.53 ± 3.97	9.69 ± 3.62	10.55 ± 3.18

The standard deviation (±) was calculated based on the entire population.
[a]Values can be found in Data S3 and represent mean values and standard deviations calculated from values without gross measurement errors (STAR Methods).
[b]ΔH, *M. thermautotrophicus* ΔH.
[c]ZZ, *M. thermautotrophicus* Z-245.
[d]MM, *M. marburgensis* Marburg.
[e]Values can be found in Data S6 and represent mean values and standard deviations calculated from values without gross measurement errors (STAR Methods).
[f]ΔH_H, *M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$ grown on H_2 and CO_2.
[g]ΔH_F, *M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$ grown on Na-formate.
[h]ZZ_F, *M. thermautotrophicus* Z-245 grown on Na-formate.

enzyme catalyzes the CoA-dependent reversible conversion of pyruvate to formate and was previously identified in *M. thermautotrophicus* ΔH.[19,48] This observation implied that *M. thermautotrophicus* ΔH primarily produces formate for biomass growth via Pfl. This hypothesis will need to be confirmed experimentally to corroborate the proteomics-derived evidence for the lack of the Fdh enzyme.

Neither the genome of *M. marburgensis* Marburg[19,48] nor *M. thermautotrophicus* Z-245 encodes a Pfl (supplemental information - Note S8). However, in contrast to both *M. thermautotrophicus* ΔH and *M. marburgensis* Marburg, *M. thermautotrophicus* Z-245 has two Fdhs. The first Fdh in *M. thermautotrophicus* Z-245 is homologous to the Fdh found in *M. thermautotrophicus* ΔH (gene groups 967 and 968). The second Fdh is encoded in the operon *fdhCAB* (i.e., $fdh_{cassette}$, ISG36_07620, ISG36_07615, ISG36_07610). Importantly, the operon includes an additional putative formate transporter (FdhC), and was previously reported to be responsible for the utilization of formate as the sole growth substrate in *M. thermautotrophicus* Z-245.[23] Accordingly, the expression of the $Fdh_{cassette}$ is regulated by the relative abundance of substrates, namely, formate, CO_2 and H_2.[23]

Based on these previous reports and because we fed the bioreactors with H_2, we had anticipated that the $Fdh_{cassette}$ would not be expressed (or found at low abundance) in our conditions. Unexpectedly, we found high absolute levels of the $Fdh_{cassette}$ in our transcriptome and proteome (Figures 2 and S4; Data S2). By contrast, the first Fdh was less abundant in our transcriptome and not detected in the proteome (Data S2). This indicated that in the presence of H_2 and CO_2 as sole growth substrates *M. thermautotrophicus* Z-245 uses the $Fdh_{cassette}$ to produce formate from H_2 and CO_2 for biomass growth. Indeed, our flux balance analysis simulations with *M. thermautotrophicus* Z-245 primarily predicted the use of the FDH_F420 ($Fdh_{cassette}$) in the direction of formate production rather than FDHf420 (the aforementioned first Fdh) (Figures 2 and S4; Data S5). Future research will need to confirm the absence of the first Fdh and activity of the $Fdh_{cassette}$ in the direction of formate production *in vivo*.

While the genome of *M. marburgensis* Marburg also encodes the first Fdh (MTBMA_c15220/ISG34_07515 and MTBMA_c15230/ISG34_07520, gene groups 967 and 968), it neither encodes the Pfl nor the $Fdh_{cassette}$ (Data S4). Thus, formate for biomass production is likely only produced from the first Fdh (gene groups 967 and 968). This was also hypothesized previously, and this hypothesis was supported by a formate-auxotrophic strain of *M. marburgensis* Marburg, which did not exhibit Fdh activity anymore.[49] All flux balance analysis simulations (including the omics-constrained ones) predicted the occurrence of the first Fdh reaction, that is the FDHf420 reaction (Data S5). The two subunits were upregulated in the transcriptome by 2.4–4.9 log_2 fold change compared to the other two strains (Figure S4; Data S4). However, we found only one subunit (FdhA) in the proteome with low abundance (gene group 967: 0.000132%). This was three orders of magnitude lower than the homologous subunit of the $Fdh_{cassette}$ in *M. thermautotrophicus* Z-245 (FdhB: 0.254%, FdhA: 0.191%, FdhC: 0.126%) (Figure 2; Data S4). These findings raise questions about how *M. marburgensis* Marburg produced formate for biomass growth as discussed below.

Table 3. Metrics of the transcriptomics and proteomics for the individual analyses and the differential expression and differential abundance analyses for the three strains

Omics	Metric	ΔH[a]	ZZ[b]	MM[c]
Individual analyses				
Transcriptomics	*Million reads*	26.3 (±2.0)	27.4 (±0.8)	26.0 (±1.1)
	Number of CDSs found (%)	1697/1796 (94.48%)	1699/1804 (94.18%)	1726/1730 (99.77%)
	Average CV[d] of all genes	35.03	33.07	26.60
	Percent of genes with CV[d] values ≤ 20%	44.12	39.61	26.83
Proteomics	*Number of proteins found (%)*	1406/1796 (78.29%)	1440/1804 (79.82%)	1374/1730 (79.42%)
	Average CV[e] of all proteins	11.37	18.73	15.51
	Percent of proteins with CV[e] values ≤ 20%[g]	88.35	66.48	77.72

Omics	Metric	MM vs. ΔH	ZZ vs. ΔH	ZZ vs. MM
Differential analyses				
Transcriptomics	*DE[f] genes (range of log_2FC)*	1482 (−13.1 to 17.8)	816 (−17.0 to 13.8)	1498 (−14.1 to 16.7)
	DE genes with $\|log_2FC\| \geq 2$	975	265	1001
	DE genes that map to a reaction in the GEM	469	249	484
Proteomics	*DA[g] proteins (range of log_2FC)*	121 (−6.6 to 6.6)	91 (−6.6 to 6.6)	93 (−6.0 to 5.9)
	DA proteins with $\|log_2FC\| \geq 2$	103	59	35
	DA proteins that map to a reaction in the GEM	43	17	38
Transcriptomics + Proteomics	*Fraction of DA proteins that were also DE genes*	110/121 (91%)	45/91 (50%)	90/102 (88%)

Statistical significance was determined by an adjusted p-value ≤0.05. The standard deviation (±) was calculated based on the entire population.
[a]ΔH, *M. thermautotrophicus* ΔH.
[b]ZZ, *M. thermautotrophicus* Z-245.
[c]MM, *M. marburgensis* Marburg.
[d]Coefficient of variance (CV) values for the transcriptomics data were calculated using transcripts per million (TPM) values and can be found in Sheets 1–3 of Data S4 and Figure S6B.
[e]Coefficient of variance (CV) values for the proteomics data were calculated using iBAQ values and can be found in Sheets 7–9 of Data S4 and Figure S6B.
[f]DE, differential expression.
[g]DA, differential abundance.

Genetically engineered *M. thermautotrophicus* ΔH can stably grow on formate in chemostat bioreactors

The finding that formate anabolism might be distinct in all three strains was unexpected and we set out to further assess formate metabolism with a second chemostat bioreactor experiment. We investigated the growth behavior of a genetically engineered *M. thermautotrophicus* ΔH strain encoding the $Fdh_{cassette}$ from *M. thermautotrophicus* Z-245 on a plasmid (*M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$).[40] *M. thermautotrophicus* Z-245 is able to use formate as the sole growth substrate via the $Fdh_{cassette}$ reaction[40] but presumably also produces formate for biomass growth by the reverse activity of the $Fdh_{cassette}$. Instead, *M. thermautotrophicus* ΔH uses Pfl for this purpose, as indicated by our omics data. Here we operated chemostat bioreactors that were fed with 355 ± 5 mM sodium formate instead of H_2 and CO_2 with *M. thermautotrophicus* Z-245 and *M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$ (Figure 3A; STAR Methods). Additionally, for *M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$, we operated also bioreactors with H_2 and CO_2 to allow a comparison to the first chemostat bioreactor experiment (Table 2; Figures 3A and S1B; STAR Methods). With sodium formate, these two strains neither showed significant differences in sodium formate uptake rates nor in CH_4, CO_2, and biomass production rates (Table 2; Figure 3B; Data S6), confirming that the episomal expression of the $Fdh_{cassette}$ enables *M. thermautotrophicus* ΔH to grow on formate as *M. thermautotrophicus* Z-245 does. During growth on H_2 and CO_2, the *M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$ exhibited a growth behavior more similar to *M. thermautotrophicus* Z-245 and *M. marburgensis* Marburg compared to wild-type *M. thermautotrophicus* ΔH (Table 2; Figure S5). These results indicate that, under the tested fermentation conditions, the reverse activity of the $Fdh_{cassette}$ enables *M. thermautotrophicus* ΔH to more efficiently form formate for biomass growth than by the sole activity of Pfl. The differences in growth behavior between the genetically engineered *M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$ strain and the wild-type *M. thermautotrophicus* ΔH might have implications for biotechnological applications, as discussed below. Notably, the plasmid was maintained in the *M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$ culture in the absence of antibiotics (Figures S1B and S1C). While we cannot completely rule out that only a subpopulation of the modified strain maintained the plasmid without antibiotic selection, this finding further supports the applicability of the genetic system for bioprocessing.

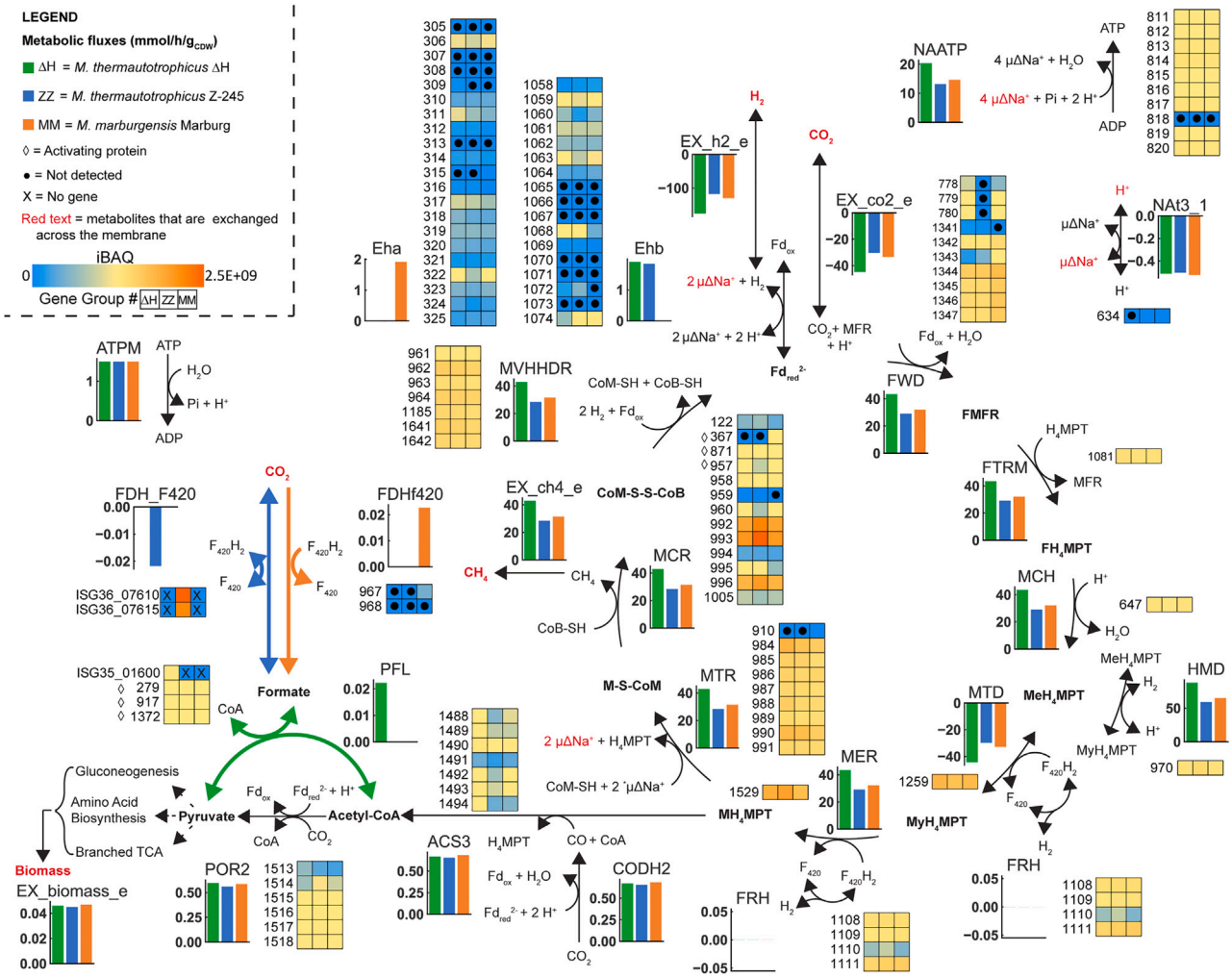


Figure 2. Wolfe Cycle adapted from Thauer[61] with reactions involved in the energy metabolism

Fluxes are from the proteomics reduced model (using intensity based absolute quantification (iBAQ) values), constrained with experimental data that was adjusted for gross measurement error. iBAQ values are given for proteins. Gene group is used as ID for the omics. For the PFL reaction, only *M. thermautotrophicus* ΔH has the gene; thus, the gene ID is used (ISG35_01600). For the FDH_F420, only *M. thermautotrophicus* Z-245 has the formate dehydrogenase cassette; thus, the gene IDs are used (ISG36_07610 and ISG36_07615). Red text refers to metabolites that are exchanged across the membrane. **Microbes**: ΔH, *M. thermautotrophicus* ΔH; MM, *M. marburgensis* Marburg; ZZ, *M. thermautotrophicus* Z-245. **Compounds**: CH_4, methane; CO, carbon monoxide; CoA, Coenzyme A; CoB, coenzyme B; CoM, coenzyme M; CoM-S-S-CoB, CoM-CoB heterodisulfide; CO_2, carbon dioxide; F, formyl; $Fd_{ox/rd}$, ferredoxin oxidized/reduced; H_2, hydrogen; H^+, proton; H_4MPT, tetrahydromethanopterin; M, methyl; Me, methenyl; MFR, methanofuran; My, methylene; Na^+, sodium ion. **Reactions/Enzymes:** ATPM, ATP maintenance (pseudo reaction); CODHr2, CO dehydrogenase/acetyl-CoA synthase; Eha/Ehb, energy converting hydrogenases; EX_biomass_e, biomass exchange (pseudo reaction); EX_ch4_e, CH_4 exchange (pseudo reaction); EX_co2_e, CO_2 exchange (pseudo reaction); EX_h2_e, H_2 exchange (pseudo reaction); FDHf420, F_{420}-dependent formate dehydrogenase; FDH_F420, F_{420}-dependent formate dehydrogenase cassette; FRH, F_{420}-reducing hydrogenase; FTRM, FMFR/H_4MPT formyltransferase; FWD, FMFR dehydrogenase (tungsten- and molybdenum-dependent isozymes); HMD, MeH_4MPT hydrogenase; MCH, MeH_4MPT cyclohydrolase; MCR, MCoM reductase (I and II); MER, MyH_4MPT reductase; MTD, MyH_4MPT dehydrogenase; MTR, MH_4MPT/CoM methyltransferase; MVHHDR, F_{420}-non-reducing hydrogenase with the heterodisulfide reductase; NAATP, ATP synthase; Nat3_1, Na^+/H^+ antiporter; PFL, pyruvate formate-lyase; POR2, pyruvate synthase. **Other:** ◊, activating protein; ●, not detected; X, no gene. For an extended figure see Figure S4.

