Characterisation of acetogen formatotrophic potential using *E. limosum*

Jamin C. Wood,[1] R. Axayacatl Gonzalez-Garcia,[2] Dara Daygon,[2,3] Gert Talbo,[2,3] Manuel R. Plan,[2,3] Esteban Marcellin,[2,3] and Bernardino Virdis[1,*]

[1] Australian Centre for Water and Environmental Biotechnology (ACWEB, *formerly AWMC*), The University of Queensland, Brisbane, QLD 4072, Australia.

[2] Australian Institute for Bioengineering and Nanotechnology, The University of Queensland, Brisbane, QLD 4072, Australia.

[3] Metabolomics Australia (Queensland node), The University of Queensland, Brisbane, QLD 4072, Australia.

* Correspondence to:

Dr Bernardino Virdis

Email: b.virdis@uq.edu.au

Key words

acetogen; formate; eubacterium limosum; omics; proteomics; metabolomics; chemostat; pyruvate

Abstract

Formate is a promising energy carrier that could be used to transport renewable electricity. Some acetogenic bacteria, such as *Eubacterium limosum*, have the native ability to utilise formate as a sole substrate for growth, which has sparked interest in the biotechnology industry. However, formatotrophic metabolism in acetogens is poorly understood, and a systems-level characterization in continuous cultures is yet to be reported. Here we present the first steady-state dataset for *E. limosum* formatotrophic growth. At a defined dilution rate of 0.4 d^{-1}, there was a high specific uptake rate of formate (280±56 mmol/gDCW/d), however, most carbon went to CO_2 (150±11 mmol/gDCW/d). Compared to methylotrophic growth, protein differential expression data and intracellular metabolomics revealed several key features of formate metabolism. Upregulation of pta appears to be a futile attempt of cells to produce acetate as the major product. Instead, a cellular energy limitation resulted in the accumulation of intracellular pyruvate and upregulation of Pfl to convert formate to pyruvate. Therefore, metabolism is controlled, at least partially, at the protein expression level, an unusual feature for an acetogen. We anticipate that formate could be an important one-carbon substrate for acetogens to produce chemicals rich in pyruvate, a metabolite generally in low abundance during syngas growth.

Introduction

Acetogens hold great promise for sustainable chemical and fuel production whilst closing the carbon cycle using feedstocks from renewable sources (Ljungdahl 2009). Acetogens can use reducing equivalents through intermediates such as hydrogen (H_2), carbon monoxide (CO), formate and methanol. However, compared to synthesis gas fermentation (syngas, a mixture of CO and H_2), as

has been commercialised by LanzaTech using offgases from the steel industry, much less is known about the liquid C_1 feedstock, formate (Cotton et al. 2020; Köpke and Simpson 2020).

Unlike the other liquid C_1 feedstock methanol, formate can be efficiently produced directly from CO_2 and renewable electricity without the need for hydrogen (*i.e.* Power-to-X), a technology that is at the pre-commercialisation stage (Spurgeon and Kumar 2018; Rabiee et al. 2021). Being a liquid, it avoids many of the transportation issues present with gaseous substrates, as well as fitting with existing supply chain and fermentation infrastructure (Cotton et al. 2020). It also can overcome key mass transfer limitations faced by gas fermentation, and have higher energy efficiencies, resulting in lower operational costs, *e.g.* for like cooling (Cotton et al. 2020). Formate could, in the future, become not only a biotechnology feedstock, but an energy carrier that avoids the pitfalls of hydrogen such as non-negligible supply chain leaks leading to additional costs and contribution to global warming as a result of its GWP over 100 years of 11 (that is, 11 times that of CO_2) (Al-Breiki and Bicer 2020; Warwick et al. 2022).

Consequently, there have been numerous efforts to engineer formatotrophy into several organisms. For example, the synthetic reductive glycine pathway has been engineered into *E. coli* (Kim et al. 2020). The so-called FORCE pathway, an orthogonal chain elongation metabolism to biomass based on formyl-CoA, has also been engineered into *E. coli* (Chou et al. 2021). Despite these advances, native acetogen formatotrophy has the highest energy efficiency, which may be a critical metric when considering it as a Power-to-X technology (Wood et al. *In press*; Neuendorf et al. 2021). Utilising the native capabilities of acetogens such as *Eubacterium limosum*, *Butyribacterium methylotrophicum* and *Acetobacterium woodii*, would avoid the need for genetic engineering or building new gas fermentation infrastructure.

Formate as a substrate in the Wood-Ljungdahl Pathway (WLP) is similar to CO metabolism in that there is excess oxidation to CO_2 in order to generate the required number of reducing equivalents. We have previously shown for *E. limosum* in batch culture that formate consumption results in

65 acetate and CO_2 production at a stoichiometric ratio of 1:1 (Wood et al. 2021). Interestingly, the
66 observed native formatotrophic acetogen maximum growth rate was similar to that seen for the
67 synthetic reduction glycine pathway in *E. coli* (*ca.* 0.1 h^{-1}) (Kim et al. 2020).

68 Analytical quantification and metabolic modelling have advanced knowledge of acetogen
69 metabolism, allowing for improved process economics and fermentation designs (Marcellin et al.
70 2016; Heffernan et al. 2022). Transcriptome and translatome data have been analysed for *E.*
71 *limosum* (Song et al. 2017, 2018), and proteomic data for closely related *E. callanderi* (Kim et al.
72 2021). Some information has been published for the model acetogen, *A. woodii* (Neuendorf et al.
73 2021), yet no formatotrophic omics datasets exist for *E. limosum* to our knowledge. Since acetogen
74 metabolism is generally acknowledged to be controlled post-transcription and translation, omics
75 data, including metabolomics, may be key to revealing growth bottlenecks (Marcellin et al. 2016;
76 Mahamkali et al. 2020; Heffernan et al. 2022). Moreover, in contrast to batch datasets, steady-state
77 chemostat cultures offer better reproducibility across experiments, allowing greater insight into
78 process optimisation (Heffernan et al. 2020).

79 To understand formatotrophic growth and limitations, in this study, we investigated steady-state C_1
80 liquid formatotrophic fermentation in chemostat for *E. limosum*. Based on our previous batch
81 observations showing distinct growth phases when grown on formate and methanol, here we
82 hypothesised there would be limitations to continuous formate assimilation, perhaps similar to
83 carbon monoxide toxicity leading to oscillatory growth observed in *Clostridium autoethanogenum*
84 (Mahamkali et al. 2020). We tested this using differential analysis between formatotrophic and
85 methylotrophic conditions. Intriguingly the differences are related to both protein expression and
86 thermodynamics. Whilst growth was poor, unexpectedly, acetate was not the major product.
87 Furthermore, we observed a biofilm formation, which would negatively impact scalable production.
88 Anecdotally, adaptive laboratory evolution could be used to address this negative phenotype,

particularly given some *E. limosum* strains are known to not form sticky polymers that support biofilms (Flaiz et al. 2021).

We understand there is currently a metabolic model under curation for *E. limosum*, and therefore differential proteomic and metabolomic data presented here may be a powerful tool to help facilitate construction of the model (Bae et al. 2021; Fackler et al. 2021).

Methods/Experimental

Bacterial Strain, Growth Medium, and Continuous Culture Conditions

Eubacterium limosum ATCC 8486 (*E. limosum)* was subject to more than 200 generations of Adaptive Laboratory Evolution under liquid C_1 fermentation conditions (500 mM methanol, 100 mM formate). A single colony from this cell lineage was used in all experiments and stored as glycerol stocks at -80°C. This strain was cultivated anaerobically at 37 °C in a chemically defined phosphate buffered medium with liquid C_1 feedstocks (**Table 1**). The medium contained per litre: 0.5 g $MgCl_2.6H_2O$, 0.5 g NaCl, 0.13 g $CaCl_2.2H_2O$, 0.75 g NaH_2PO_4, 2.05 g Na_2HPO_4, 0.25 g KH_2PO_4, 0.5 g KCl, 2.5 NH_4Cl, 0.017 g $FeCl_3.6H_2O$, 0.5 g L-cysteine hydrochloride monohydrate, 1 mL of 1 g/L resazurin, 10 mL trace metal solution (TMS), 10 mL B-vitamin solution. The TMS contained per litre: 1.5 g nitrilotriacetic acid, 3 g $MgSO_4.7H_2O$, 0.5 g $MnSO_4.H_2O$, 1 g NaCl, 0.667 g $FeSO_4.7H_2O$, 0.2 g $CoCl_2.6H_2O$, 0.2 g $ZnSO_4.7H_2O$, 0.02 g $CuCl_2.2H_2O$, 0.014 g $Al_2(SO_4)_3.18H_2O$, 0.3 g H_3BO_3, 0.03 g $Na_2MoO_4.2H_2O$, 0.028 g $Na_2SeO_3.5H_2O$, 0.02 g $NiCl_2.6H_2O$ and 0.02 g $Na_2WO_4.2H_2O$. The B-vitamin solution contained per litre: 20 mg biotin, 20 mg folic acid, 10 mg pyridoxine hydrochloride, 50 mg thiamine-HCl, 50 mg riboflavin, 50 mg nicotinic acid, 50 mg calcium pantothenate, 50 mg vitamin B_{12}, 50 mg 4-aminobenzoic acid and 50 mg thioctic acid. The Rnf complex in *E. limosum* is Na^+ dependant, and therefore to maintain a consistent concentration across all experiments, NaCl was supplemented to medium which did not contain sodium formate as a substrate (*i.e.* chemostats growing on methanol). As a more oxidised substrate is required for methanol growth, gaseous CO_2 was also provided.

