Nature AND Nurture: Enabling formate-dependent growth in *Methanosarcina acetivorans*

Jichen Bao[1,#], Tejas Somvanshi[1], Yufang Tian[1], Maxime Laird[1], Pierre Simon Garcia[2], Christian Schöne[3], Michael Rother[3], Guillaume Borrel[2], Silvan Scheller[1,#]

1. Department of Bioproducts and Biosystems, School of Chemical Engineering, Aalto University, FI-02150, Espoo, Finland
2. Institut Pasteur, Université Paris Cité, UMR CNRS6047, Evolutionary Biology of the Microbial Cell, Paris, France
3. Institute of Microbiology, Technische Universität Dresden, 01062 Dresden, Germany

#Corresponding author: Silvan Scheller (silvan.scheller@aalto.fi); Jichen Bao (jichen.bao@aalto.fi)

Abstract

Methanogens are essential players in the global carbon cycle. *Methanosarcinales* possess one of the most diverse metabolic repertoires for methanogenesis, but they have never been observed to utilize formate as a substrate.

We successfully reprogrammed the primary metabolism of *Methanosarcina acetivorans* by introducing an exogenous formate dehydrogenase derived from a closely related species. The engineered strains acquired the capacity to harness energy from formate-dependent methanogenesis pathways, including formate-dependent methyl reduction and formate-dependent carbon dioxide reduction.

The ability of *M. acetivorans* to thrive on formate suggests the existence of essential accessory machinery and metabolic redundancy for generating reduced ferredoxins from $F_{420}H_2$. This remarkable plasticity in energy metabolism raises the possibility that an ancestral lineage of *Methanosarcinales* may have possessed the capacity to utilize formate.

By combining this genetically modified strain with a disruption in methyl disproportionation, we have created a novel tool for investigating and manipulating the components of the F_{420} reduction and methanogenesis pathways independently.

Introduction

Anaerobic methanogenesis stands as one of the most ancient biochemical pathways, continuing to play a crucial role in the global carbon cycle. Its intricate network of metabolic interactions intertwines a diverse array of metabolites. Among the three primary methanogenesis pathways – hydrogenotrophic, methylotrophic, and aceticlastic – all three are present in *Methanosarcina*, with additional modifications for compounds like dimethyl sulfide, methanethiol, and carbon monoxide[1]. Remarkably, despite this versatility, *Methanosarcina* lacks the capacity to utilize formate for methanogenesis[2]. This absence is particularly intriguing considering that certain *Methanosarcina* species encode formate dehydrogenase, and strains devoid of a complete formate dehydrogenase still possess accessory proteins[3]. FdhAB has been identified in the genome of *M. barkeri*, but its expression or activity has not been demonstrated. In contrast, *M. acetivorans* encodes the *fdhD* gene. Formate-dependent carbon dioxide-reducing methanogenesis is enabled by *fdhAB*; however, it's essential to note that mere *fdhAB* presence

does not guarantee formate utilization. At least in *Thermococci* and *Methanococci*, the presence of a formate transporter is crucial for formate utilization[4].

Beyond the enzymatic oxidation of formate, another critical step involves the efficient transfer of electrons derived from formate to appropriate pathways. These electrons serve as essential energy sources for both carbon dioxide reduction and anabolic processes. In methanogens that can utilize formate as a substrate, the electron shuttle protein energy conserving hydrogenase (Eha/Ehb) converts formate-derived electrons into reduced ferredoxins[5,6]. However, *M. acetivorans* lacks this enzyme. To acquire the reduced ferredoxins required for its methylotrophic or aceticlastic growth, *M. acetivorans* employs an alternative strategy: it partially oxidizes a portion of the carbon it utilizes, generating CO_2 and releasing electrons that can be shuttled to ferredoxins. This process is primarily facilitated by the methyl-tetrahydromethanopterin: coenzyme M methyltransferase (mtr) enzyme in methylotrophic methanogenesis. The combination of substrates such as methanol and acetate or methanol and carbon monoxide has been shown to disrupt this enzyme without completely halting methanogenesis. In both cases, reduced ferredoxins can still be generated by oxidizing carbon monoxide or acetate to CO_2[7].