DISCUSSION

Although methanogenesis is well studied, there are still knowledge gaps surrounding methanogenic metabolism. In particular, understanding the metabolic strengths and limitations of different strains is necessary to select the most appropriate strain for a given biotechnological application, such as optimization of CH_4 production or engineering strains to produce value-added products. In this study, we systematically compared *M. thermautotrophicus* ΔH, *M. thermautotrophicus* Z-245, and *M. marburgensis* Marburg to identify differences in their metabolism.

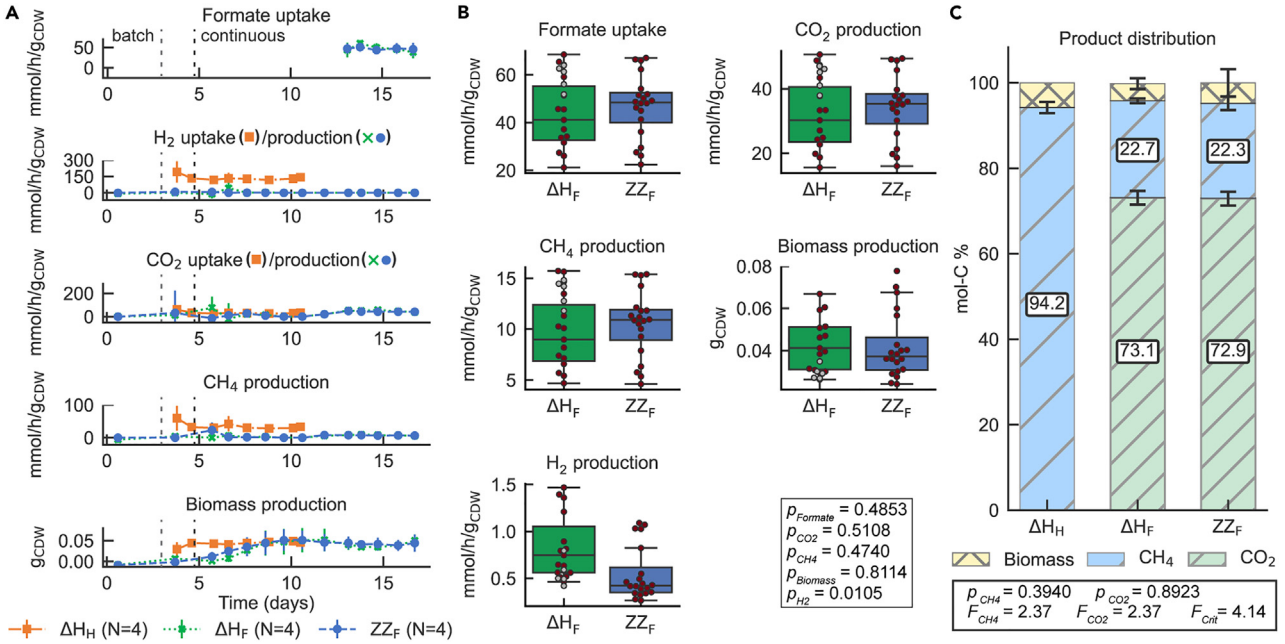


Figure 3. Fermentation data from the second experiment in chemostat bioreactors with *M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$ (ΔH) and *M. thermautotrophicus* Z-245 (ZZ)

(A) Gas consumption (H_2 and CO_2 uptake) or sodium formate consumption, and CH_4 and biomass production data from quadruplicate bioreactors for the fermentation period of 11 (ΔH_H) and 17 (ΔH_F and ZZ_F) days (mean ±1 standard deviation). Formate uptake rates were only determined once steady state was reached and all formate provided was consumed.

(B) Average sodium formate consumption and H_2, CO_2, CH_4, and biomass production data during the steady-state period (days 13–17) for ΔH_F ($N = 3$) and ZZ_F ($N = 4$). For statistical analysis in pairwise comparisons with *t*-test, data points without suspected gross measurement error (red circles) were included, and data points with suspected gross measurement error (gray circles) were excluded (STAR Methods).

(C) Average normalized product distribution (mean ±1 standard deviation), including a statistical analysis by ANOVA for ΔH_F ($N = 3$) and ZZ_F ($N = 4$) (STAR Methods). CH_4, methane; CO_2, carbon dioxide; H_2, molecular hydrogen; ΔH, *M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$; ZZ, *M. thermautotrophicus* Z-245; F, sodium formate as substrate; H, H_2 and CO_2 as substrates; *F*, *F* value; F_{crit}, *F* critical value.

Under the conditions of our bioreactor experiments, *M. thermautotrophicus* ΔH had a higher specific CH_4 production rate than the other strains, while *M. marburgensis* Marburg reached higher biomass production rates. From our modeling results (with the maximization of the ATPM reaction), we deduced that *M. thermautotrophicus* ΔH has the highest and *M. marburgensis* Marburg has the lowest non-growth-associated maintenance energy (Data S1 - Table S7). The non-growth-associated maintenance energy represents the dissipation of ATP (as no storage compounds are known).[50] The higher the non-growth-associated maintenance energy, the more CH_4 (and ATP) must be produced per biomass unit. Based on our multi-omics analysis we propose that differences in anabolism could explain this observation, while further investigations are required to confirm this hypothesis. Our data indicate that the three *Methanothermobacter* strains might use three different enzymes to produce formate for biomass growth, such as for purine biosynthesis.[43–45] From an ecological perspective, the observed differences in formate anabolism might have an impact on the adaptation to different ecological niches in which, for example, a higher CH_4-production rate over biomass yield would provide a selective benefit to *M. thermautotrophicus* ΔH. Furthermore, as ATP is the primary energy currency in the cell, these findings are important to optimize the use of *Methanothermobacter* strains as cell factories by metabolic engineering in the future.[51]

Intriguingly, while *M. marburgensis* Marburg reached higher biomass production rates (Figure 1C), none of the putative enzymes for formate anabolism were abundant in our proteomics analyses. If the low-abundant Fdh is responsible for formate production in *M. marburgensis* Marburg, questions on the kinetic properties of this enzyme remain. Alternatively, formate could be produced via other, yet unconsidered pathways, for example, through a hydrogen-dependent CO_2 reductase (HDCR)-like activity.[52] While formate is not an intermediate in CO_2 reduction during methanogenesis, a weak formate dehydrogenase activity has been ascribed to the formylmethanofuran dehydrogenases (Fmd/Fwd).[53] Recently, a large enzyme complex was shown to be involved in the electron-bifurcating step of CO_2 reduction without the release of ferredoxin.[54] This enzyme complex may provide the formate that is required for anabolism by an apparent HDCR-like activity.

Understanding the metabolism of a microbe is especially important to guide its rewiring for biotechnological purposes. Here, we demonstrated the long-term ability of the strain *M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$, encoding the catabolic $Fdh_{cassette}$ from *M. thermautotrophicus* Z-245 on a plasmid, to grow on formate in chemostat bioreactors. Introducing the $Fdh_{cassette}$ into *M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$ resulted in a growth behavior that is more similar to wildtype

M. thermautotrophicus Z-245 than to *M. thermautotrophicus* ΔH. This finding supports our hypothesis that, indeed, the different solutions for the three methanogens to produce formate in anabolism have an impact on the CH_4-to-biomass ratios. The CH_4-to-biomass ratio is an essential parameter in bioprocessing (it should be high for CH_4 production but low for chemical production). These observations have implications to design strategies for the redirection of carbon from CH_4 to other high-value products with metabolic engineering, although further investigation is needed (e.g., with biochemical investigations of the enzyme systems). The ability to combine the GEM with the genetic system will be a considerable step toward chemical production with thermophilic methanogens because the use of validated GEMs will enable phenotypes to be predicted. In summary, our combined methods lay the foundation for archaeal biotechnology to optimize existing power-to-gas processes or enable the production of value-added chemicals by metabolic engineering of microbes of the genus *Methanothermobacter*.

Limitations of the study

Based on our findings, *M. thermautotrophicus* ΔH appears the most suitable among the three investigated strains for industrially relevant power-to-gas systems due to its superior kinetics and CH_4 specificity. However, it is essential to consider the fermentation conditions of our study. We applied a volume gas per volume bioreactor per minute (vvm) of 0.08, while others used a much higher vvm of up to 2.01.[13] It must be further noted that the composition of the growth medium was not optimized for any of the three strains, to avoid a bias toward one of the strains. Whether the superior behavior of *M. thermautotrophicus* ΔH would hold at the higher vvm and with strain-optimized specific media was not evaluated here and should be considered.

Our modeling approach is based on assumptions of steady state and considers static fluxes in the metabolism. Many additional factors need to be considered, such as enzyme kinetics and reaction capacities, which may result in bottlenecks in certain pathways. Those considerations were not included in the modeling approach that we applied here, in part due to limited data on enzyme kinetics for the *Methanothermobacter* strains. However, our manually curated genome-scale reconstruction is the prerequisite for follow-up modeling approaches including but not limited to: (1) integrating transcriptomics and proteomics with alternative algorithms and methods, such as iMAT[55,56] or GIMMEp;[57] (2) model constraining with additional types of data, such as enzyme kinetic data (e.g., kinetome[58]) and ^{13}C fluxomics;[59] and (3) generating superimposed models, such as metabolic and expression models (ME-models).[60]

Our multi-disciplinary systematic approach led us to hypothesize that the three investigated *Methanothermobacter* strains produce formate for biomass growth in distinct ways. By including the genetically engineered *M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$ strain, we further support this hypothesis. However, future experiments will need to confirm our observations and hypotheses, for example, by measuring the involved intracellular metabolites or biochemically assaying the enzymatic reactions.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Microbial strains
  - Growth media
  - Growth conditions
- METHOD DETAILS
  - Genome sequencing
  - Genome-scale metabolic model reconstruction
  - Biomass and maintenance energy determination
  - Cross-contamination check
  - Fermentation gas analysis
  - Biomass concentration analysis
  - Sodium formate concentration measurements
  - RNA sample preparation
  - RNA sequencing
  - Protein sample preparation
  - Protein measurement
  - Pan-genome differential expression database
  - Integrating fermentation data in the GEMs
  - Visualization
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Genome assembly

- Genome comparisons
- COG functional annotation
- Carbon and electron balance calculations
- Normalized product distribution
- Interspecies comparison
- Raw RNA sequencing data analysis
- Protein analysis with proteome Discoverer
- Eha/Ehb ratio determination
- Methanogenesis relative abundances

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.108016.

ACKNOWLEDGMENTS

The authors would like to thank Luis Antoniotti and Jürgen Barth from the Max Planck Institute for Biology Tübingen workshop for their help with modifying the bioreactor system. The authors would like to thank Lucas Mühling, Nils Rohbohm, and Andrés Ortíz Ardilla for helpful input and support with experimentation, and Nicolai Kreitli for support with bioreactor experiments. The authors would like to acknowledge the assistance of the Genome Center of the Max Planck Institute for Biology Tübingen, particularly Christa Lanz, for their assistance in the PacBio sequencing, and the Proteome Center Tübingen (PCT) for their support with the proteomics. The authors would like to acknowledge the fruitful conversations with Theresa Ahrens and Doris Hafenbradl from Electrochaea GmbH, Planegg, Germany. The authors would like to thank Libera Lo Presti from the Cluster of Excellence – CMFI at the University of Tübingen for language editing and proofreading the manuscript. This work was supported by the Humboldt Foundation in the framework of the Humboldt professorship, which was awarded to LTA, and by the Max Planck Institute for Biology Tübingen (REL and LTA). IC would like to acknowledge support from the German Academic Exchange Service (DAAD) through the DAAD *Kurzzeitstipendien für Doktoranden*. EM acknowledges support from the Australian Research Council for the Center of Excellence in Synthetic Biology.

AUTHOR CONTRIBUTIONS

I.C.: Methodology, Software, Formal Analysis, Investigation, Data Curation, Writing - Original Draft & Review & Editing, Visualization, Project Administration, Funding acquisition; **T.M.:** Methodology, Software, Formal Analysis, Writing – Review and Editing; **S.E-E.:** Methodology, Investigation, Writing – Review and Editing; **G.G.L.:** Software, Formal Analysis, Writing – Review and Editing; **D.E.:** Formal Analysis, Investigation, Writing - Review and Editing; **C.F.:** Investigation, Writing – Review and Editing; **S.B.:** Investigation, Writing – Review and Editing; **N.D.Y.:** Methodology, Writing – Review and Editing; **L.A.:** Writing - Review and Editing; **D.H.H.:** Supervision, Writing – Review and Editing, Funding Acquisition; **A.D.:** Software, Writing – Review and Editing, Funding Acquisition; **R.E.L.:** Resources, Supervision, Writing – Review and Editing, Funding Acquisition; **E.M.:** Methodology, Resources, Writing – Review and Editing, Funding Acquisition; **L.T.A.:** Conceptualization, Resources, Writing – Review and Editing, Supervision, Project administration, Funding Acquisition; **B.M.:** Conceptualization, Writing – Original Draft, Review and Editing, Supervision, Project administration, Funding acquisition.

DECLARATION OF INTERESTS

The authors declare no conflict of interest.