Steady-state conditions were reached in 0.5 L Multifors bioreactors (infors AG) controlled by EVE software at a working volume of 350 mL (magnetic marine impeller agitation). The system was equipped with peristaltic pumps, mass flow controllers (MFCs), pH and temperature sensors, and was connected to a Hiden HPR-20-QIC mass spectrometer (Hiden Analytical) for on-line off-gas analysis. Antifoam (Sigma 435503) was continuously added to the bioreactor using a syringe pump at 10 µL/h. Results presented here are after optical density, gas uptake/production rates, and acid/base addition rates were constant for at least five working volumes. Three technical replicate samples, spaced by one dilution volume, were collected per biological replicate.

Experimental analysis and Quantification

Extracellular analysis

Optical density (OD) measurements were taken at 600 nm via a UV-Vis spectrophotometer (Thermo Fischer Scientific Genesys 10S UV-Vis Spectrophotometer, USA). A biomass formula of $C_4H_7O_2N_{0.6}$ and 0.32 gDCW/L/OD was used to convert OD to molar cell concentrations (Wood et al. 2021). Liquid samples for extracellular metabolomic analysis were collected, filtered and stored at -20 °C. Samples were analysed by high performance liquid chromatography using an Agilent 1200 HPLC System with Phenomenex Rezex RHM-Monosaccharide H+ column (7.8 x 300 mm, PN: OOH-0132-KO) and guard column (Phenomenex SecurityGuard Carbo-H, PN: AJO-4490). Analytes were eluted isocratically with 4 mM H_2SO_4 at 0.6 mL min^{-1} for 48 min, and column oven temperature of 65 °C. 30 µL of sample was injected and monitored using a UV detector (210 nm) and RID at positive polarity and 40 °C.

Bioreactor off-gas analysis was performed by an on-line mass spectrometer, monitoring the amounts of H_2, Ar and CO_2 at 2, 40 and 44 amu respectively. To achieve reliable off-gas analysis, a bypass line from the feed gas bottle was used as the calibration gas for each MS-cycle (*i.e.* 'on-line' calibration). Specific rates (mmol/gDCW/d) were calculated by accounting for the exact gas composition, bioreactor liquid working volume, feed gas flow rate, off-gas flow rate (based on a constant flow of

138 inert Argon), the ideal gas molar volume, pH and CO_2 dissolution equilibrium, and the steady-state
139 biomass concentration.

140 As a quality check, gaseous samples were also taken for 'off-line' analysis using a Shimadzu 2014 GC,
141 equipped with a ShinCarbon packed column (ST 80/100, 2mm ID, 1/8 OD Silico, Restek). H_2 and
142 Argon were detected by a thermal conductivity detector (TCD), while CO_2 was measured using a
143 flame ionization detector (FID).

144 ## Intracellular analysis

145 Quantitative proteome analysis was carried out using direct data-independent acquisition mass
146 spectrometry approach (direct-DIA). Sampling, storage and sample preparation were performed
147 based on a method previously developed for autotrophic growth of *C. autoethangenum*. Briefly, 5
148 ODs of culture was pelleted by immediate centrifugation (16,000g for 3 minutes at 4 °C) followed by
149 washing with 1 mL PBS (Sigma P4417). A further round of centrifugation was used, with the
150 supernatant then discarded and the remaining sample stored at -20 °C. Approximately 100 µg of
151 protein was resuspended in 25 µL Milli-Q water and combined with 25 µL SDS lysis buffer (10% SDS,
152 100 mM Tris). Cells grown on formate required bead beating with a small amount of 0.1 mm
153 diameter glass beads to lyse. To reduce protein disulphide bonds, dithiothreitol (DTT) was added to a
154 final concentration of 20 mM before incubating at 70 °C for 1 hour. Iodoacetamide (IAA) was added
155 to a final concentration of 40 mM and incubated in the dark for 30 minutes to alkylate cysteine. The
156 reaction was sonicated and 2.5 µL 12% phosphoric acid was added, before combining with 165 µL S-
157 Trap binding buffer (90% methanol, 100 mM Tris (aq)). The sample was centrifuged for 8 minutes at
158 13,000 g, before adding to the S-Trap spin column, spun for 1 minute at 4000g and washed thrice
159 with 150 µL S-Trap binding buffer. 1 µg of trypsin in 50 µL of 50 mM ammonium bicarbonate pH 8
160 was added to cover the top of the protein trap, incubating in a sealed bag overnight at 37 °C.
161 Peptides were eluted into a collection tube with 40 µL increasing concentration of acetonitrile in

0.1% formic acid. Samples were spun dry and resuspended in 20 µL buffer A (0.1% formic acid (aq)) for injection to the mass spectrometer.

Mass spectrometry for proteomics analysis was performed using LC-MS/MS with a Thermo Fisher Scientific UHPLC system coupled to an Exploris 480 mass spectrometer. The electrospray voltage was 2.2 kV in positive-ion mode, and the ion transfer tube temperature was 250 °C. Full MS scans were acquired in the Orbitrap mass analyser over the range of m/z 340-1110 with a mass resolution of 120,000 (at m/z 200). The automatic gain control (AGC) target value was set at 'Standard', and the maximum accumulation time was 'Auto' for the MS/MS. The MS/MS ions were measured in 12 windows from mass 350-470, in 36 windows from mass 465-645 and 10 windows from mass 640-1100. Analysis of data were performed using Spectronaut against a reference proteome (UniProt ID UP000246246), with a Q-value cutoff of 0.05 applied to differential protein expressions. Locus tags correspond to (Song et al. 2017, 2018).

Intracellular metabolome analysis was based on the method previously developed for autotrophic growth of *C. autoethangenum* (Mahamkali et al (2020)). Briefly, 5 ODs of culture was pelleted by immediate centrifugation (16,000g for 3 minutes at 4 °C) followed by resuspension in chilled 50% acetonitrile to extract intracellular metabolites. Cell debris was removed by centrifugation with the supernatant stored at -80 °C. Samples were then freeze dried and resuspended in 90 µL 2% acetonitrile containing 5 µM azidothymidine standard. To remove lipophilic compounds (such as lipids, fatty acids, oil), the extract was fractionated by adding 250 uL of chloroform, before adding 410 µL MQ water, vortexing, and then separating the organic and polar phases by centrifugation. The polar fraction was cleaned through a spin column, freeze dried and resuspended with 2% acetonitrile. LC-MS/MS analysis was performed using a Shimadzu UHPLC System (Nexera X2) coupled to a Shimadu 8060 triple quadrupole (QqQ) mass spectrometer following (Espinosa et al. 2020; Mahamkali et al. 2020) with modifications and additions to the scheduled multiple reaction monitoring (sMRM) transitions. Chromatographic separation was performed on a Gemini NX-C18

187 column (3µm x 150mm x 2mm, PN: 00F4453B0, Phenomenex) with an ion-pairing buffer system
188 consist of mobile phase A: 7.5mM tributylamine (pH 4.95 with acetic acid) and mobile phase B:
189 acetonitrile. 5 µL and 10 µL of sample were injected to ensure measured intensities fell within the
190 standard curve ranges.

191 Cell pellets were also used for PHB content analysis, following the method of (de Souza Pinto
192 Lemgruber et al. 2019).

193 Thermodynamic and kinetic metabolic flux analysis

194 We evaluated reaction thermodynamic driving forces using the thermodynamic variability analysis
195 model presented by Mahamkali et al. (2020). pH was assumed as 1 unit higher than the extracellular
196 pH (Lindley et al. 1987). Reaction concentrations were constrained by measured values from
197 intracellular metabolomics (or the lower limit of quantification, LLOQ). Carboxylate products (*i.e.*,
198 acetate, butyrate *etc.*) intracellular concentrations were limited according to between *ca.* 1 times
199 higher, or 10^6 times lower, than extracellularly, depending on possible transport mechanism (Lindley
200 et al. 1987; Mahamkali et al. 2020). For formate substrate intracellular concentration, this was
201 instead between 10 times higher, or 50 times lower, than extracellularly (Refer to Dataset A).