Given the methylotrophic origin of methanogenesis[8], the presence of formate dehydrogenase in the Last Universal Common Ancestor (LUCA)[9], and the abundance of formate on early Earth[10], it is plausible that the methylotrophic origin of methanogenesis was facilitated not only by hydrogen but also by formate.

Results

I) Impact of formate on transcription profile of *M. acetivorans* WWM73

To evaluate the potential impact of formate on wild-type cells that lack a formate dehydrogenase enzyme, *M. acetivorans* was cultivated with and without added formate in HS medium supplemented with methanol as the carbon source. *M. acetivorans* showed an inability to utilize formate present in the medium. This was confirmed by measuring formate levels throughout the growth period. Notably, the final methane yield from methanol remained unchanged regardless of formate presence or absence. However, a significant difference emerged in growth rate. *M. acetivorans* exhibited slower growth in the presence of formate compared to its growth without formate. The doubling time with formate was ~8 hours, compared to ~6 hours without formate (Fig. 1).

RNA-seq analysis of *M. acetivorans* WWM73 grown with and without formate revealed 836 differentially expressed genes (DEGs) that exhibited significant upregulation or downregulation. Most of these genes were annotated as predicted hypothetical proteins, suggesting their potential involvement in yet-uncharacterized metabolic processes. Several genes were classified as transcriptional regulators or contained DNA-binding domains, suggesting their role in regulating gene expression. Methyltransferases responsible for methyl transfer from methylated amines were also significantly downregulated in the presence of formate. The *nif* operon, responsible for nitrogen fixation, was downregulated in the presence of formate. Although formate's inhibitory effect on nitrogen fixation was evident, it did not completely abolish growth in medium devoid of ammonia up to a concentration of 60 mM.

II) In vitro characterization of fdhAB and cell suspension test

To determine whether the gene encodes a functional protein, F_{420}-reducing formate dehydrogenase (*fdhAB*) from *M. barkeri* was expressed from a plasmid under the P_{mcr}_*tet01* promoter in *M. acetivorans* WWM73. The activity of the recombinant enzyme was assessed by measuring its ability to reduce benzyl viologen in cell-free extracts. The enzyme exhibited an activity of 720 mU mg^{-1} of total protein.

Subsequent cell suspension experiments demonstrated that the fdhAB-encoded formate dehydrogenase mediated the conversion of formate to carbon dioxide and methane, but not acetate. The combination of CO and formate resulted in a significantly higher methane yield compared to either substrate alone.

III) Formate dependent methylotrophic methanogenesis

To enable formate-dependent methylotrophic methanogenesis, the *fdhAB* gene was integrated into the genome of *M. acetivorans* WWM73, resulting in the disruption of the methyl-THMPT:Coenzyme M methyltransferase (*mtr*) gene. This genetic modification rendered the strain incapable of utilizing methyl group disproportionation to obtain the reducing equivalents necessary for methanol reduction. Therefore, the strain relied on the oxidation of formate to generate $F_{420}H_2$. Consequently, the strain lost the ability to grow solely on methanol or any other single methanogenic substrate. However, it could still grow efficiently on a combination of acetate and methanol.

M. acetivorans was grown on a medium supplemented with 60 mM methanol and formate. As acetate formation was minimal in cell suspension studies, 5 mM acetate and trace pyruvate were added to the medium to provide cellular carbon. This media composition was designated as MFAcP media. The doubling time in MFAcP media was ~11 hours. Metabolite analysis revealed that all formate and methanol were completely consumed by the end of growth. When provided with 5 mM acetate and trace pyruvate for cellular carbon, the biomass was significantly higher compared to wild-type cells grown in methanol-only media (Fig 2).

To corroborate the cell suspension studies indicating that cellular carbon cannot be synthesized from methanol and formate alone, *M. acetivorans* JBA01 was grown in culture media without added acetate or pyruvate. Surprisingly, the strain was able to grow, albeit with lower OD and methane yield. This finding was unexpected given that *M. acetivorans* lacks energy-converting hydrogenase (Ech), a key enzyme involved in producing reduced ferredoxins (Fd_{red}). Metabolite analysis revealed that not all methanol was reduced, whereas all formate was consumed. This suggests that a portion of the formate was utilized for cellular carbon synthesis, thereby limiting the availability of formate for methanol reduction.

The ability of *M. acetivorans* JBA01 to grow on a medium containing only methanol and formate remains an intriguing aspect of its metabolism. Further studies are warranted to elucidate the mechanisms by which this strain utilizes formate for cellular carbon synthesis in the absence of Ech.