Received: July 27, 2023
Revised: August 29, 2023
Accepted: September 19, 2023
Published: September 22, 2023

REFERENCES

1. Zhao, Y., McDonell, V., and Samuelsen, S. (2020). Assessment of the combustion performance of a room furnace operating on pipeline natural gas mixed with simulated biogas or hydrogen. Int. J. Hydrogen Energy *45*, 11368–11379.
2. Reuß, M., Grube, T., Robinius, M., and Stolten, D. (2019). A hydrogen supply chain with spatial resolution: Comparative analysis of infrastructure technologies in Germany. Appl. Energy *247*, 438–453.
3. Melaina, M.W., Antonia, O., and Penev, M. (2013). Blending Hydrogen into Natural Gas Pipeline Networks: A Review of Key Issues (National Renewable Energy Laboratory). NREL/TP-5600-51995.
4. Hall, J.E., Hooker, P., and Jeffrey, K.E. (2021). Gas detection of hydrogen/natural gas blends in the gas industry. Int. J. Hydrogen Energy *46*, 12555–12565.
5. Schiebahn, S., Grube, T., Robinius, M., Tietze, V., Kumar, B., and Stolten, D. (2015). Power to gas: Technological overview, systems analysis and economic assessment for a case study in Germany. Int. J. Hydrogen Energy *40*, 4285–4294.
6. Guerra, L., Rossi, S., Rodrigues, J., Gomes, J., Puna, J., and Santos, M.T. (2018). Methane production by a combined Sabatier reaction/water electrolysis process. J. Environ. Chem. Eng. *6*, 671–676.
7. Müller, K., Städter, M., Rachow, F., Hoffmannbeck, D., and Schmeißer, D. (2013). Sabatier-based CO_2-methanation by

catalytic conversion. Environ. Earth Sci. *70*, 3771–3778.
8. Dry, M.E. (2002). The fischer–tropsch process: 1950–2000. Catal. Today *71*, 227–241.
9. Leonzio, G. (2016). Process analysis of biological Sabatier reaction for bio-methane production. Chem. Eng. J. *290*, 490–498.
10. Van Der Laan, G.P., and Beenackers, A.A.C.M. (1999). Kinetics and selectivity of the Fischer–Tropsch synthesis: a literature review. Catal. Rev. *41*, 255–318.
11. Rachbauer, L., Beyer, R., Bochmann, G., and Fuchs, W. (2017). Characteristics of adapted hydrogenotrophic community during biomethanation. Sci. Total Environ. *595*, 912–919.
12. Thauer, R.K., Kaster, A.-K., Seedorf, H., Buckel, W., and Hedderich, R. (2008). Methanogenic archaea: ecologically relevant differences in energy conservation. Nat. Rev. Microbiol. *6*, 579–591.
13. Pfeifer, K., Ergal, İ., Koller, M., Basen, M., Schuster, B., and Rittmann, S.K.M.R. (2021). Archaea biotechnology. Biotechnol. Adv. *47*, 107668.
14. Rusmanis, D., O'Shea, R., Wall, D.M., and Murphy, J.D. (2019). Biological hydrogen methanation systems–an overview of design and efficiency. Bioengineered *10*, 604–634.
15. Martin, M.R., Fornero, J.J., Stark, R., Mets, L., and Angenent, L.T. (2013). A Single-Culture Bioprocess of *Methanothermobacter thermautotrophicus* to Upgrade Digester Biogas by CO_2-to-CH_4 Conversion with H_2. Archaea *2013*, 157529.
16. Kaul, A., Böllmann, A., Thema, M., Kalb, L., Stöckl, R., Huber, H., Sterner, M., and Bellack, A. (2022). Combining a robust thermophilic methanogen and packing material with high liquid hold-up to optimize biological methanation in trickle-bed reactors. Bioresour. Technol. *345*, 126524.
17. Thema, M., Weidlich, T., Kaul, A., Böllmann, A., Huber, H., Bellack, A., Karl, J., and Sterner, M. (2021). Optimized biological CO_2-methanation with a pure culture of thermophilic methanogenic archaea in a trickle-bed reactor. Bioresour. Technol. *333*, 125135.
18. Wasserfallen, A., Nölling, J., Pfister, P., Reeve, J., and Conway de Macario, E. (2000). Phylogenetic analysis of 18 thermophilic *Methanobacterium* isolates supports the proposals to create a new genus, *Methanothermobacter* gen. nov., and to reclassify several isolates in three species, *Methanothermobacter thermautotrophicus* comb. nov., *Methanothermobacter wolfeii* comb. nov., and *Methanothermobacter marburgensis* sp. nov. Int. J. Syst. Evol. Microbiol. *50 Pt 1*, 43–53.
19. Kaster, A.-K., Goenrich, M., Seedorf, H., Liesegang, H., Wollherr, A., Gottschalk, G., and Thauer, R.K. (2011). More than 200 Genes Required for Methane Formation from H_2 and CO_2 and Energy Conservation Are Present in *Methanothermobacter marburgensis* and *Methanothermobacter thermautotrophicus*. Archaea *2011*, 973848.
20. de Poorter, L.M.I., Geerts, W.G., Theuvenet, A.P.R., and Keltjens, J.T. (2003). Bioenergetics of the formyl-methanofuran dehydrogenase and heterodisulfide reductase reactions in *Methanothermobacter thermautotrophicus*. Eur. J. Biochem. *270*, 66–75.
21. de Poorter, L.M.I., Geerts, W.J., and Keltjens, J.T. (2007). Coupling of *Methanothermobacter thermautotrophicus* methane formation and growth in fed-batch and continuous cultures under different H_2 gassing regimens. Appl. Environ. Microbiol. *73*, 740–749.
22. Afting, C., Kremmer, E., Brucker, C., Hochheimer, A., and Thauer, R.K. (2000). Regulation of the synthesis of H_2-forming methylenetetrahydromethanopterin dehydrogenase (Hmd) and of HmdII and HmdIII in *Methanothermobacter marburgensis*. Arch. Microbiol. *174*, 225–232. https://doi.org/10.1007/s002030000197.
23. Nölling, J., and Reeve, J.N. (1997). Growth- and substrate-dependent transcription of the formate dehydrogenase (*fdhCAB*) operon in *Methanobacterium thermoformicicum* Z-245. J. Bacteriol. *179*, 899–908.
24. Kato, S., Kosaka, T., and Watanabe, K. (2008). Comparative transcriptome analysis of responses of *Methanothermobacter thermautotrophicus* to different environmental stimuli. Environ. Microbiol. *10*, 893–905.
25. Liu, C., Mao, L., Zheng, X., Yuan, J., Hu, B., Cai, Y., Xie, H., Peng, X., and Ding, X. (2019). Comparative proteomic analysis of *Methanothermobacter thermautotrophicus* reveals methane formation from H_2 and CO_2 under different temperature conditions. MicrobiologyOpen *8*, e00715. https://doi.org/10.1002/mbo3.715.
26. Diender, M., Pereira, R., Wessels, H.J.C.T., Stams, A.J.M., and Sousa, D.Z. (2016). Proteomic Analysis of the Hydrogen and Carbon Monoxide Metabolism of *Methanothermobacter marburgensis*. Front. Microbiol. *7*, 1049. https://doi.org/10.3389/fmicb.2016.01049.
27. PATRIC (2014). *Methanothermobacter thermautotrophicus* Str. ΔH. https://www.patricbrc.org/view/Genome/187420.15.
28. Yishai, O., Lindner, S.N., Gonzalez de la Cruz, J., Tenenboim, H., and Bar-Even, A. (2016). The formate bio-economy. Curr. Opin. Chem. Biol. *35*, 1–9.
29. Feierabend, M., Renz, A., Zelle, E., Nöh, K., Wiechert, W., and Dräger, A. (2021). High-Quality Genome-Scale Reconstruction of Corynebacterium glutamicum ATCC 13032. Front. Microbiol. *12*, 750206.
30. Gu, C., Kim, G.B., Kim, W.J., Kim, H.U., and Lee, S.Y. (2019). Current status and applications of genome-scale metabolic models. Genome Biol. *20*, 121. https://doi.org/10.1186/s13059-019-1730-3.
31. Suthers, P.F., Foster, C.J., Sarkar, D., Wang, L., and Maranas, C.D. (2021). Recent advances in constraint and machine learning-based metabolic modeling by leveraging stoichiometric balances, thermodynamic feasibility and kinetic law formalisms. Metab. Eng. *63*, 13–33.
32. Lu, H., Li, F., Sánchez, B.J., Zhu, Z., Li, G., Domenzain, I., Marcišauskas, S., Anton, P.M., Lappa, D., and Lieven, C. (2019). A consensus *S. cerevisiae* metabolic model Yeast8 and its ecosystem for comprehensively probing cellular metabolism. Nat. Commun. *10*, 1–13.
33. Lu, H., Kerkhoven, E.J., and Nielsen, J. (2022). Multiscale models quantifying yeast physiology: towards a whole-cell model. Trends Biotechnol. *40*, 291–305.
34. Mendoza, S.N., Olivier, B.G., Molenaar, D., and Teusink, B. (2019). A systematic assessment of current genome-scale metabolic reconstruction tools. Genome Biol. *20*, 158.
35. Thiele, I., and Palsson, B.Ø. (2010). A protocol for generating a high-quality genome-scale metabolic reconstruction. Nat. Protoc. *5*, 93–121.
36. Baumler, D.J., Peplinski, R.G., Reed, J.L., Glasner, J.D., and Perna, N.T. (2011). The evolution of metabolic networks of *E. coli*. BMC Syst. Biol. *5*, 182.
37. Enzmann, F., Mayer, F., Rother, M., and Holtmann, D. (2018). Methanogens: biochemical background and biotechnological applications. Amb. Express *8*, 1.
38. Farkas, J.A., Picking, J.W., and Santangelo, T.J. (2013). Genetic techniques for the archaea. Annu. Rev. Genet. *47*, 539–561.
39. Fink, C., Martinez-Cano, G., Shuster, J., Panzera, A., Angenent, L.T., and Molitor, B. (2022). The targeted deletion of genes responsible for expression of the Mth60 fimbriae leads to loss of cell-cell connections in *M. thermautotrophicus* ΔH. Preprint at bioRxiv. https://doi.org/10.1101/2022.05.13.491833.
40. Fink, C., Beblawy, S., Enkerlin, A.M., Mühling, L., Angenent, L.T., and Molitor, B. (2021). A shuttle-vector system allows heterologous gene expression in the thermophilic methanogen *Methanothermobacter thermautotrophicus* ΔH. mBio *12*, e0276621.
41. Thor, S., Peterson, J.R., and Luthey-Schulten, Z. (2017). Genome-scale Metabolic Modeling of Archaea Lends Insight into Diversity of Metabolic Function. Archaea *2017*, 9763848.
42. Becker, S.A., and Palsson, B.Ø. (2008). Context-specific metabolic networks are consistent with experiments. PLoS Comput. Biol. *4*, e1000082.
43. White, R.H. (1997). Purine biosynthesis in the domain Archaea without folates or modified folates. J. Bacteriol. *179*, 3374–3377.
44. Wei, Y., Li, B., Prakash, D., Ferry, J.G., Elliott, S.J., and Stubbe, J. (2015). A ferredoxin disulfide reductase delivers electrons to the *Methanosarcina barkeri* class III ribonucleotide reductase. Biochemistry *54*, 7019–7028.
45. Brown, A.M., Hoopes, S.L., White, R.H., and Sarisky, C.A. (2011). Purine biosynthesis in archaea: variations on a theme. Biol. Direct *6*, 63.
46. Buchenau, B., and Thauer, R.K. (2004). Tetrahydrofolate-specific enzymes in *Methanosarcina barkeri* and growth dependence of this methanogenic archaeon on folic acid or p-aminobenzoic acid. Arch. Microbiol. *182*, 313–325.
47. Schauer, N.L., and Ferry, J.G. (1982). Properties of formate dehydrogenase in *Methanobacterium formicicum*. J. Bacteriol. *150*, 1–7.
48. Sawers, G., and Watson, G. (1998). A glycyl radical solution: oxygen-dependent interconversion of pyruvate formate-lyase. Mol. Microbiol. *29*, 945–954.
49. Tanner, R.S., McInerney, M.J., and Nagle, D.P. (1989). Formate auxotroph of *Methanobacterium thermoautotrophicum* Marburg. J. Bacteriol. *171*, 6534–6538.
50. Valgepea, K., de Souza Pinto Lemgruber, R., Abdalla, T., Binos, S., Takemori, N., Takemori, A., Tanaka, Y., Tappel, R., Köpke,