202 Beyond that, minimum concentrations were set to be 0.1 µM for metabolites, and 1 nM for
203 dissolved gases unless directly measured or evaluated from Henry's law (H_2, CO_2 and CO). CO_2 total
204 concentration, c, was calculated from partial pressure, p, as,

205 $$c_{co2} = p_{CO2} \times \frac{25.2mM}{atm} \times (1 + 10^{pH-6.116}) \quad (1)$$

206 Maximum values were set to 1 mM for activated metabolites, and 10 mM for others.

207 A high-level estimate for reaction fluxes was determined as follows to estimate potential bottlenecks
208 in metabolism. Considering metabolic flux (J) is controlled by protein concentration (E), Gibbs free
209 energy of reaction (ΔG), saturation by substrates (M) with affinity (K), kinetic orders (a), and other
210 sources of regulation (U), according to (Heffernan et al. 2022),

$$J \sim E \times \left(1 - e^{\Delta G/RT}\right) \times \prod_i (M_i/K_i)^{a_i} \times U(M) \quad (2)$$

The enzyme and kinetics terms can be further simplified using Michaelis-Menten enzyme kinetics for a given substrate concentration (S) to give flux (J) as,

$$J \sim E \times k_{cat} \frac{S}{K_m + S} \times \left(1 - e^{\Delta G/RT}\right) \quad (3)$$

We can then make several assumptions to get likely order of magnitude estimates for J as follows.

- Our proteomic analysis can not determine absolute quantifications due to differences in mass spectrometer protein constituent signal response. However using a spiked protein analysis, (Valgepea et al. 2022) found a linear relationship between the $\log_{10}$ of mass spectrometer intensity and $\log_{10}$ anchor protein concentrations using a similar instrument setup to here. Since we can not be certain of the injected sample protein mass, we instead scaled protein concentrations by the pta enzyme, which Valgepea et al. (2022) showed was largely consistent across different gas fermentation conditions.
- Few *E. limosum* proteins have been assayed, and therefore instead we undertook BLAST analysis of key proteins against enzymes with known kinetic parameters, k_{cat} and K_m. We then used this as verification data against deep learning models, where protein amino acid sequences were used to yield a complete set of kinetic parameters for the WLP central carbon metabolism (Kroll et al. 2021; Li et al. 2021). Where intracellular substrate concentrations, S, were unknown, instead k_{cat}/K_m was taken as a proxy for kinetic effects.

Results

Steady state fermentation

Formatotrophic grown cells reached steady-state at a dilution rate (D) of 0.4 d^{-1} (specific growth rate of 0.017 h^{-1}). The dilution rate was chosen as the highest achievable across formatotrophic and methylotrophic growth conditions, based on batch bottle growth rates. Despite the same amount of

carbon being supplied to each condition, formate reached only a biomass concentration of 0.07±0.02 gDCW/L, less than 15% of methylotrophic growth. We note however, that there was significant biofilm formation in the formate condition, and so biomass concentration as reported here may be underestimated. As a result, substrate-specific uptake rates were much higher for the formate condition but potentially overestimated: 280±56 mmol/gDCW/d compared to 75±7.0 mmol/gDCW/d for methanol in the methanol/CO_2 condition (**Figure 1A**).

Formatotrophic growth could only be achieved when no CO_2 gas was supplied to the culture. Further, a formate titre of 2.0±0.18 g/L was observed at steady-state (**Table 1**), indicating the culture was not formate limited, but rather constrained by another parameter.

As expected, formate-grown cells showed significant CO_2 evolution (150±11 mmol/gDCW/d) (a formate consumption to CO_2 production ratio of 1.9±0.4 mol/mol). However, unexpectedly, the major non-CO_2 product under formatotrophic growth was not acetate. Only trace amounts of acetate were detected throughout the entire fermentation (0.92±0.35 mmol/gDCW/d) (**Figure 1A**). Other typical *E. limosum* products such as butyrate (0.32±0.13 mmol/gDCW/d), hexanoate and butanol, were minor or below the detection limit (**Figure 1A**).

Hydrogen was produced in both cultures as a minor sink of electrons. Intriguingly, despite formate being the less reduced substrate, it produced more hydrogen at 21±7.8 mmol/gDCW/d compared to 1.1±0.56 mmol/gDCW/d for methanol/CO_2 (**Figure 1A**). We noted in Wood *et al.* (2022) that formate seems to trigger hydrogen production, which is further reinforced with this data.

Overall, this gave a carbon and electron balance of 63±15% and 24±6% for formatotrophic growth (**Figure 1B**). These values suggest the production of an unknown product with a redox number of 5.0±2.8. Given none of the known products which branch from acetyl-CoA were detected (*e.g.* acetate, butyrate, hexanoate, butanol *etc.*), we hypothesised there may be a product downstream of pyruvate. Whilst no genes are identified in the *E. limosum* genome for PHB production, given the unknown product redox number, we analysed cell pellets for intracellular PHB content. Formate

cultures showed 0.68±0.63% wPHB/wDCW. This is sufficiently low that the effect on carbon and electron balances is negligible. We also measured for 4-aminobutyrate (GABA) (potentially produced *via* partial TCA cycle) and 3-methyl-2-oxopentanoate (potentially produced *via* isoleucine biosynthesis) using HPLC analysis against standards. As neither of them could be found, we also looked for intermediates along the GABA pathway (2-ketoglutarate), and 3-methyl-2-oxopentanoate pathway (citraconate, 2-oxobutanoate). The analysis confirmed that none of these compounds were significantly detected in the samples, although the extracellular 2-ketoglutarate concentration was above the LLOQ for the formatotrophic condition, yet it was not detected for methylotrophic growth (data not shown).

Intracellular conditions

Given the lack of success in identifying the missing product, we performed proteomics and metabolomics analyses to gain an understanding of formatotrophic *E. limosum* metabolism, and compared it to methanol metabolism (Wood et al. 2022). We identified 934 proteins that met the criteria for differential analysis between the formate and methanol/CO_2 growth conditions (*i.e.* present in both samples and Q-value < 0.05). Proteome analysis detected all the key enzymes of the WLP, energy conservation, direct acetyl-CoA condensation pathway, carboxylate/alcohol production, central carbon metabolism, shikimate pathway, partial TCA cycle, isoleucine biosynthesis, and biofilm switches (**Figure 2** and **Figure 3**). Across these pathways, 31 intracellular metabolites were identified in both conditions. Coenzyme A was below the lower limit of quantification (LLOQ) for all conditions.

Formate is assimilated via the formate dehydrogenase elim_c2470 and FTHFS elim_c0957 which have $\log_2$ fold changes of 1.7 and 2.1 and are respectively the 66th and 8th most 'abundant' proteins. Given the higher specific flux through the WLP for formate, it is not surprising there is upregulation generally (elim_c0959, 1648, 1651, 1655). That said, the $\log_2$ fold changes are mostly small and less than 2. The WLP proteins which are more regulated in the methanol/CO_2 condition mostly relate to

284 the thermodynamic difficulty in reversing methylene reductase (elim_c0960-63). Several of the WLP
285 acs/codh complex were the most abundant in the formatotrophic proteome – elim_c1655, 1651.
286 Yet, acetyl-CoA intracellular concentration was one order of magnitude lower (0.014±0.003
287 µmol/gDCW compared to 0.300±0.133 µmol/gDCW). For reference, *C. autoethanogenum* cells under
288 syngas growth had an acetyl-CoA concentration of 0.358±0.140 µmol/gDCW (Mahamkali et al.
289 2020).

290 Strangely, mtaB (elim_c3293), part of the methyltransferase operon for methanol assimilation into
291 the WLP was also one of the most 'abundant' proteins in the formatotrophic proteome. Song and
292 colleagues showed no transcriptome fold change in mtaB between heterotrophic and autotrophic
293 (H_2/CO_2) growth, and low relative expression, suggesting that our observation here shows
294 formatotrophic and methylotrophic growth may share features not seen in other autotrophic
295 conditions (Song et al. 2018).