IV) Whole transcriptome analysis of *M. acetivorans* JBA01

The whole transcriptome was analyzed to understand how the strain was able to synthesize Fd_{red} and in turn cellular carbon when grown in methanol and formate only as *mtr* was disrupted. Without *mtr* carbon from methanol cannot be oxidized in *M. acetivorans* as no bypass

exists[11]. The transcriptome revealed some interesting changes with or without added acetate and pyruvate. A total of 1814 DEGs were significantly upregulated or downregulated. Firstly, the *fdhAB*, inserted in the *mtr* operon was upregulated in media without acetate and pyruvate, which might be an artifact since *fdhAB* was inserted in the same location as *mtr*. Along with, *rnf* was significantly upregulated. A lot of the subunits of the multi-subunit *fpo* complex were also significantly upregulated although crucially *fpoF* was not one of them. The hydrogenases that were shown to be silent[12] in *M. acetivorans* were also slightly upregulated in absence of acetate and pyruvate. Other interesting significantly upregulated DEGs were *cdh, mcr, mta, fwd, coo*. As was the case previously, a lot of the DEGs were annotated as predicted/conserved hypothetical proteins with some of them being transcriptional regulators.

V) Formate dependent CO_2 reduction

After successfully attempting formate-dependent methylotrophic methanogenesis, we wanted to attempt formate dependent CO_2 reducing methanogenesis. This methanogenesis pathway has never been shown in *Methanosarcinales*. To attempt this, we integrated *fdhAB* without disrupting *mtr* as it would be essential for this methanogenesis pathway. The obtained strain was designated as *M. acetivorans* JBAF01. We used adaptive lab evolution to increase the fitness of JBAF01 to be able to grow and perform methanogenesis solely on formate. The strain with good growth fitness on formate was designated as JBAF02. Repeated serial transfers decreased the doubling time of methanogenesis from couple of days to around 12 hours. The transfers were performed over around one year. As expected, midway through the transfers the doubling time dropped off a cliff. The final doubling time of biomass of JBAF02 obtained was around 12 hours (Fig 3). The genome of select transfer strains was sequenced and analyzed to identify if the sharp drop-in doubling time could correspond to any change in the genome.

Although the strain exhibits similar metabolism to non-cytochrome containing methanogens, cytochromes are still integral to the growth of *M. acetivorans*. This was confirmed by addition of Diphenyleneiodonium chloride to the media, which inhibited the growth of *M. acetivorans* JBAF02 but did not do the same for *M. maripaludis*.[13]

Finally, we have a strain of *M. acetivorans* that can perform formate dependent CO_2 reducing methanogenesis, a novel metabolism in the genus.

VI) Phylogeny of the formate dehydrogenase

We were interested in understanding the origin of this formate dehydrogenase in *Methanosarcinales*. It was interesting to see that the gene was interspersed in the group but could not provide the capability of formate utilization.

We searched for the homologues of FdhA in a database built in a previous study[14] and encompassing bacterial and archaeal proteomes. After alignment of the FdhA sequences, the phylogeny showed that the sequences of methanogens class I and class II cluster together (Figure X). The sequences of FdhA form monophyletic groups corresponding to orders (*Methanobacteriales*, *Methanococcales*, *Methanocellales*, *Methanoflorentales*, *Methanosarcinales*), except for *Methanomicrobiales* that are in two groups. The relative placement of the different clades suggests that methanogens class II acquired FdhA by horizontal gene transfer (HGT) from methanogens class I. Still, it remains unclear if the HGT occurred at the ancestor of methanogens class II or at the ancestor of *Methanomicrobiales*, *Methanoflorentales*, *Methanocellales*, and *Halobacteria*. The former case would imply a loss at the ancestor of *Methanosarcinales*. Finally, considering the very low number of sequences present in *Methanosarcinales* and their grouping with *Methanomicrobiales*, it is likely that *Methanosaeta* and *Methanosarcina* acquired FdhA by a second HGT from Methanomicrobiales.