M., Simpson, S.D., et al. (2018). H_2 drives metabolic rearrangements in gas-fermenting Clostridium autoethanogenum. Biotechnol. Biofuels *11*, 55.
51. Chen, Y., and Nielsen, J. (2019). Energy metabolism controls phenotypes by protein efficiency and allocation. Proc. Natl. Acad. Sci. USA *116*, 17592–17597.
52. Schuchmann, K., and Müller, V. (2013). Direct and reversible hydrogenation of CO_2 to formate by a bacterial carbon dioxide reductase. Science *342*, 1382–1385.
53. Bertram, P.A., Karrasch, M., Schmitz, R.A., Böcher, R., Albracht, S.P., and Thauer, R.K. (1994). Formylmethanofuran dehydrogenases from methanogenic Archaea Substrate specificity, EPR properties and reversible inactivation by cyanide of the molybdenum or tungsten iron-sulfur proteins. Eur. J. Biochem. *220*, 477–484.
54. Watanabe, T., Pfeil-Gardiner, O., Kahnt, J., Koch, J., Shima, S., and Murphy, B.J. (2021). Three-megadalton complex of methanogenic electron-bifurcating and CO_2-fixing enzymes. Science *373*, 1151–1156.
55. Shlomi, T., Cabili, M.N., Herrgård, M.J., Palsson, B.Ø., and Ruppin, E. (2008). Network-based prediction of human tissue-specific metabolism. Nat. Biotechnol. *26*, 1003–1010.
56. Zur, H., Ruppin, E., and Shlomi, T. (2010). iMAT: an integrative metabolic analysis tool. Bioinformatics *26*, 3140–3142.
57. Bordbar, A., Mo, M.L., Nakayasu, E.S., Schrimpe-Rutledge, A.C., Kim, Y.M., Metz, T.O., Jones, M.B., Frank, B.C., Smith, R.D., Peterson, S.N., et al. (2012). Model-driven multi-omic data analysis elucidates metabolic immunomodulators of macrophage activation. Mol. Syst. Biol. *8*, 558.
58. Nilsson, A., Nielsen, J., and Palsson, B.O. (2017). Metabolic models of protein allocation call for the kinetome. Cell Syst. *5*, 538–541.
59. Birkel, G.W., Ghosh, A., Kumar, V.S., Weaver, D., Ando, D., Backman, T.W.H., Arkin, A.P., Keasling, J.D., and Martín, H.G. (2017). The JBEI quantitative metabolic modeling library (jQMM): a python library for modeling microbial metabolism. BMC Bioinf. *18*, 205.
60. Fang, X., Lloyd, C.J., and Palsson, B.O. (2020). Reconstructing organisms in silico: genome-scale models and their emerging applications. Nat. Rev. Microbiol. *18*, 731–743.
61. Thauer, R.K. (2012). The Wolfe cycle comes full circle. Proc. Natl. Acad. Sci. USA *109*, 15084–15085.
62. Smith, D.R., Doucette-Stamm, L.A., 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* ΔH: functional analysis and comparative genomics. J. Bacteriol. *179*, 7135–7155.
63. Rinke, C., Chuvochina, M., Mussig, A.J., Chaumeil, P.-A., Davín, A.A., Waite, D.W., Whitman, W.B., Parks, D.H., and Hugenholtz, P. (2021). A standardized archaeal taxonomy for the Genome Taxonomy Database. Nat. Microbiol. *6*, 946–959.
64. Liesegang, H., Kaster, A.-K., Wiezer, A., Goenrich, M., Wollherr, A., Seedorf, H., Gottschalk, G., and Thauer, R.K. (2010). Complete genome sequence of *Methanothermobacter marburgensis*, a methanoarchaeon model organism. J. Bacteriol. *192*, 5850–5851.
65. (1988). National Center for Biotechnology Information. https://www.ncbi.nlm.nih.gov/.
66. Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. *30*, 207–210.
67. Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D.J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M., et al. (2019). The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. *47*, D442–D450.
68. Deutsch, E.W., Csordas, A., Sun, Z., Jarnuczak, A., Perez-Riverol, Y., Ternent, T., Campbell, D.S., Bernal-Llinares, M., Okuda, S., and Kawano, S. (2016). The ProteomeXchange Consortium in 2017: Supporting the Cultural Change in Proteomics Public Data Deposition. Nucleic Acids Res. *45*, gkw936.
69. Deutsch, E.W., Bandeira, N., Sharma, V., Perez-Riverol, Y., Carver, J.J., Kundu, D.J., García-Seisdedos, D., Jarnuczak, A.F., Hewapathirana, S., Pullman, B.S., et al. (2020). The ProteomeXchange consortium in 2020: enabling 'big data'approaches in proteomics. Nucleic Acids Res. *48*, D1145–D1152.
70. Bergmann, F.T., Adams, R., Moodie, S., Cooper, J., Glont, M., Golebiewski, M., Hucka, M., Laibe, C., Miller, A.K., Nickerson, D.P., et al. (2014). COMBINE archive and OMEX format: one file to share all information to reproduce a modeling project. BMC Bioinf. *15*, 369.
71. Glont, M., Nguyen, T.V.N., Graesslin, M., Hälke, R., Ali, R., Schramm, J., Wimalaratne, S.M., Kothamachu, V.B., Rodriguez, N., Swat, M.J., et al. (2018). BioModels: expanding horizons to include more modelling approaches and formats. Nucleic Acids Res. *46*, D1248–D1253.
72. Malik-Sheriff, R.S., Glont, M., Nguyen, T.V.N., Tiwari, K., Roberts, M.G., Xavier, A., Vu, M.T., Men, J., Maire, M., Kananathan, S., et al. (2020). BioModels—15 years of sharing computational models in life science. Nucleic Acids Res. *48*, D407–D415.
73. Keating, S.M., Waltemath, D., König, M., Zhang, F., Dräger, A., Chaouiya, C., Bergmann, F.T., Finney, A., Gillespie, C.S., Helikar, T., et al. (2020). SBML Level 3: an extensible format for the exchange and reuse of biological models. Mol. Syst. Biol. *16*, e9110.
74. Lieven, C., Beber, M.E., Olivier, B.G., Bergmann, F.T., Ataman, M., Babaei, P., Bartell, J.A., Blank, L.M., Chauhan, S., Correia, K., et al. (2020). MEMOTE for standardized genome-scale metabolic model testing. Nat. Biotechnol. *38*, 272–276.
75. König, M. (2022). fbc_curation: Reproducibility of constraint-based models. https://fbc-curation.readthedocs.io/en/latest/index.html.
76. King, Z.A., Dräger, A., Ebrahim, A., Sonnenschein, N., Lewis, N.E., and Palsson, B.O. (2015). Escher: a web application for building, sharing, and embedding data-rich visualizations of biological pathways. PLoS Comput. Biol. *11*, e1004321.
77. Balch, W.E., Fox, G.E., Magrum, L.J., Woese, C.R., and Wolfe, R.S. (1979). Methanogens: reevaluation of a unique biological group. Microbiol. Rev. *43*, 260–296.
78. Jensen, J.N. (2001). Approach to steady state in completely mixed flow reactors. J. Environ. Eng. (New. York). *127*, 13–18.
79. Cold Spring Harbor Protocols (2009). CTAB extraction buffer. Cold Spring Harbor. http://cshprotocols.cshlp.org/content/2009/10/pdb.rec11984.full.
80. Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. *28*, 27–30. https://doi.org/10.1093/nar/28.1.27.
81. Seaver, S.M.D., Liu, F., Zhang, Q., Jeffryes, J., Faria, J.P., Edirisinghe, J.N., Mundy, M., Chia, N., Noor, E., Beber, M.E., et al. (2021). The ModelSEED Biochemistry Database for the integration of metabolic annotations and the reconstruction, comparison and analysis of metabolic models for plants, fungi and microbes. Nucleic Acids Res. *49*, D575–D588. https://doi.org/10.1093/nar/gkaa746.
82. UniProt Consortium (2019). UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. *47*, D506–D515. https://doi.org/10.1093/nar/gky1049.
83. Jeske, L., Placzek, S., Schomburg, I., Chang, A., and Schomburg, D. (2019). BRENDA in 2019: a European ELIXIR core data resource. Nucleic Acids Res. *47*, D542–D549. https://doi.org/10.1093/nar/gky1048.
84. Karp, P.D., Billington, R., Caspi, R., Fulcher, C.A., Latendresse, M., Kothari, A., Keseler, I.M., Krummenacker, M., Midford, P.E., Ong, Q., et al. (2019). The BioCyc collection of microbial genomes and metabolic pathways. Brief. Bioinform. *20*, 1085–1093. https://doi.org/10.1093/bib/bbx085.
85. Caspi, R., Billington, R., Keseler, I.M., Kothari, A., Krummenacker, M., Midford, P.E., Ong, W.K., Paley, S., Subhraveti, P., and Karp, P.D. (2020). The MetaCyc database of metabolic pathways and enzymes-a 2019 update. Nucleic Acids Res. *48*, D445–D453.
86. Norsigian, C.J., Pusarla, N., McConn, J.L., Yurkovich, J.T., Dräger, A., Palsson, B.O., and King, Z. (2020). BiGG Models 2020: multi-strain genome-scale models and expansion across the phylogenetic tree. Nucleic Acids Res. *48*, D402–D406.
87. Sayers, E.W., Beck, J., Bolton, E.E., Bourexis, D., Brister, J.R., Canese, K., Comeau, D.C., Funk, K., Kim, S., Klimke, W., et al. (2021). Database resources of the national center for biotechnology information. Nucleic Acids Res. *49*, D10–D17.
88. Satish Kumar, V., Ferry, J.G., and Maranas, C.D. (2011). Metabolic reconstruction of the archaeon methanogen *Methanosarcina Acetivorans*. BMC Syst. Biol. *5*, 28.
89. Benedict, M.N., Gonnerman, M.C., Metcalf, W.W., and Price, N.D. (2012). Genome-scale metabolic reconstruction and hypothesis testing in the methanogenic archaeon *Methanosarcina acetivorans* C2A. J. Bacteriol. *194*, 855–865.
90. Nazem-Bokaee, H., Gopalakrishnan, S., Ferry, J.G., Wood, T.K., and Maranas, C.D. (2016). Assessing methanotrophy and carbon fixation for biofuel production by *Methanosarcina acetivorans*. Microb. Cell Fact. *15*, 10.
91. Peterson, J.R., Thor, S., Kohler, L., Kohler, P.R.A., Metcalf, W.W., and Luthey-Schulten, Z. (2016). Genome-wide gene expression and RNA half-life measurements allow

predictions of regulation and metabolic behavior in *Methanosarcina acetivorans*. BMC Genom. *17*, 924.
92. Feist, A.M., Scholten, J.C.M., Palsson, B.Ø., Brockman, F.J., and Ideker, T. (2006). Modeling methanogenesis with a genome-scale metabolic reconstruction of *Methanosarcina barkeri*. Mol. Syst. Biol. *2*, 2006.0004.
93. Gonnerman, M.C., Benedict, M.N., Feist, A.M., Metcalf, W.W., and Price, N.D. (2013). Genomically and biochemically accurate metabolic reconstruction of *Methanosarcina barkeri* Fusaro, iMG746. Biotechnol. J. *8*, 1070–1079.
94. Hamilton, J.J., Calixto Contreras, M., and Reed, J.L. (2015). Thermodynamics and H_2 transfer in a methanogenic, syntrophic community. PLoS Comput. Biol. *11*, e1004364.
95. Goyal, N., Widiastuti, H., Karimi, I.A., and Zhou, Z. (2014). A genome-scale metabolic model of *Methanococcus maripaludis* S2 for CO_2 capture and conversion to methane. Mol. Biosyst. *10*, 1043–1054.
96. Richards, M.A., Lie, T.J., Zhang, J., Ragsdale, S.W., Leigh, J.A., and Price, N.D. (2016). Exploring hydrogenotrophic methanogenesis: a genome scale metabolic reconstruction of *Methanococcus maripaludis*. J. Bacteriol. *198*, 3379–3390.
97. Shoaie, S., Karlsson, F., Mardinoglu, A., Nookaew, I., Bordel, S., and Nielsen, J. (2013). Understanding the interactions between bacteria in the human gut through metabolic modeling. Sci. Rep. *3*, 2532.
98. Selkov, E., Maltsev, N., Olsen, G.J., Overbeek, R., and Whitman, W.B. (1997). A reconstruction of the metabolism of *Methanococcus jannaschii* from sequence data. Gene *197*, GC11–GC26.
99. Hucka, M., Bergmann, F.T., Dräger, A., Hoops, S., Keating, S.M., Le Novère, N., Myers, C.J., Olivier, B.G., Sahle, S., Schaff, J.C., et al. (2018). The systems biology markup language (SBML): Language specification for level 3 version 1 core. J. Integr. Bioinform. *15*, 20170081.
100. Olivier, B.G., and Bergmann, F.T. (2018). SBML level 3 package: flux balance constraints version 2. J. Integr. Bioinform. *15*, 20170082.
101. Hucka, M., and Smith, L.P. (2016). SBML level 3 package: groups, version 1 release 1. J. Integr. Bioinform. *13*, 290.
102. Carey, M.A., Dräger, A., Beber, M.E., Papin, J.A., and Yurkovich, J.T. (2020). Community standards to facilitate development and address challenges in metabolic modeling. Mol. Syst. Biol. *16*, e9235.
103. Juty, N., Le Novère, N., and Laibe, C. (2012). Identifiers. org and MIRIAM Registry: community resources to provide persistent identification. Nucleic Acids Res. *40*, D580–D586.
104. Mahamkali, V., McCubbin, T., Beber, M.E., Noor, E., Marcellin, E., and Nielsen, L.K. (2021). multiTFA: a Python package for multi-variate thermodynamics-based flux analysis. Bioinformatics *37*, 3064–3066.
105. Henry, C.S., DeJongh, M., Best, A.A., Frybarger, P.M., Linsay, B., and Stevens, R.L. (2010). High-throughput generation, optimization and analysis of genome-scale metabolic models. Nat. Biotechnol. *28*, 977–982.
106. Marcellin, E., Behrendorff, J.B., Nagaraju, S., DeTissera, S., Segovia, S., Palfreyman, R.W., Daniell, J., Licona-Cassani, C., Quek, L.-E., Speight, R., et al. (2016). Low carbon fuels and commodity chemicals from waste gases–systematic approach to understand energy metabolism in a model acetogen. Green Chem. *18*, 3020–3028.
107. Ebrahim, A., Lerman, J.A., Palsson, B.O., and Hyduke, D.R. (2013). COBRApy: constraints-based reconstruction and analysis for python. BMC Syst. Biol. *7*, 74.
108. Duboc, P., Schill, N., Menoud, L., Van Gulik, W., and Von Stockar, U. (1995). Measurements of sulfur, phosphorus and other ions in microbial biomass: influence on correct determination of elemental composition and degree of reduction. J. Biotechnol. *43*, 145–158.
109. Valgepea, K., Loi, K.Q., Behrendorff, J.B., Lemgruber, R.d.S.P., Plan, M., Hodson, M.P., Köpke, M., Nielsen, L.K., and Marcellin, E. (2017). Arginine deiminase pathway provides ATP and boosts growth of the gas-fermenting acetogen *Clostridium autoethanogenum*. Metab. Eng. *41*, 202–211.
110. Fagbadebo, F.O., Kaiser, P.D., Zittlau, K., Bartlick, N., Wagner, T.R., Froehlich, T., Jarjour, G., Nueske, S., Scholz, A., Traenkle, B., et al. (2022). A Nanobody-Based Toolset to Monitor and Modify the Mitochondrial GTPase Miro1. Front. Mol. Biosci. *9*, 835302. https://doi.org/10.3389/fmolb.2022.835302.
111. Broadbent, J.A., Broszczak, D.A., Tennakoon, I.U.K., and Huygens, F. (2016). Pan-proteomics, a concept for unifying quantitative proteome measurements when comparing closely-related bacterial strains. Expert Rev. Proteomics *13*, 355–365.
112. Maillet, N. (2020). Rapid Peptides Generator: fast and efficient in silico protein digestion. NAR Genom. Bioinform. *2*, lqz004.
113. Heirendt, L., Arreckx, S., Pfau, T., Mendoza, S.N., Richelle, A., Heinken, A., Haraldsdóttir, H.S., Wachowiak, J., Keating, S.M., Vlasov, V., et al. (2019). Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v. 3.0. Nat. Protoc. *14*, 639–702.
114. Gurobi Optimization LLC (2020). Gurobi Optimizer Reference Manual. https://www.gurobi.com.
115. Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H., and Phillippy, A.M. (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. *27*, 722–736.
116. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078–2079.
117. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., and Li, H. (2021). Twelve years of SAMtools and BCFtools. GigaScience *10*, giab008.
118. Kuznetsov, A., and Bollin, C.J. (2021). NCBI genome workbench: desktop software for comparative genomics, visualization, and GenBank data submission. Methods Mol. Biol. *2231*, 261–295.
119. Li, T., Zhang, X., Luo, F., Wu, F.-X., and Wang, J. (2020). MultiMotifMaker: a multi-thread tool for identifying DNA methylation motifs from Pacbio reads. IEEE/ACM Trans. Comput. Biol. Bioinform. *17*, 220–225.
120. Cock, P.J.A., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., and de Hoon, M.J.L. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics *25*, 1422–1423.
121. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009). BLAST+: architecture and applications. BMC Bioinf. *10*, 421–429.
122. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389–3402.
123. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.
124. Madden, T. (2002). The BLAST Sequence Analysis Tool. In The NCBI Handbook, J. McEntyre and J. Ostell, eds. (National Center for Biotechnology Information (US)).
125. Leimbach, A. (2016). Bac-Genomics-Scripts: Bovine *E. coli* Mastitis Comparative Genomics Edition. https://github.com/aleimba/bac-genomics-scripts/blob/master/README.md#citation.
126. Galperin, M.Y., Wolf, Y.I., Makarova, K.S., Vera Alvarez, R., Landsman, D., and Koonin, E.V. (2021). COG database update: focus on microbial diversity, model organisms, and widespread pathogens. Nucleic Acids Res. *49*, D274–D281.
127. Mojica Prieto, F.J., and Millero, F.J. (2002). The values of pK1+ pK2 for the dissociation of carbonic acid in seawater. Geochem. Cosmochim. Acta *66*, 2529–2540.
128. Wang, N.S., and Stephanopoulos, G. (1983). Application of macroscopic balances to the identification of gross measurement errors. Biotechnol. Bioeng. *25*, 2177–2208.
129. Michael, L.S., and Kargi, F. (2002). Bioprocess Engineering: Basic Concepts (Prentice-Hall International).
130. Bryant, C. (2018). ResearchPy. https://github.com/researchpy/researchpy.
131. Bushnell, B. (2014). BBMap: A Fast, Accurate, Splice-Aware Aligner (Lawrence Berkeley National Lab).
132. Gordon, A., and Hannon, G. (2017). Fastx-toolkit. FASTQ/A Short-Reads Pre-processing Tools. 2010. http://hannonlab.cshl.edu/fastx_toolkit.
133. Li, H. (2013). Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM. Preprint at arXiv. https://doi.org/10.48550/arXiv.1303.3997.
134. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930.
135. Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods *14*, 417–419.
136. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.
137. Waardenberg, A.J., and Field, M.A. (2019). consensusDE: an R package for assessing consensus of multiple RNA-seq algorithms with RUV correction. PeerJ *7*, e8206.