296 When it comes to energy conservation, there was little change in etf proteins (elim_c0229, 0230),
297 nor Rnf proteins detected (elim_c3882, 3884, 1224), yet ATPase was strongly downregulated in the
298 formate condition (elim_c3452, 3454, 3455, 3670) with $\log_2$ fold changes greater than 2. ATP/ADP
299 intracellular ratios were relatively consistent at 1.4±0.07 for the formate condition, and 1.3±0.11 for
300 methanol/CO_2, noting however the ATP/ADP/AMP pool size was one order of magnitude higher in
301 the latter case. The sporomusa type Nfn complex was upregulated in the formate condition
302 (elim_c1993 (Kremp et al. 2020)) with a $\log_2$ fold change of 2. The NADPH/NADP ratio was 0.40±0.25
303 and 1.0±0.056 for formate and methanol/CO_2 respectively, with the same order of magnitude total
304 pool size. The NADH/NAD ratios were 0.035 and 0.0011±0.0006 respectively, however the total pool
305 size for methanol/CO_2 was one order of magnitude higher. The formate condition ratio is similar to a
306 typical syngas fermentation at NADH/NAD of 0.019 and NADPH/NADP of 0.276 (Mahamkali et al.
307 2020), which probably reflects formate and syngas operate the WLP and Nfn in the same direction,
308 unlike methanol.

Considering pathways to produce carboxylates from acetyl-CoA (elim_c0231-0234, 2806, 2339) there is mostly little change in expression (log_2 fold change < 1), reinforcing that this metabolism is regulated post-translation as acetate and butyrate were not major products under formatotrophic growth. Both acetoacetyl-CoA and 3-hydroxybutyryl-CoA were not detected, supporting the observation of no butyrate production. *E. limosum* codes for two different pta enzymes, one of which is significantly upregulated under methylotrophic growth (elim_c1218; log_2 fold change 2.4), and the other significantly upregulated under formatotrophic growth (elim_c1884; log_2 fold change 2.7). Despite the upregulation of elim_c1884, no acetyl-phosphate was detected for the formate condition, yet 0.294±0.177µmol/gDCW was detected under methanol/CO_2 growth. This probably reflects the difference in acetyl-CoA/CoA driving force. Whilst cells did not produce alcohols under any condition, it is interesting to note identification of acetylating aldehyde dehydrogenase (elim_c1223), AOR (elim_c3128, 0919) and alcohol dehydrogenase (elim_c2439, 2953, 0460). None of these had log_2 fold changes greater than 1.5 between the two conditions, yet it is worth noting elim_c3128 was in the top 5% of proteome 'abundance'.

Carbon flux from the WLP is linked to pyruvate in glucogenesis via PFOR (elim_c2885) or alternatively Pfl (elim_c0889), the latter of which is upregulated with a log_2 fold change of 5.4 for formatotrophic growth. This is the 9th most upregulated formatotrophic gene indicating that cells may utilise formate directly to synthesise pyruvate, rather than oxidise it to generate ferredoxin and CO_2 for use by PFOR. Pyruvate was not detected in the methanol/CO_2 condition (LLOQ 0.0016 µmol/gDCW), yet it was 0.060±0.044 µmol/gDCW in formate growth.

The remainder of the glucogenesis pathway (elim_c3055, 0452, 3997, 0104, 0106, 0107, 0421, 2869) all have log_2 fold changes less than 1.5. Two notable exceptions are TPI (elim_c0105) and FBA (elim_1932, 2267) which are downregulated for formatotrophic growth (log_2 *ca.* 2), along with parts of the pentose phosphate pathway (elim_c0386, 0631, 1191) (log_2 fold change *ca.* 2-3). Of all the measured intermediates, only FBP was higher under the formate growth condition (yet still below

the LLOQ). This seems counterintuitive based on protein expression, unless FBP flux also goes elsewhere for methylotrophic growth.

Branching off the glucogenesis and pentose phosphate pathways, the shikimate pathway proteins all have $\log_2$ fold changes less than 1.5, with the sole exception of aroKL (elim_c0110) being downregulated in formatotrophic growth with a $\log_2$ fold change of 1.8. PEP and E4P, which enter the shikimate pathway, have lower concentrations for the formate condition, compared to methylotrophic growth (0.028±0.009 µmol/gDCW v. 0.060±0.018 µmol/gDCW, and N.D. v 0.058±0.012 µmol/gDCW). Downstream of the shikimate pathway, anthranilate was not detected in the former condition, and was 0.002±0.0004 µmol/gDCW in the latter case.

The partial TCA cycle, and isoleucine biosynthesis pathways both have three of the most upregulated proteins under formatotrophic growth (11^{th}, 28^{th} and 43^{rd}). GAD (elim_c2464) and its antiporter (elim_c2465) have $\log_2$ fold changes of 5.0 and 3.6 respectively, suggesting 4-aminobutyrate (GABA) may be an extracellular metabolite. Similarly, LeuDH (elim_c1793) has a $\log_2$ fold change of 3.9, suggesting 3-methyl-2-oxopentanoate may also be an extracellular metabolite. ACN (elim_c1420), which reversibly converts citrate and isocitrate was upregulated in the formate condition by a $\log_2$ fold change of 1.7. Yet, intracellular metabolite concentrations were the same in the two conditions (0.127±0.027 µmol/gDCW v. 0.136±0.026 µmol/gDCW, and 0.093±0.017 µmol/gDCW v 0.103±0.022 µmol/gDCW). Interestingly FUM (elim_c3042, 3043) were both downregulated with a $\log_2$ fold change of 1.4 and 2.0 respectively, which seems to correlate with the substrate, malate, having a higher intracellular concentration for formate (0.186±0.077 µmol/gDCW) compared to methanol/CO_2 (0.056±0.018 µmol/gDCW). Isoleucine biosynthesis cimA (elim_c3964) and leuCD (elim_c3962) were downregulated by $\log_2$ fold changes of 4.3 and 2.6 respectively. The remainder of the partial TCA cycle and isoleucine biosynthesis pathways showed no significant change in regulation.

We found several significantly 'abundant' proteins that have not been annotated. These include, for example, elim_c2694 (log_2 fold change 7.8, 22nd most 'abundant'), yet we note that it had no significant transcriptome fold changes ($log_2 < 1.5$) between heterotrophic and autotrophic growth, indicating it may be specific to C_1 liquid metabolism (Song et al. 2018). Further, there are also upregulated proteins for which the annotation is ambiguous. For example, elim_c2880 is the most upregulated protein with a log_2 fold change of 10.4, as well as being the 5th most 'abundant' protein. It is annotated as a BadM/Rrf2 transcriptional regulator, associated with benzoyl-CoA degradation, even though benzoate was not present in the medium. We do note BadL, used in PABA metabolism, is part of this operon although not annotated in *E. limosum* (Vandrisse and Escalante-Semerena 2018). Therefore, this could *potentially* mean there is a PABA nutrient limitation (noting that PABA was not detected intracellularly in either condition). Similarly, elim_c1937 is a folate transport upregulated for formatotrophic growth (log_2 fold change of 3.8 compared to methylotrophic growth), indicating a potential THF limitation.

Discussion

Metabolic model and redox balances

Thus far we have presented 'omics data for formatotrophic *E. limosum* cultures, with differential analysis against methylotrophic growth. Whilst valuable conclusions can be drawn, the data takes on further meaning when integrated together in a complete kinetic model. Since such a model does not exist for *E. limosum*, we have instead undertaken thermodynamic calculations and various kinetics assumptions to identify several potential metabolic bottlenecks (refer to materials and methods).

Unlike NADH/NAD and NADPH/NADP, we could not measure ferredoxin. However, upper and lower bounds on the ferredoxin state could be calculated. From the Stn, Fd^{2-}/Fd must be > 0.075. Using Henry's law, the dissolved hydrogen must be *ca.* 0.6 µM and so the hydrogenase suggests Fd^{2-}/Fd must be > 2.4. With the Rnf/ATPase complex operating in the forward direction for formate metabolism (*i.e.*, generating ATP), Fd^{2-}/Fd should be > 1.5. For PFOR to generate pyruvate, and

assuming a CoA concentration of <43 µM (LLOQ), Fd^{2-}/Fd must be <60. However, considering the predicted kinetic parameters, it is likely that the Fd^{2-}/Fd ratio is on the lower end (*ca.* 4) to generate sufficient flux through the mthfr, codh and fdh. However, given that ferredoxin is mainly used in the WLP and pyruvate synthesis, this wide range of possible values ($2.4 < Fd^{2-}/Fd < 60$) does not provide further clues as to identification of the unknown product.

Figure 4 illustrates how methanol and formate metabolism differ in the WLP by reversing key enzymes. The hydrogenase thermodynamic driving force is much larger under formatotrophic growth, which may explain why we observed higher hydrogen evolution. The methanol condition has an acetyl-CoA/CoA ratio three orders of magnitude higher than the formate condition, yet we calculated a similar thermodynamic driving force. This is because, in the latter condition, a much lower acetate concentration is possible, hence why it is in fact possible for pta to reverse under formatotrophic growth. We calculated this acetate limit to be in the *ca.* 1mM range (using eQuilibrator (Flamholz et al. 2012) with acetyl-P = 12.5µM, ATP/ADP = 1.5, pH = 7.9, pMg = 3.0, ionic strength = 0.25). This confirms, that despite significant upregulation of the pta (elim_c1884), acetate production is limited by post-translation effects. Additionally, we cannot identify how thl is feasible to satisfy butyrate flux, without the acetoacetyl-CoA concentration being below our assumed minimum value of 0.1 µM, which would surely impose kinetic limitations. Thus, the model confirms our suspicion that acetyl-CoA must be diverted elsewhere (*i.e.,* towards pyruvate).