Discussion

I) *M. acetivorans* WWM73 and formate

i. Formate as transcriptional effector

M. acetivorans has 49 Histidine kinases identified and about 11 response regulators[15]. The transcription levels for a lot of these regulators were significantly affected by the presence of formate in the media. Also, other genes with DNA binding domains, possibly belonging to the family of transcriptional regulators were affected by formate. This implies that formate brings about changes in the transcriptome of *M. acetivorans* and as such the presence of formate is sensed by and responded to. This would make sense as *M. acetivorans* has been shown to produce formate when growing on carbon monoxide and would probably have mechanisms necessary to sense the shift in the environment. This would also be beneficial as formate would be competed for by other anaerobes in the vicinity where it grows. Alternatively, it is also possible that these responses are remnants of past metabolism where *Methanosarcina* would have been able to use formate for growth and methanogenesis.

II) Functional *fdhAB*

i. Environmental cues to trigger *fdh* expression

The *fdhAB* encoded by *M. barkeri* was shown to be functional, at least when expressed using different promoter in *M. acetivorans*. A possible reason why *M. barkeri* was not shown to grow using formate could be because of the regulatory elements or environmental cues not identified in *M. barkeri*. Recent studies have shown FPKM values for *fdhA* and *fdhB* around 40 and 80 respectively in control samples[16].

ii. Formate transporter

Apart from *fdhD*, there are no other formate dehydrogenase associated proteins annotated in the genome of *M. acetivorans* (not accounting for the formylmethanofuran dehydrogenases). There are no formate transporters annotated in the genome. Recent studies have shown the importance of formate transporters for growth on formate. In *Methanococci* and *Thermococci*, growth on formate was not dependent on the formate dehydrogenase genes but rather on the presence of a formate transporter *fdhC*[4]. As there is no *fdhC* found in the genome of *M. acetivorans*, it either uses promiscuity in some other transporter for formate intake or has a yet-unidentified formate transporter in the genome.

iii. Annotated fdhD gene

The genome of *M. acetivorans* encodes a *fdh*D gene which is a formate dehydrogenase family accessory protein catalyzing the sulfur transfer from *IscS* to formate dehydrogenase. In E. coli, *fdh*D has been shown to be essential for formate dehydrogenase activity[17]. The fact that this gene was already present and expressed in *M. acetivorans* could have been what made *M. acetivorans* ammenable for growth and methanogenesis from formate.

III) *M. acetivorans* JBA01

i. Mtr disruption applies pressure to obtain reducing equivalents

In *M. acetivorans* JBA01, *mtr* disruption applied pressure on *fdh*AB for formate oxidation to obtain reducing power. Apart from *fdh*AB, there appears to be a redundancy in energy metabolism natively present to obtain Fd_{red}. Given this redundancy, *M. acetivorans* JBA01 can be

used to check oxidation capability of other compounds or for directed evolution of F_{420} reducing enzymes by applying selection pressure.

IV) Whole transcriptome of *M. acetivorans* JBA01

i. Ferredoxins for anabolism

M. acetivorans on methylotrophic methanogenesis, synthesizes Fd_{red} by oxidizing one molecule of methanol to CO_2 for three molecules of methanol reduced. These Fd_{red} either reduce methanophenazine or are used for acetyl-CoA and pyruvate synthesis. In *M. acetivorans* JBA01, because of disruption of *mtr*, the cell needs to generate Fd_{red} from $F_{420}H_2$ only. There are three possible pathways for this. One could be HdrA2B2C2 complex, that has been shown in vitro to catalyze the endergonic reduction of fd by coupling oxidation of $F420H_2$ to exergonic reduction of heterodisulfide[18]. *hdrA2* has been shown to be not essential in *M. acetivorans*. Hence, with *mtr* disruption, the cell would be incentivized to use HdrA2B2C2 that it would not normally use. The other could be the Rnf complex running in reverse to catalyze formation of Fd_{red} from reduced methanophenazine. Rnf complex running in reverse has been observed in other bacterial [19]. The final possibility is the *fpo* complex. It has been shown in *M. mazei*[20] that *fpoF,* present separately in cytoplasm, is able to catalyze reduction of F_{420} using Fd_{red}. It could also be all 3 in tandem working together as increased expression is seen for all 3. Surprisingly, of all the subunits of *fpo*, *fpoF* was the only not significantly upregulated.