138. Mölder, F., Jablonski, K.P., Letcher, B., Hall, M.B., Tomkins-Tinch, C.H., Sochat, V., Forster, J., Lee, S., Twardziok, S.O., and Kanitz, A. (2021). Sustainable Data Analysis with Snakemake. F1000Research *10*.

139. Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337–342.

140. Wagner, T., Watanabe, T., and Shima, S. (2019). Hydrogenotrophic Methanogenesis. In Biogenesis of Hydrocarbons, A.J. Stams and D. Sousa, eds. (Springer Cham), pp. 79–107.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Methanothermobacter thermautotrophicus ΔH	DSMZ	DSM 1053
Methanothermobacter thermautotrophicus Z-245	DSMZ	DSM 3720
Methanothermobacter marburgensis Marburg	DSMZ	DSM 2133
Methanothermobacter thermautotrophicus ΔH pMVS1111A:PhmtB-fdhZ-245	Bastian Molitor, Fink et al.[40]	https://doi.org/10.1128/mBio.02766-21
Chemicals, peptides, and recombinant proteins		
Anti Foam SE-15	Sigma-Aldrich	A8582
ROTI®Phenol/Chloroform/Isoamyl alcohol	Carl Roth	Cat# A156.2
Vacuum grease Dow Corning®	VWR International GmbH	Cat# 44224.KT
Chloroform/Isoamyl alcohol (24:1)	VWR International GmbH	Cat# SERA39554.01
SYBR™ Safe DNA Gel Stain	Thermo Fischer Scientific	Cat# S33102
RNAlater™ Stabilization Solution	Thermo Fischer Scientific	Cat# AM7021
Critical commercial assays		
Genomic DNA 165 kb Kit	Agilent Technologies	Cat# FP-1002-0275
Qubit® dsDNA HS Assay Kit	Thermo Fischer Scientific	Cat# Q32851
SMRTbell® Express Template Preparation Kit	Pacific Biosciences	PN 101-397-100 Version 3
Thermo Scientific™ Phire Plant Direct PCR Master Mix	Thermo Fischer Scientific	Cat# F160S
RNA Clean & Concentrator Kit	Zymo Research	Cat# R1019
Agilent RNA 6000 Nano Kit	Agilent Technologies	Cat# 5067-1511
TruSeq™ Stranded Total RNA Kit	Illumina	Cat# 20020597
Illumina Ribo-Zero Plus rRNA Depletion Kit	Illumina	Cat# 20037135
Deposited data		
Supplemental data sets Data S1, S3–S7	This manuscript	https://github.com/isacasini/Casini_GEM_2023
Raw and analyzed genome sequencing data	This manuscript	NCBI Bioproject: PRJNA674001
Raw and analyzed RNA sequencing data	This manuscript	GEO: GSE218145
Raw and analyzed Proteomics data	This manuscript	PRIDE: PXD039574
GEM modeling related files *M. thermautotrophicus ΔH*	This manuscript	Biomodels: MODEL2211290001
GEM modeling related files *M. thermautotrophicus Z-245*	This manuscript	Biomodels: MODEL2211290002
GEM modeling related files *M. marburgensis Marburg*	This manuscript	Biomodels: MODEL2211290003
Experimental models: Organisms/strains		
M. thermautotrophicus ΔH genome sequence	NCBI, Smith et al.[62]	NC_000916.1
M. marburgensis Marburg genome sequence	NCBI, Liesegang et al.[64]	NC_014408.1
Oligonucleotides		
Mth_RS05590_selec_FW2_opt	This manuscript	*M. thermautotrophicus* ΔH Forward
Mth_RS05590_selec_RV_2_opt	This manuscript	*M. thermautotrophicus* ΔH Reverse
Seq_specific_MtZ_pFZ1_FW	This manuscript	*M. thermautotrophicus* Z-245 Forward
Seq_specific_MtZ_pFZ1_RV	This manuscript	*M. thermautotrophicus* Z-245 Reverse
Primer_pME2001_XhoI_FW	Bastian Molitor, Fink et al.[40]	*M. marburgensis* Marburg Forward
pME2001_Rev_1440_1000Xho	Bastian Molitor, Fink et al.[40]	*M. marburgensis* Marburg Reverse

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Canu Versions 1.7 & 2.0	Koren et al.[115]	https://github.com/marbl/canu; RRID: SCR_015880
Smrttools Version 9.0.0.92188	Pacific Biosciences	https://downloads.pacbcloud.com/public/software/installers/smrtlink_9.0.0.92188.zip
Samtools Versions 1.6 & 1.11	Li et al.[116],Danecek et al.[117]	http://www.htslib.org/
Blast *via* Genome Workbench 3.5	Kuznetsov and Bollin[118]	https://ftp.ncbi.nlm.nih.gov/toolbox/gbench/ver-3.5.0/
Biopython Version 1.77	Cock et al.[120]	http://biopython.org/DIST/biopython-1.77.zip
MotifMaker Version 0.3.1	Li et al.[119]	https://github.com/bioinfomaticsCSU/MultiMotifMaker
Smrtlink Version 10.2.0.133434	Pacific Biosciences	https://downloads.pacbcloud.com/public/software/installers/smrtlink_10.2.1.143962.zip
BaseModFunction.v2.1.R	Pacific Biosciences	https://github.com/PacificBiosciences/Bioinformatics-Training/tree/master/basemods (*link no longer works*)
Basic Local Alignment Search Tool (BLASTp+) Version 2.10.0	Camacho et al.[121],Altschul et al.[122],Altschul et al.[123]	https://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/2.10.0/
cdd2cog from bac-genomics-scripts Version v0.2	Leimbach[125]	https://github.com/aleimba/bac-genomics-scripts
Microsoft® Excel® Version 2202	Microsoft 365 MSO	https://www.office.com/
cxcalc from MarvinSketch Version 18.8.0	ChemAxon	https://www.chemaxon.com
MEMOTE Version 0.13.0	Lieven et al.[74]	https://github.com/opencobra/memote
multiTFA	Mahamkali et al.[104]	https://github.com/biosustain/multitfa
Gurobi Optimizer v.9.0.1	Gurobi Optimization LLC[114]	https://www.gurobi.com/downloads/gurobi-software/
COBRApy version 0.22.1	Ebrahim et al.[107]	https://github.com/opencobra/cobrapy
Escher	King et al.[76]	https://escher.github.io/#/
Python 3.6.13	Python Software Foundation	https://www.python.org/downloads/release/python-3613/
Researchpy package Version 0.3.2	Bryant[130]	https://github.com/researchpy/researchpy
Illumina bcl2fastq Version 2.20	Illumina	https://emea.support.illumina.com/downloads/bcl2fastq-conversion-software-v2-20.html
BBMap Version 38.93	Bushnell[131]	https://anaconda.org/bioconda/bbmap/files?sort=uploader.name&sort_order=desc&version=38.93
FastX Toolkit Version 0.0.14	Gordon and Hannon[132]	http://hannonlab.cshl.edu/fastx_toolkit/download.html
BWA-MEM Version 0.7.17	Li[133]	https://github.com/lh3/bwa
Subread package Version 2.0.1	Liao et al.[134]	https://github.com/ShiLab-Bioinformatics/subread
Salmon Version 1.5.2	Patro et al.[135]	https://github.com/COMBINE-lab/salmon/releases
DESeq2 Version 1.32.0	Love et al.[136]	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
ConsensusDE Version 1.10.0	Waardenberg and Field[137]	https://anaconda.org/bioconda/bioconductor-consensusde/files?version=
Snakemake Version 6.8.0	Mölder et al.[138]	https://github.com/snakemake/snakemake
Proteome Discoverer Software Version 2.5.0.400	Thermo Fisher Scientific	https://www.thermofisher.com/de/en/home/industrial/mass-spectrometry/liquid-chromatography-mass-spectrometry-lc-ms/lc-ms-software/multi-omics-data-analysis/proteome-discoverer-software.html

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Rapid Peptides Generator Version 1.2.4	Maillet[112]	https://gitlab.pasteur.fr/nmaillet/rpg
GIMME	Becker and Palsson[42]	https://github.com/opencobra/cobratoolbox/tree/master/src/dataIntegration/transcriptomics/GIMME
COBRA Toolbox v.3.1	Heirendt et al.[113]	https://github.com/opencobra/cobratoolbox
MATLAB Version R2018b	The MathWorks, Inc.	https://de.mathworks.com/products/matlab.html
Other		
UNIlab Pro Eco glovebox	MBraun	*Not listed*
Centrifuge 5920 R	Eppendorf	Cat# 5948000010
Vortex-Genie® 2	VWR International GmbH	Cat# 444-0873P
Centrifuge 5424	Eppendorf	*No longer sold*
ThermoMixer® C	Eppendorf	Cat# 5382000015
Femto Pulse System 1.0.0.32	Agilent	Cat# M5330AA
Qubit® 2.0 Fluorometer	Invitrogen	*No longer sold*
Sequel I System	Pacific Biosciences	*No longer sold*
BioXplorer 100 bioreactor & WinISO version 2.3.149.1 software	H.E.L	https://helgroup.com/products/bioreactors/bioxplorer-100/
pH sensors	I&L Biosystems GmbH	Cat# Z001013510
ORP sensors	I&L Biosystems GmbH	Cat# Z061013510
Red-y smart min - mass-flow controller	Vögtlin	https://www.voegtlin.com/massedurchflussmesser-und-massedurchflussregler/red-y-smart-series/
Bronkhorst EL-Flow® Prestige mass flow controllers	Bronkhorst Deutschland Nord GmbH	https://www.bronkhorst.com/int/products/gas-flow/el-flow-prestige/
MilliGascounter MGC-1 V3.4 PMMA	Dr.-Ing. RITTER Apparatebau GmbH & Co. KG	https://www.ritter.de/en/milligascounter-mgc-1-pmma/
FiveEasy™ Plus pH/mV Benchtop meter	Mettler-Toledo GmbH	Cat#: 30266627
Micro pH electrode LE422	Mettler-Toledo GmbH	Cat#: 30089747
Mastercycler® pro S	Eppendorf	*No longer sold*
Gel Doc™ XR+ visualizer	Bio-Rad	https://www.bio-rad.com/de-de/product/gel-doc-xr-gel-documentation-system?ID=O494WJE8Z
490 Micro Gas Chromatograph	Agilent	*No longer sold (G3581-90001)*
High-Pressure Liquid Chromatographer (HPLC) Nexera Series SIL-40C	Shimadzu Europa	https://www.ssi.shimadzu.com/products/liquid-chromatography/hplcuhplc/nexera-series/index.html
Multifuge X3R with the rotor TX-1000	Fischer Scientific	*No longer sold*
Lysing Matrix B	MP Biomedicals Germany GmbH	Cat# 116911050-CF
FastPrep-24TM 5G Bead-Beater	MP Biomedicals Germany GmbH	Cat# 116005500
2100 Agilent Bioanalyzer	Agilent Technologies	Cat# G2939BA
NovaSeq™ 6000	Illumina	https://emea.illumina.com/systems/sequencing-platforms/novaseq.html
Easy-nLC™ 1200	Thermo Fisher Scientific	Cat# LC140
Orbitrap Exploris™ 480 Mass Spectrometer	Thermo Fisher Scientific	Cat# BRE725539

RESOURCE AVAILABILITY

Lead contact

Further information and request for resources should be directed to Bastian Molitor, bastian.molitor@uni-tuebingen.de.

Materials availability

This study did not generate new unique materials.