Interestingly, the upregulated Pfl is probably the major source of pyruvate, considering thermodynamics (**Figure 4**), protein abundance and k_{cat}/K_m (**Table 2**). Yet, the model indicates PFOR may in fact operate in the pyruvate consuming direction during formatotrophic growth, causing a futile metabolic cycling of pyruvate. Considering thermodynamics, protein 'abundance' and k_{cat}/K_m (**Table 2, Figure 4**), these reinforce a mix of pre- and post-translation effects control metabolism and may explain the lack of production of GABA (no feasible export) and 3-methyl-2-oxopentanoate (kinetic bottleneck).

Given the Stn expression, NADPH likely has greater flux in the formate condition than methanol. This is interesting because the NADPH/NADP is lower for formate. This could be a clue for product identification, as production may be tied to the need to regenerate NADP. However, it is noted this would come at an ATP cost, as ferredoxin is diverted away from Rnf/ATPase chemiosmotic ATP generation, instead towards Stn to generate NADPH.

Cellular stress and pyruvate

Formate metabolism in acetogens appears partially controlled by protein expression in response to cellular state. For example, since formatotrophic growth compared to methanol/CO_2 has a much lower CO_2 driving force, it makes sense that most pyruvate flux would be through heavily upregulated Pfl, despite PFOR being the normal mechanism for pyruvate synthesis (Neuendorf et al. 2021). Given the low acetyl-CoA/pyruvate ratio, Pfl would require a high formate concentration as driving force, which may explain why a formate titre existed at steady-state. Previously, researchers have suggested an increased pyruvate concentration is regulated by PPDK activity and ATP availability, causing a drop in glucogenic intermediates for *E. limosum* (Lebloas et al. 1996). Similarly, we found that formatotrophic growth had lower glucogenic intermediates and no change in PPDK levels (elim_c3055) compared to methylotrophic growth. Specifically, we can see that the PEP / pyruvate ratio is respectively 1.3±0.61 and >39±12 (adopting LLOQ for pyruvate where not detected). This means PPDK activity is being limited by ATP availability, and so increasing pyruvate concentration and hence the required formate driving force. This associated energy shortage allows redirection of carbon (Lebloas et al. 1996).

We can see this redirection of carbon in the lack of acetate production, and futile upregulation of the pta enzyme. Given that metabolites on the butyrate production pathway were not detected, this supports the idea that a significant amount of carbon goes through pyruvate.

Of all measured intracellular metabolites, most were relatively similar for both methanol and formate conditions. However, several which were much lower or below LLOQ for formate only.

Some of these are on the glucogenesis pathway, which is controlled by ATP availability discussed above. We found that pools of NAD, ATP and acetyl-CoA to be one order of magnitude lower for formatotrophic growth. This is also significantly lower than typical acetogens, despite the NADH/NAD, ATP/ADP ratios being conserved relative to other conditions (Mahamkali et al. 2020). Clearly, these ratios are tightly controlled, despite the lower ATP gain from growth, but the overall pool size is not.

It is hard to understand why ATP, NAD and CoA pool sizes are smaller for formatotrophic growth compared to methylotrophic growth based on the proteomics data given that two proteins involved in NAD biosynthesis (elim_c1173, 3031) were upregulated. The remaining two (elim_c1197, 1880) were not detected. This includes elim_c1880, which is used to synthesise NADP from NAD, despite NADP(H) being detected as a metabolite. In CoA biosynthesis, only one (elim_c1486) of the four proteins were identified in the formate condition, compared to three (elim_c1486, 3523, 2435) in the methanol condition. The lack of protein detection may correlate with the reduced ability of cells to synthesise NAD and CoA. GTP is required for AMP synthesis (elim_c1406, 2669), and since the GMP flux is not directed to GTP in formatotrophic growth, this may affect the ATP pool size. Some researchers have found pool size can be reduced by cellular stress conditions, such as pH and temperature (Chohnan et al. 1998).

Formate itself may be cause of stress

Carboxylate transport in acetogens is possible *via* passive diffusion, symport of anion with a proton, transport of anion *via* a uniport and ATP-consuming ABC transporter (Mahamkali et al. 2020). Given we are already energetically limited, we can discount the latter process. An anion uniport (*e.g.* elim_c1361 $\log_2$ fold change of 0.7) must balance the charge with independent proton uptake, which could be *via* an ATPase, thus assisting with the energetic requirements. This would yield a lower intracellular formate concentration driving force (1 mM; Dataset A), however, our model predicts the resulting intracellular formate driving force is still sufficient for cell function.

Given acid was added throughout the experiment to maintain extracellular pH, there must have been a net alkalinisation of the cell, which may be the cause of cellular stress. This would lead to a more negative membrane potential that might be balanced by importing protons through an ATPase. This could theoretically improve the overall ATP gain. It is important to note *E. limosum* ATPases are generally thought to be sodium-dependent and we also did not find a significant upregulation of any ATPase under formatotrophic growth (Song et al. 2017; Kang et al. 2020). Whilst there was clearly an ATP limitation, the ATP balance itself thus remains unclear.

Additionally, formate metabolism is known to lead to a loss of the acetyl-phosphate pool via formyl-CoA (and hence ATP) (Sly and Stadtman 1963). However, we have not identified the culprit CoA-hydrolase enzyme in the proteome. Overall, we speculate that cellular stress caused an energy shortage, increased pyruvate concentration, which then allowed redirection of carbon.

What does AOR do?

We found high 'abundance' of AOR, with little difference between the two conditions. Previous researchers have noted that acetogens are regulated post-translationally (Marcellin et al. 2016; Mahamkali et al. 2020; Heffernan et al. 2022), and we believe this could be the case for AOR. Whilst the enzyme is known to be promiscuous, AOR is generally acknowledged not to have C_1 specificity for formate (Bertsch and Müller 2015). The same can be said for adh. The kinetic deep learning model, however, suggests formate consumption may be possible (as k_{cat} and K_m are similar to C_2 and C_4 substrates, but adh k_{cat} is one order of magnitude lower for C_1; data not shown), but would result in only *ca.* 0.2 µM accumulation of methanol, which is below our limit of detection. In principle, it is possible, which may explain our previous observation of methanol production from formate during early growth stages, presumably when there is a highly reduced redox pool (Wood et al. 2021).

CO_2 must be maintained low to allow growth in chemostat

We found cells were unable to grow formatotrophically when 15% CO_2 gas was present in the headspace. Growth was restored, and steady state was reached when the gas feed contained zero

483 CO_2. If this requirement was thermodynamics related, the bottleneck reaction would therefore
484 produce CO_2. All else being equal, increasing the CO_2 concentration means the formate driving force
485 must increase, and there must be an upper limit to intracellular formate concentration in terms of
486 toxicity.

487 Biofilm formation

488 Two *E. limosum* cyclic-di-GMP-I riboswitches are known to regulate virulence, motility, quorum
489 sensing and biofilm (Song et al. 2017). Here, under formatotrophic growth we found upregulation of
490 two proteins which synthesise cyclic-di-GMP from GMP – elim_c1151, 3120 (log_2 fold change of 6.2
491 and 3.1 respectively). Intracellular GMP concentrations were similar for both formatotrophic growth
492 and methylotrophic growth at 0.008±0.005 µmol/gDCW and 0.007±0.002 µmol/gDCW, respectively.
493 GMP can also be used for GDP and GTP synthesis and interestingly these proteins (elim_c1967,
494 1274) are downregulated for formatotrophic growth (log_2 fold change 1.2 and 2 respectively).
495 Additionally, unlike the methanol condition, no GDP or GTP were detected under formate growth,
496 reinforcing that GMP flux went elsewhere. Therefore, we confirm cyclic-di-GMP has a correlation
497 with biofilm formation strength.

498 We also note GlmU (elim_c1471), which produces a precursor for the adhesion of biofilms in some
499 bacteria, was upregulated in the formate condition (log_2 fold change of 1.1) (Burton et al. 2006).
500 Interestingly GlmU transcriptome shows upregulation in heterotrophic growth, compared to
501 autotrophic, which matches the past *E. limosum* biofilm formation observations (Song et al. 2018).
502 LuxR (elim_c1099) is a transcriptional regulator known to affect biofilm formation through quorum
503 sensing, and here it was upregulated in formatotrophic growth (log_2 fold change of 2.4) (Chen and
504 Xie 2011). There was little difference in one identified protein related to cell motility (twitching)
505 (elim_c0255).