V) Phylogeny of *fdhAB*

Formate has been shown to be abiotically formed near hydrothermal vents [4,10] and was present in early earth. Methanogenesis has been hypothesized to have originated methylotrophically by using H_2 as the reductant prior to evolution of mtr that enabled disproportionation of methyl groups. Given the presence of formate on early earth, formate dehydrogenase or at least its ancestor being present in LUCA and LUCA having the F420-cofactor biosynthesis pathway, it is possible that formate-dependent methylotrophic methanogenesis could have co-evolved with hydrogen-dependent methylotrophic methanogenesis.

Even if not in the origin of methanogenesis but somewhere along the line there would have possibly existed a methanogen ancestor to *M. acetivorans* capable of reducing methylated substrates with formate. The question remains as to why this ability was lost? Was it competition for formate or lack of fitness due to formate?

VI) Case study for ALE

Efforts to make *M. barkeri* grow on formate have failed in the past. As mentioned above, proper environmental cues and/or regulatory parts of the pathway might need optimization to allow *M. barkeri* to grow on formate as well. The ALE method that we used to get *M. acetivorans* to grow solely on formate provides a roadmap for future to obtain a strain of *M. barkeri* like that of *M. acetivorans* that would be able to solely grow on formate.

Methods

Strains and media

Lysogenic broth containing 50 mg L^{-1} ampicillin was used for plasmid construction in *Eshcherichia coli* NEB5α. *M. acetivorans* was cultured in a high-salt medium tailored to the specific requirements of the experiment[21]. Depending on the experimental conditions,

substrates were added as follows: 60 mM methanol alone; 60 mM methanol combined with 60 mM formate; a mixture of 60 mM methanol, 60 mM formate, along with 5 mM each of acetate and pyruvate; or 120 mM formate alone.

Molecular methods

Gibson assembly and golden gate assembly were used to construct the plasmids[22,23]. Liposome-mediated methods and polyethylene glycol-mediated methods were used to transform DNA into M. acetivorans[24,25].

Nucleic acid isolation, sequencing, and analysis

The genomic DNA was extracted with Wizard® Genomic DNA Purification Kit (Promega). The Microbial WGS was performed using Illumina NovaSeq 6000 by Novogene (UK) Company. The resultant data was cleaned using Trimmomatic[26], aligned using BWA-mem2[27] to *M. acetivorans* C2A genome and analyzed for SNPs and SVs using Samtools[28] and BreakDancer[29]. For transcriptomics, the cultures were grown to an OD of 0.3-0.4 and pelleted in RNAprotect. RNA was isolated using Qiagen RNA isolation kit. The total RNA was sent to Novogene where library prep and sequencing was performed. The resultant data was cleaned using Trimmomatic and analyzed using Geneious Prime 2023.1.2 Workbench. The differential expression was calculated using DESEQ2[30] inside Geneious Prime 2023.1.2.

Phylogenetic analysis

Detailed Phylogenetic analysis is available in supplementary materials

Figures

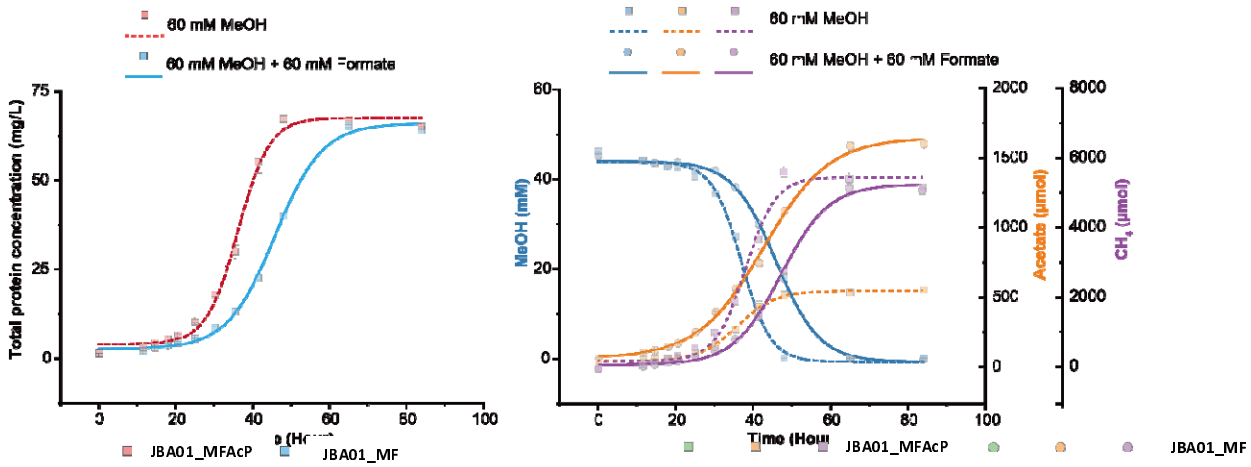

Figure 1 Grow[illegible]es of M. acetivorans WWM73 growth in prese[illegible]e of formate