Data and code availability

- Supplemental tables can be found in Data_S1.xlsx. The **genome** sequences, annotations, and methylation patterns were deposited to NCBI[65] and can be found under Bioproject ID PRJNA674001 (Data S1 - Table S1). The **transcriptomic** data (gene expression data) discussed in this publication have been deposited in NCBI's Gene Expression Omnibus[66] and are accessible through GEO Series accession number GSE218145 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE218145). The mass spectrometry **proteomic** DDA data were deposited to the ProteomeXchange Consortium *via* the PRIDE partner repository[67–69] with the dataset identifier PXD039574. The **modeling**-related files (Data S2) were deposited in an Open Modeling EXchange format (OMEX)[70] to BioModels[71,72] and assigned the identifiers MODEL2211290001 (*M. thermautotrophicus* ΔH), MODEL2211290002 (*M. thermautotrophicus* Z-245), and MODEL2211290003 (*M. marburgensis* Marburg). The pan-model is provided as an Excel® model (.xlsx, densely commented). The GEMs for each strain are provided in the following formats: .xml (SBML[73]), .json, and .mat. Additional files include: the benchmarking reports (.html) from MEMOTE,[74] the FROG analysis files (.omex) from fbc_curation,[75] and the map files for visualization (.json) with Escher.[76] Further, the data used to constrain the models are provided as supplemental data**.** This includes: the gas fermentation data (as Excel® files, Data S3 and S6) and the individual strain and differential expression analyses data for both the transcriptomics and proteomics (as an Excel® file, Data S4). Data S1 and S3–S6 are available on GitHub at https://github.com/isacasini/Casini_2023_GEM.
- Additional original code (scripts and programs) are found in Data S7. Data S7 is available on GitHub at https://github.com/isacasini/Casini_2023_GEM.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Microbial strains

M. thermautotrophicus ΔH (DSM 1053), *M. thermautotrophicus* Z-245 (DSM 3720), and *M. marburgensis* Marburg (DSM 2133) were obtained from the DSMZ (Braunschweig, Germany). *M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$ was previously generated in our lab.[40]

Growth media

Preparation of batch and continuous media for bioreactor runs was adjusted from the mineral media of Balch et al.[77] and Martin et al.[15] The mineral medium contained (per liter): nitrilotriacetic acid (NTA), 0.096 g; trisodium nitrilotriacetate, 0.275 g; sodium chloride, 0.675 g; di-potassium hydrogen phosphate, 0.255 g; ammonium chloride, 2.006 g; magnesium chloride hexahydrate, 0.12 g; calcium chloride dihydrate, 0.090 g; potassium di-hydrogen phosphate, 0.345; ammonium nickel sulfate, 1.5 mL (0.2% w/v); iron(II) chloride tetrahydrate, 1.5 mL (0.2% w/v); resazurin indicator solution, 4 mL (0.025% w/v); and trace element solution, 1.5 mL. The trace element solution was prepared 10-fold as stated by Balch et al.[77] with minor modifications, and contained (per liter): nitrilotriacetic acid (NTA), 2.0 g; magnesium sulfate heptahydrate, 30.0 g; manganese(II) sulfate, 5.0 g; sodium chloride, 10.0 g; iron(II) sulfate heptahydrate, 1.0 g; cobalt(II) chloride hexahydrate, 1.8 g; calcium chloride dihydrate, 1.0 g; zinc sulfate heptahydrate, 1.8 g; copper sulfate pentahydrate, 0.1 g; aluminum potassium sulfate dodecahydrate, 0.18 g; boric acid, 0.1 g; sodium molybdate dihydrate, 0.1 g; ammonium nickel(II) sulfate hexahydrate, 2.8 g; sodium tungstate dihydrate, 0.1; and sodium selenate, 0.1 g. The pH of the trace element solution was adjusted to 6.0 with 5 M potassium hydroxide. In all the continuous media, 0.02 mL/L of Anti Foam SE-15 (Sigma-Aldrich, Merck, Darmstadt, Germany) was supplemented. For growth with sodium formate, serum bottles for batch cultivation were sparged with N_2/CO_2 (80/20 %, v/v), and 100 mM sodium formate was added after autoclaving. Further, the serum bottles were supplemented with 0.001 mM sodium selenate and 0.01 mM sodium molybdate dihydrate. For growth with sodium formate, the continuous media contained 355 $\pm$ 5 mM sodium formate and was supplemented with 0.0015 mM sodium selenate and 0.015 mM sodium molybdate dihydrate (final pH was 5.85). A concentrated sodium formate solution was prepared and sterilely added to the continuous media after autoclaving, sparging, and reducing the media. All media were prepared with Millipore water (18.2 MΩ·cm). The medium was autoclaved either in Schott bottles with butyl septa with Masterflex® L/S Norprene Food-Grade Tubing, L/S 14 tubing (Cole-Parmer GmbH, Wertheim, Germany) or directly in the bioreactor vessels.

For anaerobic handling of media and materials the atmosphere of the anaerobic chamber (UniLab Pro Eco, MBraun, Garching, Germany) contained 95% v/v N_2 and 5% v/v CO_2.

Growth conditions

Cultivations in serum bottles for all strains were essentially as described in Fink et al.[40] Continuous fermentations were carried out in the BioXplorer 100 bioreactor platform controlled with the WinISO version 2.3.149.1 software (H.E.L, London, England). Each bioreactor was

equipped with temperature, pH (part number Z001013510), and ORP (part number Z061013510) sensors (I&L Biosystems GmbH, Königswinter, Germany); a 0.15 μm sparging stone; a magnetic coupled stirring system; three peristaltic pumps for media feed-in, base feed-in (1 N NaOH was used for pH control), and effluent-out; a mass-flow controller (Red-y smart min; Vögtlin, Muttenz, Switzerland) to control the inlet gas flow rate; a condenser for the exhaust-gas line; and a separate sampling and inoculum port (fitted with a rubber butyl stopper). The bioreactors were fitted with Masterflex® L/S Norprene Food-Grade Tubing, L/S 14 (Cole-Parmer GmbH, Wertheim-Mondfel, Germany) except for the gas inlet lines, which were fitted with Masterflex® C-Flex ULTRA tubing, L/S 16 (Cole Parmer, Wertheim-Mondfel, Germany). The upstream gas mixture was set using Bronkhorst EL-Flow® Prestige mass flow controllers (Bronkhorst Deutschland Nord GmbH, Kamen, Germany) and mixed in a doubled-ended cylinder (Swagelok® Stuttgart, Reutlingen, Germany). The exhaust gas flow rate was measured offline using a MilliGascounter MGC-1 V3.4 PMMA (Dr.-Ing. RITTER Apparatebau GmbH & Co. KG, Bochum, Germany). The pH and ORP sensors, the pumps, and the MFCs were calibrated before each experiment. The bioreactors were filled with the mineral medium and autoclaved for one hour at 121°C. Afterward, the bioreactors were connected to the bioreactor platform, and the temperature was set to 65°C with agitation at 700 rpm.

For the first bioreactor experiment, the bioreactors were then sparged through sterile filters Minisart® HY (0.2 μm pore size; Sartorius AG, Göttingen, Germany) for two hours with H_2/CO_2 (80/20% v/v) at a gas flow rate of 10 mL/min. Before inoculation, the mineral medium was reduced with sterile anaerobic L-cysteine-HCl (0.5 g/L) and disodium sulfide nonahydrate (0.3 g/L), and the pH control was set to 7.3. Each bioreactor was inoculated with either 4 mL (*M. thermautotrophicus* ΔH and *M. thermautotrophicus* Z-245) or 3.6 mL (*M. marburgensis* Marburg) of preculture grown in serum bottles to an OD_{600} of 0.35-0.36. The bioreactors were operated in batch mode for one day until an OD_{600} of approximately 1.00 was reached, at which point continuous mode was started with a medium feed at a dilution rate of 0.83 d^{-1}. The first continuous operating period was conducted for approximately 12 days, after which the bioreactors were emptied except for about 3-5 mL, which were used as inoculum for a second period with a starting OD_{600} of 0.20-0.30. When an OD_{600} of approximately 1.00 was reached, the bioreactors were switched into continuous mode with a medium feed at a dilution rate of 1.11 d^{-1}. Steady-state was reached after three hydraulic retention times (HRT),[78] which was 2.7 days in our setup. After an additional 3 HRTs (day 6.8), the transcriptomics, proteomics, and gram cell-dry weight determination samples were taken.

For the second bioreactor experiment, the bioreactor setup was the same as in the first experiment, except for the following adjustments for bioreactors fed with sodium formate: 1) pH adjusted using 1 N HCl instead of 1 N NaOH; and 2) sparged with N_2/CO_2 (80/20 %, v/v) rather than H_2/CO_2 (80/20 %, v/v). Each bioreactor was inoculated with 6 mL (*M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$ and *M. thermautotrophicus* Z-245) of preculture grown in serum bottles (OD_{600} ~0.2). The bioreactors were operated in batch mode until an OD_{600} of approximately 1.00 for bioreactors grown on H_2 and CO_2 and 0.10-0.15 for bioreactors grown on sodium formate was reached. At this point, the continuous mode was started with a media feed at a dilution rate of ~1.00 d^{-1} (this rate was achieved by ramping over several days for the bioreactors grown on sodium formate). After maintaining steady-state for three HRTs (10.6 days for H_2/CO_2 and 16.8 days for sodium formate), samples for transcriptomics, proteomics, and gram cell-dry weight determination were taken.

Daily samples of OD_{600}, pH, exhaust gas flow rate, and inlet and exhaust gas composition were taken during both experiments and periods. For the liquid culture samples, 1 mL of dead volume was first removed, and then 1 mL was used for OD_{600} and pH measurements. To measure pH, the samples were maintained at 65°C (ThermoMixer® C 460-0223, Eppendorf, Hamburg, Germany) and measured within 1 min of taking the sample (FiveEasy™ Plus pH/mV Benchtop meter with the micro pH electrode LE422 [Mettler-Toledo GmbH, Gießen, Germany] calibrated at 25°C and set to 65°C).

METHOD DETAILS

Genome sequencing

M. thermautotrophicus ΔH, *M. thermautotrophicus* Z-245, and *M. marburgensis* Marburg were grown in mineral medium overnight. For sequencing, we used strains directly from the DSMZ without excessive subcultivation. The entire biomass (50 mL) was collected by centrifugation at 3170 × g and room temperature for 20 min (5920 R Eppendorf, Hamburg, Germany), and the genomic DNA was extracted using a phenol-chloroform extraction method. For this, to the biomass pellet, 500 μL of cetyltrimethylammonium bromide buffer (CTAB, made according to Cold Spring Harbor Protocols,[79] however, without the Polyvinylpyrrolidone) was added, and the pellet was resuspended. The mixture was transferred to sterile 2 mL bead beating tubes (containing 500 μL of 0.1 μm BeadBeater® zirconia beads, Carl Roth, Karlsruhe, Germany) and vortexed (Vortex-Genie® 2, VWR International GmbH, Darmstadt, Germany) at 2700 min^{-1} for 5 s on, then 1 s off, repeatedly for 1 min, after which the tube was placed on ice. 500 μL of ROTI®Phenol/Chloroform/Isoamyl alcohol (ratio of 25:24:1, Carl Roth, Karlsruhe, Germany) was added, the tube inverted, and then centrifuged at 4°C and 16000 × g for 10 min (5424, Eppendorf, Hamburg, Germany). The top layer was transferred (using wide orifice pipette tips) to a phase lock tube, which was prepared ahead by adding 2 mm^3 of vacuum grease Dow Corning® (VWR International GmbH, Darmstadt, Germany) into 2 mL tubes, which were centrifuged until 9391 × g was reached, and then autoclaved to sterilize. Another 500 μL of ROTI®Phenol/Chloroform/Isoamyl alcohol was added, the tube was inverted to mix, and centrifugation was performed as in the previous step. The supernatant was transferred to new tubes, 500 μL of Chloroform/Isoamyl alcohol (24:1, VWR International GmbH, Darmstadt, Germany) was added, the tube was inverted to mix, and centrifugation was performed as in the previous step. The supernatant was added to a fresh tube, and the gDNA was precipitated by adding 0.1 volumes of cold 3 M sodium acetate and 2 volumes of ice-cold absolute ethanol, inverting the tube to mix, and then incubating it overnight at -20°C. The following day, the tube was centrifuged again at 4°C and 16000 × g for 10 min, the supernatant was removed, and the gDNA was washed with 300 μL of ice-cold 70% v/v ethanol. Centrifugation was performed as in the previous step, the supernatant was removed, and the tube was air-dried for

approximately 1.5 h at 50°C in a ThermoMixer® C (Eppendorf, Hamburg, Germany). The pellet was resuspended with 44 μL elution buffer (10 mM Tris·Cl, pH 8.5, and nuclease-free water [New England Biolabs, Ipswich, United States]) and allowed to rest for an hour to resolve at room temperature. To remove RNA, 1 μL of Bovine Ribonuclease A (VWR International GmbH, Darmstadt, Germany) was added, and the tube was allowed to rest for 30 min at room temperature. The quality of the gDNA was then checked using the Femto Pulse System (1.0.0.32, Agilent, Santa Clara, United States) according to the manufacturer's instructions, using the Genomic DNA 165 kb Kit (Agilent, Santa Clara, United States) and a 70 min separation time. The quantity of double-stranded DNA was measured with the Qubit® 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA) using the Qubit® dsDNA HS Assay Kit (Thermo Fischer Scientific, Dreieich, Germany). Library preparation was conducted with the SMRTbell® Express Template Preparation Kit (PN 101-397-100 Version 3, January 2018, Pacific Biosciences, Menlo Park, United States) as specified in the manufacturer's instructions. The genomes were then sequenced using the Sequel I System (Pacific Biosciences, Menlo Park, United States).

Genome-scale metabolic model reconstruction

A pan-model for *M. thermautotrophicus* ΔH, *M. thermautotrophicus* Z-245, and *M. marburgensis* Marburg, which included all reactions from the three strains, was built in Microsoft® Excel® (Microsoft 365 MSO, Version 2202, Washington, United States), following the protocol established by Thiele and Palsson.[35] This pan-model consists of 618 reactions (including 46 exchange reactions, 56 transport reactions, and seven biomass-associated reactions, with eleven transport and eleven exchange reactions that act as pseudo reactions for orphan metabolites to refer to compounds that are either only produced or consumed), 555 metabolites, and 545 genes (Data S2). The pan-model used genome sequences, assemblies, and annotations as its backbone. Reactions and pathways were added based on data from KEGG,[80] ModelSEED,[81] UniProt,[82] Brenda,[83] BioCyc,[84] MetaCyc,[85] BIGG,[86] and NCBI.[87] Metabolite protonation was determined (for pH 7.0) using the command line tool cxcalc and MarvinSketch 18.8.0, ChemAxon (https://www.chemaxon.com, accessed in 2018).

When possible, reactions and genes were verified with literature. Genes for which no genus-specific evidence was found were BLAST searched to genes of species with stronger evidence. These BLAST results are specified in the comment section of the reconstruction. Over 790 references were cited for which the microbe and type of evidence were recorded using Evidence & Conclusion Ontology (categories: biochemical, genetic, physiological, sequence, modeling, and no data), which were then used to determine the confidence of each reaction for each strain.[35] Published GEMs from the following strains were used to gap-fill and validate pathways in the *Methanothermobacter* pan-model: 1) *Methanosarcina acetivorans*: *i*VS941,[88] *i*MB745,[89] *i*MAC868a,[90] *i*ST807;[91] 2) *Methanosarcina barkeri*: *i*AF692,[92] *i*MG746;[93] 3) *Methanospirillum hungatei*: *i*Mhu428;[94] 4) *Methanococcus maripaludis* S2: *i*MM518;[95,96] 5) *Methanobrevibacter smithii*: *i*Msi385;[97] 6) *Methanocaldococcus jannaschii*: *i*TS436.[98]

The pan-model was converted to three strain-specific GEMs written in SBML Level 3 Version 1,[99] including the extension packages for flux balance constraints (fbc) version 2[100] and groups,[101] and verified in MEMOTE 0.13.0.[74] We followed recent recommendations for best practices, including for the GEM nomenclature (*i*MTD22IC, *i*MTZ22IC, *i*MMM22IC).[102] For verification in MEMOTE, the GEMs were constrained with Conditions 1, 8, and 15 for *M. thermautotrophicus* ΔH, *M. thermautotrophicus* Z-245, and *M. marburgensis* Marburg, respectively (Data S2 and S5). Each model scored 85% (Data S2).[74] The GEMs aimed to be MIRIAM-compliant, including metabolite and reaction annotations with Compact Identifiers[103] for various databases (gene annotations were not available given the newly annotated genomes; however, the old gene annotations can be found in Data S2).