506 From the 'omics data, biofilm formation in *E. limosum* formatotrophic growth is most likely caused
507 by cyclic-di-GMP production, adhesive compounds and/or quorum sensing. Inhibitors for adhesive

compounds produced by GlmU include N-acetyl glucosamine-1-phosphate and iodoacetamide (Chen et al. 2015). The latter reacts with cysteine in our medium, and so is not an option here. Anthranilate, a tryptophan metabolite, is known to reduce intracellular cyclic-di-GMP, which should reduce biofilm formation (Li et al. 2017). Other metabolites on the tryptophan pathway, including indole and D-tryptophan have also been shown to influence biofilm formation. However what works for one species, can even enhance formation in other species (Li et al. 2017).

Interestingly, we did not detect anthranilate in the formate condition, but it was 0.002±0.0004 µmol/gDCW under methylotrophic growth. Anthranilate is produced from S3P, which was three times higher in the formate condition (0.006±0.003 µmol/gDCW versus 0.002±0.0003 µmol/gDCW for methanol). This trend suggests that there is more flux to anthranilate, but its concentration is being kept low, perhaps to regulate intracellular cyclic-di-GMP. Hence more biofilm formed in the formate condition.

To investigate this relationship, we supplemented the medium with 0.8 mM anthranilate in a culture with a pre-formed biofilm. Li and colleagues noted approximately this amount of anthranilate reduced pre-formed biofilm coverage by up to 30% for three different species of bacteria (Li et al. 2017). However, to our surprise, this concentration of anthranilate had no visual effect on the formed biofilm (data not shown). It is possible anthranilate may only prevent biofilm formation in *E. limosum*, which we have not tested here. Alternatively, other biofilm inhibitors may need to be tried, such as N-acetyl glucosamine-1-phosphate to prevent adhesive compound synthesis (Chen et al. 2015).

Conclusion

We have established a baseline formatotrophic steady-state dataset for *E. limosum* covering phenomics, proteomics and metabolomics. Cells had a high formate-specific uptake rate. However, unexpectedly they did not produce acetate as a major product. There is evidence of a cellular energy limitation, which resulted in an accumulation of pyruvate intracellularly, and depletion of the total

533 CoA and NAD pools. Together, these redirected carbon flux away from acetate, despite upregulation
534 of pta, and towards products downstream of pyruvate. We contend this state of cellular stress
535 ultimately relates to formate itself as a substrate, which cells attempt to overcome through protein
536 expression, for example, upregulation of Pfl. This is interesting because acetogen metabolism is
537 usually controlled post-translation.

538 Despite having high energy efficiency, there are significant challenges to overcome for *E. limosum*
539 formatotrophy to be a scalable technology. Firstly, limiting biofilm formation, either through
540 medium supplementation or adaptive laboratory evolution, will be important to reduce operational
541 costs. Formatotrophy also needs to overcome poor yields and limit excess CO_2 production,
542 necessitating another substrate such as hydrogen or methanol. Adopting either of those would
543 partially defeat the purpose of using formate (*i.e.,* liquid feedstock that can be directly synthesised
544 from CO_2, fits with existing infrastructure and has no mass transfer limitations). Throughout this
545 work, we have identified exciting aspects of formate metabolism that could have significance for
546 acetogen applications, which include more targeted products downstream of pyruvate.

547

Abbreviations

General: Dilution rate; D, Lower Limit of Quantification; LLOQ; Optical Density, OD; Wood-Ljungdahl Pathway, WLP

Compounds: 2PG, 2-phosphoglycerate; 3DHQ, 3-dehydroquinate; 3DHSA, 3-dehydroshikimate; 3PG, 3-phosphoglycerate; Ac-CoA, acetyl-CoA; BPG, 1,3-biphosphoglycerate; DAHP, 3-deoxy-arabino-heptulonate 7-phosphate; DHAP, Dihydroxy acetone phosphate; E4P, Erythrose 4-phosphate; EPSP, 3-enolpyruvyl-shikimate 5-phosphate; F6P, Fructose 6-hosphate; FBP, Fructose 1,6-biphosphate; G3P, Glyceraldehyde 3-phosphate; G6P, Glucose 6-phosphate; Glc, Glucose; PEP, Phosphoenolpyruvate; PRPP, 5-phosphoribosyl diphosphate; PYR, Pyruvate; R5P, Ribose 5-phosphate; RL5P, Ribulose 5-phosphate; S3P, Shikimate 3-phosphate; S7P, Sedoheptulose 7-phosphate; THF, tetrahydrofolate; X5P, Xylulose 5-phosphate

Enzymes: ALD, aldehyde dehydrogenase; ACK, acetate kinase; ACN, anonitase; ACS, acetyl-CoA synthase; ADH, alcohol dehydrogenase; AOR, aldehyde:ferredoxin oxidoreductase; aroA, DAHP synthase; aroA, 3-phosphoshikimate 1-carboxyvinyltransferase; aroB, 3-dehydroquinate synthase; aroC, chorismate synthase; aroD, 3-dehydroquinate dehydratase; aroKL, shikimate kinase; BK, butyrate kinase; cimA, citramalate synthase; CODH, CO dehydrogenase; crt, crotonase; CS, citrate synthase; ENO, enolase; Etf-Bcd, butyryl-CoA dehydrogenase; FBA, Fructose biphosphate aldose; FBP, fructose 1; FDH, formate dehydrogenase; FTHFS, formyl-THF synthetase; FUM, fumarase; GAD, glutamate decarboxylase; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; GDH, glutamate dehydrogenase; GLK, Glucokinase; hbd, 3-hydroxybutyryl-CoA dehydrogenase; IDH, isocitrate dehydrogenase; ilvBN, pyruvate:2-oxobutanoate acetaldehydetransferase; ilvC, 2-aceto-2hydroxy-butanoate:NADP+ oxidoreductase; ilvD, dihydroxy acid dehydratase; ilvE, branched chain amino acid aminotransferase; KOR, oxoglutarate ferredoxin oxidoreductase; leuB, methylmalate dehydrogenase; leuCD, methylmalate hydrolase; LeuDH, leucine dehydrogenase; MDH, malate dehydrogenase; MtaABC, methanol dependent methyl transferase; MTHFC, methenyl-THF

cylcohydrolase; MTHFD, methylene-THF dehydrogenase; MTHFR, methyltransferase / methylene-THF reductase; PC, pyruvate carboxylase; PD, pyruvate dehydrogenase; PFKA, Phosphofructose kinase; PFL, pyruvate formate ligase; PFOR, pyruvate ferredoxin oxidoreductase; PGAM, phosphoglycerate mutase; PGI, Glucose 6-phosphate isomerase; PGK, phosphoglycerate kinase; PK, pyruvate kinase; PPDK, phosphoenolpyruvate synthetase; PRPS, Ribose phosphat epyrophosphokinase; PTA, phosphotransacetylase; PTB, phosphotransbutyrylase; PTS, phosphotransferase system; RPE, Ribulose phosphate epimerase; RPI, Ribose phosphate isomerase; Stn, Sporomusa type Nfn transhydrogenase; TAL, Transaldolase; thl, thiolase; TKT, Transketolase; TPI, Trisephosphate isomerase; trpE, anthranilate synthase; ydiB, shikimate dehydrogenase

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests

Funding

JW acknowledges support of the Warwick and Nancy Olsen Scholarship. BV acknowledges the support of the Australian Research Council (ARC) through grant FL170100086 and LP200200136. EM acknowledges the support of the ARC CoE in Synthetic Biology CE200100029 and LP200200136. The research utilised equipment and support provided by the BPA-funded facility, Queensland Metabolomics And Proteomics (Q-MAP), an Australian Government initiative being conducted as part of the National Collaborative Research Infrastructure Strategy (NCRIS) National Research Infrastructure for Australia .

Authors' contributions

JCW designed and conducted experiments, interpreted results and was a major contributor in writing the manuscript. RAGG jointly conducted experiments. DD performed intracellular metabolomics analysis. GT performed proteomics analysis. MRP performed extracellular metabolomics and PHB analysis. EM and BV conceived experiments and provided substantiative manuscript review. All authors read and approved the final manuscript.