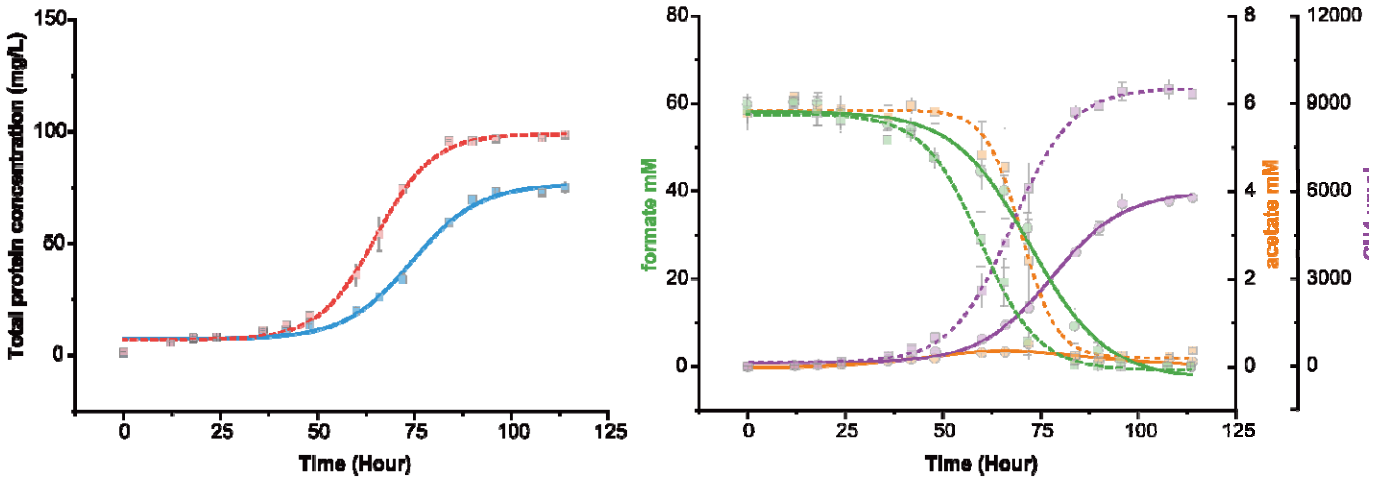

Figure 2 Growth curves and metabolite profiles of M. acetivorans JBA01 growth in presence and absence of acetate and pyruvate