The directionality of the reactions was determined using thermodynamics-based flux variability analysis (TFVA)[104] and the ModelSEED database.[105] For the TFVA method, a sink reaction (reversible uptake and production possible) was added to each metabolite in the GEMs. A core model was then generated using the tmodel function with the following constraints for cytosol and extracellular: 1) pH (7.6, 7.3); 2) ionic strength (0.1, 0); 3) temperature (338.15 K, 338.15 K); and 4) membrane potentials ([0,150], [-150, 0] for [cytosol, extracellular]). The minimum and maximum flux values were found with Gurobi as the solver, using the box or univariate method.[104] If the flux values spanned zero, the reaction bounds were set to (-1000, 1000). If both the minimum and maximum flux values were less than zero, the reaction bounds were set to (-1000, 0). Lastly, if both the minimum and maximum flux values were greater than zero, the reaction bounds were set to (0, 1000). If directionality could not be determined with either TFVA or the ModelSEED database, directionality was set to reversible, except for reactions that caused loops with ATP (*e.g.*, with ATP losing a phosphate group,[35] or between redox carriers).[106]

Seven constrain conditions per strain were designed that depict relevant cultivation conditions: Conditions 1-7, Conditions 8-14, and Conditions 15-21 for *M. thermautotrophicus* ΔH, *M. thermautotrophicus* Z-245, and *M. marburgensis* Marburg, respectively (Data S5). Flux balance analysis simulations with maximization of biomass exchange (EX_biomass_e) as the objective function and these different conditions were performed to verify the ability of the three strain-specific GEMs to grow under these cultivation conditions (Data S5 [Simulations 1-24]). The flux balance analysis simulations were run using COBRApy version 0.22.1 (Data S5 and S7).[107] We produced metabolic pathway maps to visualize the GEMs with Escher (Data S2).[76]

Biomass and maintenance energy determination

The biomass composition for the three strains in the GEMs is assumed to be the same (Data S2). Briefly, the fraction for each of the molecules that make up the biomass was found by averaging those from previously published methanogen GEMs[41] and incorporating empirically found data for *M. thermautotrophicus* ΔH (reaction ID, BIOMASS_X, where X is ΔH, ZZ, or MM; Data S2). The elemental composition of the biomass, $CH_{1.681}O_{0.418}N_{0.222}S_{0.004}$ (molecular weight of 23.502 g/mol) was taken from Duboc et al.[108] Growth-associated maintenance (GAM) energy is

included in the biomass macromolecule synthesis reactions. Non-growth-associated ATP maintenance costs (NGAM) are represented using the ATP hydrolysis reaction (ATPM).

Cross-contamination check

The bioreactors were checked for cross-contamination by polymerase chain reaction (PCR) using the Thermo Scientific™ Phire Plant Direct PCR Master Mix (Thermo Fischer Scientific, Dreieich, Germany) and custom primers (Data S1 - Table S9). Cells were lysed by boiling 300 μL of the bioreactor culture (with an approximate OD_{600} of 1.00) at 100°C for 10 min (ThermoMixer® C 460-0223, Eppendorf, Hamburg, Germany); 1 μL was directly taken as a template for the PCR reaction. Primers were used at a 10 μM concentration. The PCR was carried out for 28 cycles in a Mastercycler® pro S (Eppendorf, Hamburg, Germany). The results of the reactions were visualized with gel electrophoresis (1% w/v agarose and SYBR™ Safe DNA Gel Stain (Thermo Fischer Scientific, Dreieich, Germany) and Gel Doc™ XR+ visualizer (Bio-Rad, Feldkirchen, Germany).

Fermentation gas analysis

A 490 Micro Gas Chromatograph (microGC; Agilent, Santa Clara, United States) fitted with a multi-valve port system (Teckso GmbH, Neukirchen-Vluyn, Germany) was used to analyze the inlet and outlet gas compositions. The microGC was equipped with two columns, the Molecular Sieve 5A PLOT 0.25 mm, 10 m (Agilent, Santa Clara, United States) that used Argon as a carrier gas to measure H_2, O_2, N_2, CH_4, and CO, and the PoraPLOT Q PLOT, 0.25 mm, 10 m (Agilent, Santa Clara, United States) that used Helium as a carrier gas to measure CO_2, N_2 (combined with O_2), and H_2S. The microGC was calibrated before the run using six and four calibration levels for experiments 1 and 2, respectively (Data S1 - Table S10). Each level was sampled for four replicates, and the average was taken as the calibration point. The total method lasted 180 s with a sample time of 20 s. The injector and sample line temperatures were set at 110°C, and column temperatures and pressures at 60°C and 150 kPa, respectively.

Biomass concentration analysis

The biomass correlation coefficient (K in g/L/OD_{600}), as defined in Valgepea et al.,[109] was found by sampling 60 mL from each bioreactor (6 Falcon™ tubes of 10 mL) at the end of each steady-state period. The samples were centrifuged, the supernatant removed, and two Falcon™ tubes combined with 0.5 mL of Millipore water in pre-weighed glass vials (548-0028; VWR International GmbH, Darmstadt, Germany), resulting in three technical replicates per bioreactor. The vials were dried at 200 mbar (absolute pressure) and 80°C for three days, and the weight of the biomass was recorded. The slope of the measured biomass weight to the corresponding OD_{600} was taken and divided by the volume (0.02 L) to give the following K values (g/L/OD_{600}): First experiment: *M. thermautotrophicus* ΔH, 0.32; *M. thermautotrophicus* Z-245, 0.31; and *M. marburgensis* Marburg, 0.28. Second experiment: *M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$ (H_2 and CO_2), 0.31; *M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$ (sodium formate), 0.89; and *M. thermautotrophicus* Z-245 (sodium formate), 0.95.

Sodium formate concentration measurements

Sodium formate concentrations were analyzed *via* high-pressure liquid chromatography (HPLC) (SIL-40C, Shimadzu Europa, Duisburg, Germany) system that was equipped with an Aminex HPX-87H column (300 by 7.8 mm; Bio-Rad, CA, USA) and a refractive index detector (RID-20A). A 5 mM sulfuric acid solution was used as the eluent, with a flow rate of 0.6 mL min^{-1} and a sample run time of 30-60 min. The oven temperature was set to 60°C, while the sample rack of the attached autosampler to 4°C. For HPLC sample preparation, all culture samples (0.5 mL ± 0.1 mL) were filtered using 0.22 μm filters (ROTILABO® PVDF, 13 mm, Carl Roth, Karlsruhe, Germany). Sodium formate calibration curves were prepared with concentrations ranging from 0.5-10 mM and 20-400 mM.

RNA sample preparation

When the steady state was reached, four technical replicates of 9 mL (n=4) of bioreactor sample were placed into 5 mL of prechilled (overnight at 4°C) RNAlater™ Stabilization Solution (Thermo Fischer Scientific, Dreieich, Germany). The samples were stored overnight at 4°C and then frozen at -20°C until RNA isolation. The samples were thawed on ice and centrifuged at 4°C and 4100 × g for 10 min (Multifuge X3R TX-1000, Fischer Scientific, USA), the supernatant was discarded, and the samples were resuspended in 800 μL of RNase-free water. The 800 μL were mixed with 950 μL of saturated phenol (Sigma-Aldrich, Merck, Germany) and 115 μL of a lysis solution in Lysing Matrix B (MP Biomedicals Germany GmbH, Eschwege, Germany). The lysis solution contained sodium acetate (20 mM pH 5.2), sodium dodecyl sulfate (SDS; 0.5 % v/v), ethylenediaminetetraacetic acid (EDTA; 1 mM), and DNase/RNase-free distilled water (DI; Invitrogen, Thermo Fischer Scientific, Germany). The cells were homogenized for one cycle (5 x 40 s at 6 m/s and 20 s off) in a FastPrep-24™ 5G bead-beater (MP Biomedicals Germany GmbH, Eschwege, Germany) and then centrifuged at room temperature and 21130 × g for 10 min (5424 Eppendorf, Hamburg, Germany). The top layer was placed into 600 μL of phenol-chloroform-isoamyl alcohol (ROTI® Aqua-P/C/I, Carl Roth, Karlsruhe, Germany) and centrifuged as in the previous step. This step was repeated twice. The top layer was added to 1.2 mL of ethanol (undenatured absolute, SERVA Electrophoresis GmbH, Heidelberg, Germany) and left overnight at -80°C to precipitate the RNA. The following day, the tubes were centrifuged at 4°C and 21130 × g for 10 min, and then the supernatant was removed. The RNA pellet was washed with 1 mL of 75% v/v ethanol solution and centrifuged as in the previous step. The supernatant was pipetted off, and the pellet was resuspended with 53 μL of DNase/RNase-free distilled water. The RNA was cleaned (which included DNA depletion) and concentrated using the RNA Clean & Concentrator Kit (Zymo Research, Irvine, CA, United States) according to the manufacturer's instructions. The cleaning and concentrating protocols were repeated

four times. The quantity (6-16 μg) and quality (RNA integrity index > 9) of the purified RNA were measured with a 2100 Agilent Bioanalyzer (Agilent Technologies, Santa Clara, United States) using the Agilent RNA 6000 Nano Kit (Agilent Technologies, Santa Clara, United States) before freezing the samples at -80°C until sequencing.

RNA sequencing

The library preparation was performed using 100 ng of RNA and the TruSeq™ Stranded Total RNA Kit with Ribo-Zero™ Plus (Illumina, San Diego, United States). Pair-ended sequencing was performed using the NovaSeq™ 6000 with the Flow Cell Type 2 x 100 bp (Illumina, San Diego, United States). Demultiplexing of the sequences was performed with Illumina bcl2fastq (2.20) software. The RNA sequencing was performed by CeGaT GmbH (Tübingen, Germany).

Protein sample preparation

The proteomics sampling and sample preparation methods from Valgepea et al.[50] were used with the following modifications. Screw-cap 2-mL tubes with lysing matrix B (MP Biomedicals Germany GmbH, Eschwege, Germany) were used in the homogenizer (FastPrep-24™ 5G, MP Biomedicals Germany GmbH, Eschwege, Germany). For cell lysis, the one cycle of bead beating as described for transcriptomics sample preparation was repeated three times. The following protein precipitation method from the Proteome Center Tübingen (PCT) at the University of Tübingen was applied. The supernatant was transferred to a 15-mL Falcon™ tube with 4 mL and 0.5 mL of ice-cold 100% acetone and 100% methanol, respectively, and precipitated overnight at -20°C. The following day, the tubes were centrifuged at 2200 × g and 4°C for 20 min (Multifuge X3R TX-1000, Thermo Fischer Scientific, Waltham, United States). The supernatant was removed, and the pellet was washed with 1 mL of ice-cold acetone (80/20% v/v in water) and centrifuged as in the previous step. The supernatant was removed, and the protein pellet was left to air dry on ice for 15 min before freezing at -20°C.

Protein measurement

Pellets were resuspended in a denaturation buffer (6 M urea, 2 M thiourea in 10 mM Tris pH 8.0), and 20 μg of protein were subjected to tryptic in-solution digestion (0.2 μg trypsin). The samples were run for an LC-MS/MS analysis on an Easy-nLC™ 1200 (Thermo Fisher Scientific, Dreieich, Germany) coupled to an Orbitrap Exploris™ 480 mass spectrometer (Thermo Fisher Scientific, Dreieich, Germany). The LC-MS analysis was performed as described previously in Fagbadebo et al.,[110] with the exception of the duration and gradient of the peptide elution. Solvent A was 0.1% formic acid, and solvent B was 80% acetonitrile in 0.1% formic acid.[110] The LC-MS was operated at 40°C.[110] The peptides were eluted from 0-113 min with a linear gradient from 10% to 33% of solvent B at a flow rate of 200 nL/min. Then, from 113-116 minutes, the gradient increased from 33% to 50% of solvent B at 200 nL/min. To wash the remaining peptides from the column, from minutes 116-119, the gradient increased from 50% to 90% of solvent B with a gradually increasing flow rate from 200 nL/min to 500 nL/min. Lastly, the column was then washed from 119-127 min with 90% solvent B at 500 nL/min.

Pan-genome differential expression database

With the *M. thermautotrophicus* ΔH genes as a reference, an intersection of homologous genes/proteins based on the best BLASTp+ hits from *M. thermautotrophicus* Z-245 and *M. marburgensis* Marburg was used (Equation 3). *M. thermautotrophicus* ΔH and *M. thermautotrophicus* Z-245 had 81 additional homologous genes/proteins, which were not considered when *M. marburgensis* Marburg was being analyzed, while *M. thermautotrophicus* ΔH and *M. marburgensis* Marburg had 11. This pan-genome of homologous gene groups can be found in Data S4.

$$(ZZ \cup MM) \cap \Delta H \qquad \text{(Equation 3)}$$

A protein FASTA file that represented the homologous clusters was required to perform relative proteomics across different species. Each homologous cluster FASTA entry needs to account for genetic heterogeneity that leads to variation in protein sequences across strains. Therefore, peptide sequences from the other strains had to be added to the gene-group sets.[111] An *in-silico* tryptic digestion of the homologous protein FASTA sequences for *M. thermautotrophicus* ΔH, *M. thermautotrophicus* Z-245, and *M. marburgensis* Marburg was performed using Rapid Peptides Generator version 1.2.4 to retrieve all peptides for each strain.[112] A new enzyme was defined to cleave after lysine (K) or arginine (R) except if proline (P) follows, with the cleaving rule as (K or R,) with the exception (K or R,)(P). Then, using the *M. thermautotrophicus* ΔH protein FASTA sequence as a reference, unique peptides (with a length greater than six amino acids) from homologous proteins of *M. thermautotrophicus* Z-245 and *M. marburgensis* Marburg that were not already in the *M. thermautotrophicus* ΔH protein were inserted *prior* to the C-terminal peptide (Data S7). If *M. thermautotrophicus* Z-245 and *M. marburgensis* Marburg shared a C-terminus different from the one of *M. thermautotrophicus* ΔH, then that C-terminus was additionally amended as the C-terminus for the new combined protein, otherwise the C-terminus from *M. thermautotrophicus* ΔH was kept. This new protein FASTA file was then used for the proteomics analysis (Section 1.21).