Acknowledgements

Not applicable

References

Al-Breiki M, Bicer Y (2020) Comparative cost assessment of sustainable energy carriers produced from natural gas accounting for boil-off gas and social cost of carbon. Energy Reports 6:1897–1909. https://doi.org/10.1016/j.egyr.2020.07.013

Bae J, Song Y, Lee H, Cho B (2021) Reconstruction of a genome-scale model of Eubacterium limosum converting C1 substrates to multi-carbon products

Bertsch J, Müller V (2015) Bioenergetic constraints for conversion of syngas to biofuels in acetogenic bacteria. Biotechnol. Biofuels 8

Burton E, Gawande P V., Yakandawala N, et al (2006) Antibiofilm activity of GlmU enzyme inhibitors against catheter-associated uropathogens. Antimicrob Agents Chemother 50:1835–1840. https://doi.org/10.1128/AAC.50.5.1835-1840.2006

Buttlaire DH (1980) Purification and properties of Formyltetrahydrofolate Synthetase. Methodsin Enzymol 66:585–599

Chen J, Xie J (2011) Role and regulation of bacterial LuxR-like regulators. J Cell Biochem 112:2694–2702. https://doi.org/10.1002/jcb.23219

Chen S, He N, Yu J, et al (2015) Inhibitory effect of 2-mercaptoethane sulfonate on the formation of Escherichia coli biofilms in vitro. Mol Med Rep 12:5223–5230. https://doi.org/10.3892/mmr.2015.4112

Chohnan S, Izawa H, Nishihara H, Takamura Y (1998) Changes in Size of Intracellular Pools of Coenzyme A and Its Thioesters in Escherichcia coli K-12 Cells to Various Carbon Sources and Stresses. Biosci Biotechnol Biochem 62:1122–1128

Chou A, Lee SH, Zhu F, et al (2021) An orthogonal metabolic framework for one-carbon utilization. Nat Metab 3:1385–1399. https://doi.org/10.1038/s42255-021-00453-0

Clark J, Ljungdahl LG (1982) Purification and properties of 5,10-methenyltetrahydrofolate cyclohydrolase from Clostridium formicoaceticum. J Biol Chem 257:3833–3836. https://doi.org/10.1016/s0021-9258(18)34857-9

Cotton CA, Claassens NJ, Benito-Vaquerizo S, Bar-Even A (2020) Renewable methanol and formate as microbial feedstocks. Curr Opin Biotechnol 62:168–180. https://doi.org/10.1016/j.copbio.2019.10.002

Davies JA (2017) Characterisation of the reversible formate dehydrogenases of Shewanella. University of East Anglia

de Souza Pinto Lemgruber R, Valgepea K, Tappel R, et al (2019) Systems-level engineering and characterisation of Clostridium autoethanogenum through heterologous production of poly-3-hydroxybutyrate (PHB). Metab Eng 53:14–23. https://doi.org/10.1016/j.ymben.2019.01.003

Espinosa MI, Gonzalez-Garcia RA, Valgepea K, et al (2020) Adaptive laboratory evolution of native methanol assimilation in Saccharomyces cerevisiae. Nat Commun 11:5564. https://doi.org/10.1038/s41467-020-19390-9

Fackler N, Heijstra BD, Rasor BJ, et al (2021) Annual Review of Chemical and Biomolecular Engineering. Annu Rev Chem Biomol Eng. https://doi.org/https://doi.org/10.1146/annurev-chembioeng-120120-021122

Flaiz M, Ludwig G, Bengelsdorf FR, Durre P (2021) Production of the Biocommodities Butanol and

644 Acetone from Methanol with Fluorescent FAST-tagged Proteins using Metabolically Engineered
645 Strains of Eubacterium Limosum. Biotechnol Biofuels 14:117:
646 https://doi.org/https://doi.org/10.21203/rs.3.rs-171102/v1

647 Flamholz A, Noor E, Bar-Even A, Milo R (2012) EQuilibrator - The biochemical thermodynamics
648 calculator. Nucleic Acids Res 40:770–775. https://doi.org/10.1093/nar/gkr874

649 Heffernan JK, Mahamkali V, Valgepea K, et al (2022) Analytical tools for unravelling the metabolism
650 of gas-fermenting Clostridia. Curr Opin Biotechnol 75:102700:
651 https://doi.org/10.1016/j.copbio.2022.102700

652 Heffernan JK, Valgepea K, Souza R De, et al (2020) Enhancing CO_2 -valorization using Clostridium
653 autoethanogenum for sustainable fuel and chemicals production. Front Bioeng Biotechnol
654 8:204. https://doi.org/https://doi.org/10.3389/fbioe.2020.00204

655 Kang S, Song Y, Jin S, et al (2020) Adaptive Laboratory Evolution of Eubacterium limosum ATCC 8486
656 on Carbon Monoxide. Front Microbiol 11:. https://doi.org/10.3389/fmicb.2020.00402

657 Kim J, Park S, Jeong J, et al (2021) Bioresource Technology Methanol supply speeds up synthesis gas
658 fermentation by methylotrophic-acetogenic bacterium , Eubacterium limosum KIST612.
659 Bioresour Technol 321:124521. https://doi.org/10.1016/j.biortech.2020.124521

660 Kim S, Lindner SN, Aslan S, et al (2020) Growth of E. coli on formate and methanol via the reductive
661 glycine pathway. Nat Chem Biol 16:538–545. https://doi.org/10.1038/s41589-020-0473-5

662 Köpke M, Simpson SD (2020) Pollution to products: recycling of 'above ground' carbon by gas
663 fermentation. Curr Opin Biotechnol 65:180–189. https://doi.org/10.1016/j.copbio.2020.02.017

664 Kremp F, Roth J, Müller V (2020) The Sporomusa type Nfn is a novel type of electron-bifurcating
665 transhydrogenase that links the redox pools in acetogenic bacteria. Sci Rep 10:1–14.
666 https://doi.org/10.1038/s41598-020-71038-2

667 Kroll A, Engqvist MKM, Heckmann D, Lercher MJ (2021) Deep learning allows genome-scale
668 prediction of Michaelis constants from structural features. PLoS Biol 19:1–21.
669 https://doi.org/10.1371/journal.pbio.3001402

670 Lebloas P, Lindley ND, Loubiere P (1996) Regulation of carbon and energy metabolism during the
671 linear growth phase in batch fermentations of the acetogenic methylotroph Eubacterium
672 limosum on methanol/CO2. Enzyme Microb Technol 19:187–195.
673 https://doi.org/10.1016/0141-0229(95)00230-8

674 Li F, Yuan L, Lu H, et al (2021) Deep learning based kcat prediction enables improved enzyme
675 constrained model reconstruction. bioRxiv. https://doi.org/10.1101/2021.08.06.455417

676 Li XH, Kim SK, Lee JH (2017) Anti-biofilm effects of anthranilate on a broad range of bacteria. Sci Rep
677 7:1–12. https://doi.org/10.1038/s41598-017-06540-1

678 Lindley ND, Loubiere P, Pacaud S, et al (1987) Novel Products of the Acidogenic Fermentation of
679 Methanol during Growth of Eubacterium limosum in the Presence of High Concentrations of
680 Organic Acids. Microbiology 133:3557–3563. https://doi.org/10.1099/00221287-133-12-3557

681 Ljungdahl LG (2009) A Life with Acetogens, Thermophiles, and Cellulolytic Anaerobes. Annu Rev
682 Microbiol 63:1–25. https://doi.org/10.1146/annurev.micro.091208.073617

683 Mahamkali V, Valgepea K, de Souza Pinto Lemgruber R, et al (2020) Redox controls metabolic
684 robustness in the gas-fermenting acetogen Clostridium autoethanogenum. Proc Natl Acad Sci
685 117:13168–13175. https://doi.org/10.1073/pnas.1919531117

686 Marcellin E, Behrendorff JB, Nagaraju S, et al (2016) Low carbon fuels and commodity chemicals
687 from waste gases-systematic approach to understand energy metabolism in a model acetogen.
688 Green Chem 18:3020–3028. https://doi.org/10.1039/c5gc02708j

689 Neuendorf CS, Vignolle GA, Derntl C, et al (2021) A quantitative metabolic analysis reveals
690 Acetobacterium woodii as a flexible and robust host for formate-based bioproduction. Metab
691 Eng. https://doi.org/10.1016/j.ymben.2021.09.004

692 Rabiee H, Ge L, Zhang X, et al (2021) Gas diffusion electrodes (GDEs) for electrochemical reduction of
693 carbon dioxide, carbon monoxide, and dinitrogen to value-added products: a review. Energy
694 Environ Sci 1959–2008. https://doi.org/10.1039/d0ee03756g

695 Ragsdale SW, Ljungdahl LG (1984) Purification and properties of NAD-dependent 5,10-
696 methylenetetrahydrofolate dehydrogenase from Acetobacterium woodii. J Biol Chem
697 259:3493503. https://doi.org/10.1016/s0021-9258(17)43122-x

698 Robb FT, Park JB, Adams MWW (1992) Characterization of an extremely thermostable glutamate
699 dehydrogenase: a key enzyme in the primary metabolism of the hyperthermophilic
700 archaebacterium, Pyrococcus furiosus. Biochim Biophys Acta (BBA)/Protein Struct Mol
701 1120:267–272. https://doi.org/10.1016/0167-4838(92)90247-B