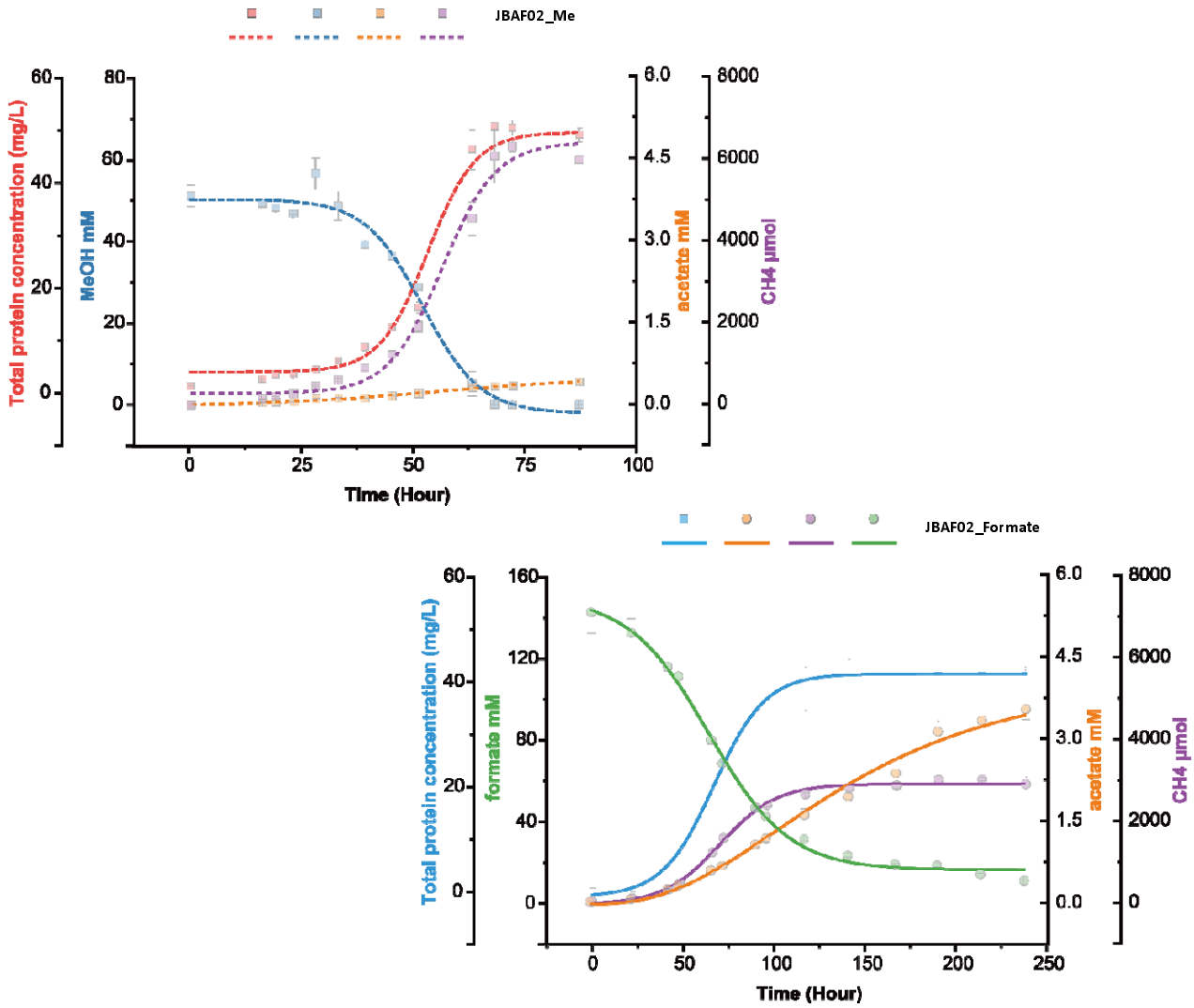

Figure 3 Growth curves and metabolite profiles of M. acetivorans JBAF02 growth ion formate vs methanol

References

1. Kurth, J. M., Op den Camp, H. J. M. & Welte, C. U. Several ways one goal—methanogenesis from unconventional substrates. *Appl Microbiol Biotechnol* **104**, 6839–6854 (2020).

2. Mand, T. D. & Metcalf, W. W. Energy Conservation and Hydrogenase Function in Methanogenic Archaea, in Particular the Genus Methanosarcina . *Microbiology and Molecular Biology Reviews* **83**, (2019).

3. Maeder, D. L. *et al.* The Methanosarcina barkeri genome: Comparative analysis with Methanosarcina acetivorans and Methanosarcina mazei reveals extensive rearrangement within methanosarcinal genomes. *J Bacteriol* **188**, 7922–7931 (2006).

4. Holden, J. F. & Sistu, H. Formate and hydrogen in hydrothermal vents and their use by extremely thermophilic methanogens and heterotrophs. *Front Microbiol* **14**, 1–17 (2023).

5. Welte, C. & Deppenmeier, U. Bioenergetics and anaerobic respiratory chains of aceticlastic methanogens. *Biochim Biophys Acta Bioenerg* **1837**, 1130–1147 (2014).

6. Thauer, R. K., Kaster, A. K., Seedorf, H., Buckel, W. & Hedderich, R. Methanogenic archaea: Ecologically relevant differences in energy conservation. *Nat Rev Microbiol* **6**, 579–591 (2008).

7. Schone, C. *et al.* Deconstructing Methanosarcina acetivorans into an acetogenic archaeon. *Proc Natl Acad Sci U S A* **119**, 1–7 (2022).