Integrating fermentation data in the GEMs

Experimental fermentation data (Conditions 22-24) were used to constrain the models for the flux balance analysis simulations (with biomass maximization as the objective function) to validate the three strain-specific GEMs (Data S5 and S7). The data did not include time points with suspected gross measurement error and was adjusted with the maximum likelihood estimates.

Reduced experimentally-constrained GEMs were created by applying the GIMME[42] algorithm with the transcriptomics and proteomics data (thresholds set at the lower quartile). This step was accomplished in the COBRA Toolbox v.3.1[113] using MATLAB (R2018b) and the Gurobi Optimizer v.9.0.1.[114] Also in the COBRA Toolbox, flux balance analysis simulations (with and without loops allowed) were run on the resulting reduced GEMs, using either the maximization of biomass exchange (EX_biomass_e) or the maximization of ATP dissipation (ATPM) as the objective function (Data S5 and S7).

Visualization

The graphical abstract was created with Biorender.

QUANTIFICATION AND STATISTICAL ANALYSIS

Genome assembly

De novo genome assembly (including the plasmid when applicable) was conducted, and methylation patterns were deciphered from the PacBio sequencing results. The utilized pipeline included seven main aspects: 1) contamination control by DIAMOND alignment against the NCBI-nr database to confirm the absence of any non-*Methanothermobacter* reads; 2) assembly by Canu[115] (which includes error correction, trimming, and assembly); 3) first polishing by BlastR, samtools,[116,117] and Arrow; 4) circularizing by Blast[118] and BioPython; 5) second polishing by BlastR, samtools,[116,117] and Arrow; and 6) methylation patterns prediction by ipdSummary, MotifMaker,[119] and BaseModFunctions. The final annotation was performed by National Center for Biotechnology Information (NCBI) upon genome sequence submission. Details on the applied programs can be found in Data S1 - Table S8.

Genome comparisons

Protein FASTA files were generated using the GenBank files from NCBI and the BioPython version 1.77.[120] Genome-wide comparisons were performed with protein Basic Local Alignment Search Tool (BLASTp+) version 2.10.0 locally.[121–123] The new *M. thermautotrophicus* ΔH and *M. marburgensis* Marburg sequences were compared to the previously published ones, accession numbers NC_000916.1 [62] and NC_014408.1,[64] respectively. The alignments were performed twice, both times with an expectation value (*e*-value) cut-off of 10E-30. The first time, the new sequences were set as the queries and the old sequences as the subjects. The second time the query and subject were reversed. The protein hit with the highest bit score was taken as the matching protein (and gene). For homologous gene comparisons (which were used for the pan-genome and pan-model construction), the *e*-value cut-off was set to 0.001. The following pairwise comparisons were performed with the following query and subject pairs: 1) *M. thermautotrophicus* ΔH and *M. marburgensis* Marburg; 2) *M. thermautotrophicus* ΔH and *M. thermautotrophicus* Z-245; and 3) *M. marburgensis* Marburg and *M. thermautotrophicus* Z-245. The protein hit with the highest bit score was taken as the homologous protein (and gene). The three comparisons were merged with *M. thermautotrophicus* ΔH as the backbone *via* the LOOKUP function in Excel®. Some supplementary gene-wise annotations were completed with the online BLASTp tool.[124]

COG functional annotation

The clusters of orthologous genes (COG) functional annotation was performed on the protein FASTA file for each strain with the script cdd2cog v0.2 from the bac-genomics-scripts according to the author's instructions.[125] A modification to the code, as suggested in issue #14 (https://github.com/aleimba/bac-genomics-scripts/issues/14, accessed on 12/2021), was applied to allow the use of the COG2020 database.[126]

Carbon and electron balance calculations

For the first experiment, initial carbon balances using molar flow rates ($\dot{n}$, mmol/h) were calculated according to conservation of mass (Equation 4) using the ideal gas law (Equation 5) with the following assumptions:

- The ideal gas constant of 8.3144 L·kPa/K/mol
- A temperature (T) of 25°C (298.15 K)
- A pressure (P) of 1 atm (101.325 kPa).

$$C_{in} - C_{out} = 0 \quad \text{(Equation 4)}$$

$$CO_{2,in} - \left(CO_{2,gas\_out} + CO_{2,aq\_out} + CH_{4,gas\_out} + biomass_{aq\_out}\right) = 0$$

$$\dot{n} = \frac{P\dot{V}}{RT} \quad \text{(Equation 5)}$$

The volumetric flow rate ($\dot{V}$, L/h) was found by multiplying the gas concentration (the fraction measured by the microGC) by the measured gas flow rate. This was done for the CO_2 of the inlet gas (CO_{2,gas_in}), and the CO_2 (CO_{2,gas_out}) and CH_4 (CH_{4,gas_out}) of the outlet gas. In addition, Henry's law was used to account for the soluble CO_2 (CO_{2,aq_out}) that was removed from the bioreactors in the effluent. To find the

quantities of soluble bicarbonate and carbonate in the medium, the acid dissociation constants (and pKa) were found using the equations provided by Prieto and Millero[127] (Data S3). The salinity of the medium (4.565 ppt), required for the equations in Prieto and Millero,[127] was estimated by summing the amount of cations and anions of the salts that were added to the medium. The estimated concentrations of solubilized and removed carbon species were subtracted from the calculated gas uptake values. Lastly, to calculate the biomass concentration in the bioreactor ($biomass_{aq\_out}$), the biomass correlation coefficient (K in g/L/OD_{600}) was divided by the measured bioreactor OD_{600} and volume. The resulting g_{CDW} was divided by the molecular weight (23.502 g/mol)[108] and multiplied by the measured dilution rate of the bioreactors (h^{-1}). Given inconsistencies in carbon balances in our measurements, a gross measurement error analysis was performed following the macroscopic balance method described in Wang and Stephanopoulos[128] (Data S7). In brief, first, the probable gross measurement error was identified *via* statistical hypothesis testing with a chi-square test. Second, maximum likelihood estimates were calculated (dropping the data points with gross measurement error) to find corrected values for the gas and biomass data.

For the second experiment, on H_2 and CO_2, the carbon balances were performed as described above with an adjusted salinity value of 4.568 ppt rather than 4.565 ppt. However, given more considerable inconsistencies in carbon and electron balances with the raw data for bioreactor runs on sodium formate, a different method was applied in this case. No sodium formate was detected in the liquid media during steady state. Thus, we assumed that the microbes consumed the entire sodium formate from the feed in the bioreactor ($Na\text{-}formate_{in}$). However, the quantity of 1 M HCl solution added to the bioreactors for pH control was not insignificant (unlike the amount of NaOH added for fermentations on H_2 and CO_2). The feed rate of the pH control pumps was determined, and a dilution factor was calculated. Approximately 21.88 $\pm$ 2.93% of the liquid feed into the bioreactors was due to the acid feed, thus lowering the effective sodium formate feed concentration to 291.38 $\pm$ 10.55 mM (Data S1 - Table S11). Further, the estimation of dissolved carbon species is very sensitive to salinity. The increased salt concentrations from the sodium formate (new value estimated at 17.80 $\pm$ 0.44 ppt) were insufficient for closing both carbon and electron balances. It is also noted that variation in pH was more frequent due to the production of NaOH with sodium formate consumption, and at times, there was a noticeable shift in the online pH probe measurement that required a manual offset after offline pH measurements. Furthermore, the dissolved CO_2 values are sensitive to changes in pH. Thus, to solve for the quantities of produced CH_4 and CO_2, Equations 6 and 7, which represent the carbon and electron (using degree of reduction) balances, were solved. The biomass coefficient in Equation 7 was calculated assuming a molecular formula of $CH_{1.681}O_{0.481}N_{0.222}$, as described in Michael and Kargi.[129] H_2 gas was detected in the outlet gas stream (H_{2,gas_out}), and thus it was also included in Equation 7.

$$C_{in} - C_{out} = 0 \quad \text{(Equation 6)}$$

$$Na - formate_{in} - \left(CO_{2,gas\_out} + CO_{2,aq\_out} + CH_{4,gas\_out} + biomass_{aq\_out}\right) = 0$$

$$Na - formate_{in} - \left(CO_{2,gas+aq\_out} + CH_{4,gas\_out} + biomass_{aq\_out}\right) = 0$$

$$E_{in} - E_{out} = 0 \quad \text{(Equation 7)}$$

$$2 * Na - formate_{in} - \left(8 * CH_{4,gas\_out} + 4.053 * biomass_{aq\_out} + 2 * H_{2,gas\_out}\right) = 0$$

Normalized product distribution

Product distributions were calculated using CH_4 and biomass data (mmol/h) for bioreactors and time points that did not have suspected gross measurement errors (STAR Methods – carbon and electron balance calculations). The quantity of CH_4 or biomass was divided by the sum of the CH_4 and biomass and then multiplied by 100, resulting in the normalized product distribution percentage. For fermentations on sodium formate, the CO_2 production was also accounted for in the products. An analysis of variance (ANOVA) test was conducted in Excel® using the ANOVA:Single Factor (α = 0.05) function of the Data Analysis Add-in to analyze statistically significant differences between the product distribution ratios of the three different strains (first experiment). Further, product distribution ratios were compared for: 1) *M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$ and *M. thermautotrophicus* Z-245 grown on sodium formate (second experiment); and 2) product distribution ratios of *M. thermautotrophicus* ΔH, *M. thermautotrophicus* Z-245, and *M. marburgensis* Marburg grown on H_2/CO_2 (from the first experiment) with *M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$ grown on H_2/CO_2 (from the second experiment).

Interspecies comparison

For the first experiment, all gas and biomass data points without suspected gross measurement error were used for the comparative analysis (independent *t*-tests) between the three strains, except for one bioreactor replicate for *M. thermautotrophicus* Z-245 that experienced a pump malfunction and wash-out of bioreactor broth during the fermentation.

For the second bioreactor experiment, one replicate of *M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$ grown on sodium formate was discarded because of a mixer malfunction, and thus, the gas and biomass data points were not included in the *t*-tests. The rest of the gas and biomass data points were used for the comparative analysis between *M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$ and *M. thermautotrophicus* Z-245 grown on sodium formate. The second set of *t*-tests was conducted between *M. thermautotrophicus* ΔH,

M. thermautotrophicus Z-245, and *M. marburgensis* Marburg grown on H_2/CO_2 from the first experiment with *M. thermautotrophicus* ΔH pMVS1111A:P_{hmtB}-$fdh_{Z\text{-}245}$ grown on H_2/CO_2 from the second experiment.

The *t*-tests (two-tailed and confidence intervals of 95%) were carried out in Python 3.6.13 using the Researchpy package Version 0.3.2 and the ttest function.[130] Three-way comparisons were performed in Excel® using the ANOVA:Single Factor (α = 0.05) function of the Data Analysis Add-in. The maximum likelihood estimates of the measurements for time points without gross measurement error were calculated,[128] which resulted in closed carbon balances for those time points (assuming 90% confidence intervals for the test function). These adjusted values were then averaged across time points for each strain and used to constrain the GEMs (Data S3).

Raw RNA sequencing data analysis

Sequencing reads were quality filtered and trimmed with BBMap version 38.93[131] and FastX Toolkit version 0.0.14[132] with specific parameters for paired-end reads. Reads from each bioreactor culture were mapped to the reference genomes of *M. thermautotrophicus* ΔH (CP064324), *M. thermautotrophicus* Z-245 (CP064336 and CP064337), and M. *marburgensis* Marburg (CP069376 and CP069377) using BWA-MEM version 0.7.17.[133] Mapped reads were assigned to genomic features using the FeatureCounts program from the Subread package version 2.0.1.[134] Additionally, an alignment-free quantification of gene features was performed using Salmon version 1.5.2.[135] Pairwise differential expression analysis was performed on the Salmon counts using DESeq2 version 1.32.0[136] with the homologous genes present in the pan-genome database (STAR Methods - pan-genome differential expression database) and ConsensusDE version 1.10.0.[137] The adjusted *P*-value cut-off threshold was set to 0.05. This process was orchestrated using Snakemake version 6.8.0[138] to achieve reproducibility (Data S7). The output files from Salmon were used to analyze each strain individually, and only genes with at least two non-zero transcript values were considered. The averages of each gene were then used.

Protein analysis with proteome Discoverer

Raw data-dependent acquisition (DDA) data from LC-MS/MS analysis was performed using Thermo Fisher Proteome Discoverer software (version 2.5.0.400, Thermo Fisher Scientific, Scoresby, Australia). Files were processed using two different methods: 1) for intensity-based absolute quantification (iBAQ) analysis, each strain was processed individually through the precursor label-free quantification workflow with their corresponding FASTA files and then converted to iBAQ as described in Schwanhäusser et al.[139]; and 2) for differential expression analysis, all files were processed together using the custom pan-genome FASTA file (STAR Methods - pan-genome differential expression database). Both methods used the same workflows and only differed in their choice of FASTA file. Briefly, spectra were searched against the respective databases with carbamidomethylation set as a fixed modification and up to three methionine oxidations. Acetylation, methionine loss, and acetylation after methionine loss at the N-terminus were also permitted. The mass tolerance of precursors and fragments were 10 ppm and 0.02 Da, respectively. The minimum and maximum peptide lengths were six and 144 amino acids, respectively. Two missed cleavages were allowed per peptide. Proteins were filtered to a 1% FDR cut-off threshold and for an adjusted *P*-Value of 0.05. For quantification, only peptides unique to a protein group were used. All other parameters were kept at their default settings.

Eha/Ehb ratio determination

The ratio of the energy-converting hydrogenases Eha and Ehb subunits was found by summing the abundances (either transcripts or relative protein abundances) of the subunits detected for each energy-converting hydrogenase and dividing those sums by the number of subunits found. This was done to normalize against subunits that were not found (the stoichiometry between all subunits was considered as equimolar based on current knowledge from the literature).[140] The normalized abundance of Eha was then divided by that of Ehb. This ratio was then averaged across the replicates (bioreactor samples) of each strain. An ANOVA test was conducted in Excel® using the ANOVA:Single Factor (α = 0.05) function of the Data Analysis Addin to test if there was a statistically significant difference between the Eha/Ehb ratios for the three different strains (Data S4).

Methanogenesis relative abundances

Gap-filled relative abundances were found for the different proteins (or complexes) involved in the Wolfe cycle of methanogenesis. The iBAQ values of each subunit were summed and divided by the number of detected subunits (the maximum number of subunits missing per strain per protein complex was one). This sum was then multiplied by the theoretical number of subunits that should be found (effectively assigning the non-detected subunit the average iBAQ value for the complex) to produce the entire protein complex normalized iBAQ value. This value was then divided by the sum of all iBAQ values in the proteome and multiplied by 100, resulting in relative (normalized) protein abundance percentages (Data S4). The stoichiometry between all subunits was considered as equimolar based on current knowledge from the literature.[140]