8. Wang, Y. *et al.* A methylotrophic origin of methanogenesis and early divergence of anaerobic multicarbon alkane metabolism. *Sci Adv* **7**, 1–12 (2021).

9. Ouzounis, C. A., Kunin, V., Darzentas, N. & Goldovsky, L. A minimal estimate for the gene content of the last universal common ancestor - Exobiology from a terrestrial perspective. *Res Microbiol* **157**, 57–68 (2006).

10. Mohammadi, E. *et al.* Formic Acid, a Ubiquitous but Overlooked Component of the Early Earth Atmosphere. *Chemistry - A European Journal* **26**, 12075–12080 (2020).

11. Schöne, C., Poehlein, A. & Rother, M. Genetic and Physiological Probing of Cytoplasmic Bypasses for the Energy-Converting Methyltransferase Mtr in Methanosarcina acetivorans. *Appl Environ Microbiol* (2023) doi:10.1128/aem.02161-22.

12. Guss, A. M., Kulkarni, G. & Metcalf, W. W. Differences in hydrogenase Gene expression between methanosarcina acetivorans and methanosarcina barkeri. *J Bacteriol* **191**, 2826–2833 (2009).

13. Steiniger, F., Sorokin, D. Y. & Deppenmeier, U. Process of energy conservation in the extremely haloalkaliphilic methyl-reducing methanogen Methanonatronarchaeum thermophilum. *FEBS Journal* **289**, 549–563 (2022).

14. Garcia, P. S., Gribaldo, S. & Borrel, G. Diversity and Evolution of Methane-Related Pathways in Archaea. *Annu Rev Microbiol* **76**, 727–755 (2022).

15. Galagan, J. E. *et al.* The genome of M. acetivorans reveals extensive metabolic and physiological diversity. *Genome Res* **12**, 532–542 (2002).

16. Harris, R. L. *et al.* Transcriptional response to prolonged perchlorate exposure in the methanogen Methanosarcina barkeri and implications for Martian habitability. *Sci Rep* **11**, 1–16 (2021).

17. Thomé, R. *et al.* A sulfurtransferase is essential for activity of formate dehydrogenases in Escherichia coli. *Journal of Biological Chemistry* **287**, 4671–4678 (2012).

18. Yan, Z., Wang, M. & Ferry, J. G. A Ferredoxin- and F420H2-Dependent, Electron-Bifurcating, Heterodisulfide Reductase with Homologs in the Domains Bacteria and Archaea. *mBio* (2017).

19. Buckel, W. & Thauer, R. K. Flavin-based electron bifurcation, ferredoxin, flavodoxin, and anaerobic respiration with protons (Ech) or NAD+ (Rnf) as electron acceptors: A historical review. *Front Microbiol* **9**, (2018).

20. Welte, C. & Deppenmeier, U. Re-evaluation of the function of the F420 dehydrogenase in electron transport of Methanosarcina mazei. *FEBS Journal* **278**, 1277–1287 (2011).

21. Sowers, K. R., Boone, J. E. & Gunsalus, R. P. *Disaggregation of Methanosarcina spp. and Growth as Single Cells at Elevated Osmolarity. APPLIED AND ENVIRONMENTAL MICROBIOLOGY* vol. 59 (1993).

22. Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS One* **3**, (2008).

23. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**, 343–345 (2009).

24. Oelgeschläger, E. & Rother, M. In vivo role of three fused corrinoid/methyl transfer proteins in Methanosarcina acetivorans. *Mol Microbiol* **72**, 1260–1272 (2009).

25. Metcalf, W. W., Zhang, J. K., Apolinario, E., Sowers, K. R. & Wolfe, R. S. *A genetic system for Archaea of the genus Methanosarcina: Liposome-mediated transformation and construction of shuttle vectors*. vol. 94 www.pnas.org. (1997).

26. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, (2014).

27. Md, V., Misra, S., Li, H. & Aluru, S. Efficient architecture-aware acceleration of BWA-MEM for multicore systems. in *Proceedings - 2019 IEEE 33rd International Parallel and Distributed Processing Symposium, IPDPS 2019* (2019). doi:10.1109/IPDPS.2019.00041.

28. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, (2009).

29. Chen, K. *et al.* BreakDancer: An algorithm for high-resolution mapping of genomic structural variation. *Nat Methods* **6**, (2009).

30. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, (2014).