702 Sly WS, Stadtman ER (1963) Formate Metabolism. J Biol Chem 238:2632–2638.
703 https://doi.org/10.1016/s0021-9258(18)67878-0

704 Song Y, Shin J, Jeong Y, et al (2017) Determination of the Genome and Primary Transcriptome of
705 Syngas Fermenting Eubacterium limosum ATCC. Sci Rep 7:1–11.
706 https://doi.org/10.1038/s41598-017-14123-3

707 Song Y, Shin J, Jin S, et al (2018) Genome-scale analysis of syngas fermenting acetogenic bacteria
708 reveals the translational regulation for its autotrophic growth. BMC Genomics 19:837.
709 https://doi.org/10.1186/s12864-018-5238-0

710 Spurgeon JM, Kumar B (2018) A comparative technoeconomic analysis of pathways for commercial
711 electrochemical CO 2 reduction to liquid products. Energy Environ Sci 11:1536–1551.
712 https://doi.org/10.1039/C8EE00097B

713 Valgepea K, Talbo G, Takemori N, et al (2022) Absolute Proteome Quantification in the Gas-
714 Fermenting Acetogen Clostridium autoethanogenum. mSystems 7:1–19.
715 https://doi.org/10.1128/msystems.00026-22

716 Vandrisse CM, Escalante-Semerena JC (2018) Small-Molecule Acetylation Controls the Degradation
717 of Benzoate and Photosynthesis in Rhodopseudomonas palustris. MBio 9:1–17.
718 https://doi.org/https://doi.org/10.1128/ mBio.01895-18

719 Wang VCC, Ragsdale SW, Armstrong FA (2013) Investigations of two bidirectional carbon monoxide
720 dehydrogenases from carboxydothermus hydrogenoformans by protein film electrochemistry.
721 ChemBioChem 14:1845–1851. https://doi.org/10.1002/cbic.201300270

722 Warwick N, Griffiths P, Keeble J, et al (2022) Atmospheric implications of increased Hydrogen use.
723 United Kingdom Department for Business, Energy & Industrial Strategy

724 Wood JC, Gonzalez-garcia RA, Daygon D, et al (2022) Molecular understanding of Eubacterium
725 limosum chemostat methanol metabolism. bioRxiv.
726 https://doi.org/https://doi.org/10.1101/2022.11.04.514945

727 Wood JC, Marcellin E, Plan MR, Virdis B (2021) High methanol-to-formate ratios induce butanol
728 production in Eubacterium limosum. Microb Biotechnol. https://doi.org/10.1111/1751-

729 7915.13963

730 Wood JC, Yuan Z, Virdis B (In press) Towards carbon neutral chemicals production: opportunities for
731 combining fermentation with electrochemical processes. Curr Opin Electrochem COELEC-D-22-
732 00174R2

733

734

Tables

Table 1 Summary of Eubacterium limosum fermentations. N = stirrer speed; BR = biological replicates; BC = biomass concentration. Methylotrophic data in the first row is provided for comparison from (Wood et al. 2022).

Condition	Gas	N	BR	D	pH	BC	Acetate	Butyrate
		rpm	#	Day^{-1}	Setpoint	gDCW/L	g/L	g/L
100mM Methanol	*85% Ar, 15% CO_2*	*400*	*2*	*0.4*	*6.8*	*0.54±0.05*	*1.60±0.08*	*0.99±0.05*
100 mM Formate	100% Ar	400	2	0.4	6.8	0.07±0.02	0.014±0.006	0.006±0.001

Table 2 Kinetic parameters of enzymes with highest, to our knowledge, BLAST similarity to E. limosum enzymes. Deep learning models have been used as a comparison, and for all remaining pathways.

Enzyme	Substrate	Cofactor	K_M (µM)	V_{max} (U/mg)	Assay k_{cat}/K_M ($M^{-1}s^{-1}$)	BLAST similarity	Deep learning prediction k_{cat}/K_M ($M^{-1}s^{-1}$)
metFV	methyl-THF						4.96E+06
MTHFD	methylene-THF	NADH	2000	1650	4.46E+05	*A woodii* (86% to elim_c0959)(Ragsdale and Ljungdahl 1984)	8.91E+04
MTHFC	methenyl-THF	NADH				*C formicaceticum* (60% to elim_c0958)(Clark and Ljungdahl 1982)	1.33E+05
FTHFS	formyl-THF	ATP	6700	58.75	3.51E+04	*Clostridium cylindrosporum* (To elim_c0957)(Buttlaire 1980)	1.32E+03
Fdh	Formate	MV	39	0.74	3.34E+04	*Shewanella oneidensis* (44% to elim_c2470)(Davies 2017)	1.65E+04
codh	CO_2	Electrode	8100 (Wang et al. 2013)		4.81E+03	*Carboxydothermus hydorgenoformas* (42% to elim_c1653) (Wang et al. 2013) *Clostridium autoethanogenum* (55% to elim_c1653)(Valgepea et al. 2022)	8.23E+04
PFOR	Acetyl-CoA	Fd^{2-}, CO_2				elim_c2885	1.47E+05
Pfl	Acetyl-CoA	Formate				elim_c0889	2.63E+05
PC	pyruvate	ATP				elim_c0058	2.89E+04
CS	Acetyl-CoA					elim_c3511	2.20E+05
ACN	Citrate					elim_c1420	2.88E+03
IDH	Isocitrate	NAD				elim_c2883	2.24E+04
GDH	2-oxoglutarate	NAD(P)H	4000	32.5	3.66E+04	*P. furiosus* (49% to elim_c2504)(Robb et al. 1992)	9.88E+04
GAD	Glutamate					elim_c2464	9.85E+03
cimA	Acetyl-CoA	Pyruvate				elim_c3964	8.26E+05
leuCD	Citramalate					elim_c3963	1.34E+05

leuCD	Citraconate					elim_c3962	1.01E+07
leuB	D-erythro-3-methylmalate	NAD				elim_c3961	1.61E+05
ilvBN	2-oxobutanoate	Pyruvate				elim_c3959/ elim_c1989	7.06E+02
ilvC	2-aceto-2-hydroxybutanoate	NAD(P)H				elim_c3965	5.22E+02
ilvD	2,3-dihydroxy-3-methylpentanoate					elim_c3960	3.24E+03
ilvE	3-methyl-2-oxopentanoate	Glutamate				elim_c0806	2.97E+03
LeuDH	isoleucine	NAD				elim_c1793	7.26E+02

Figures

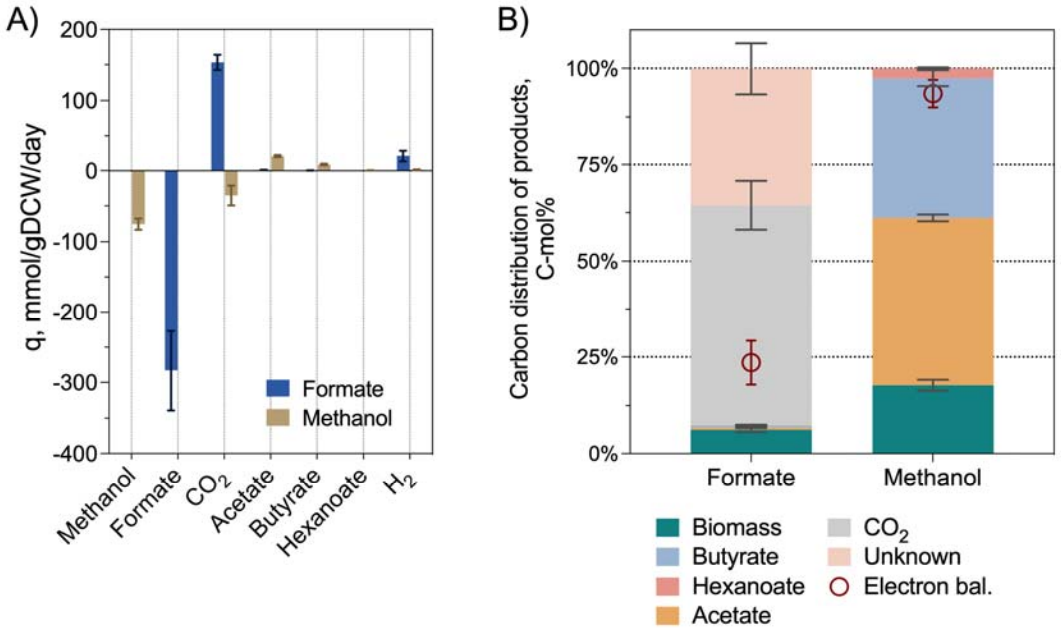
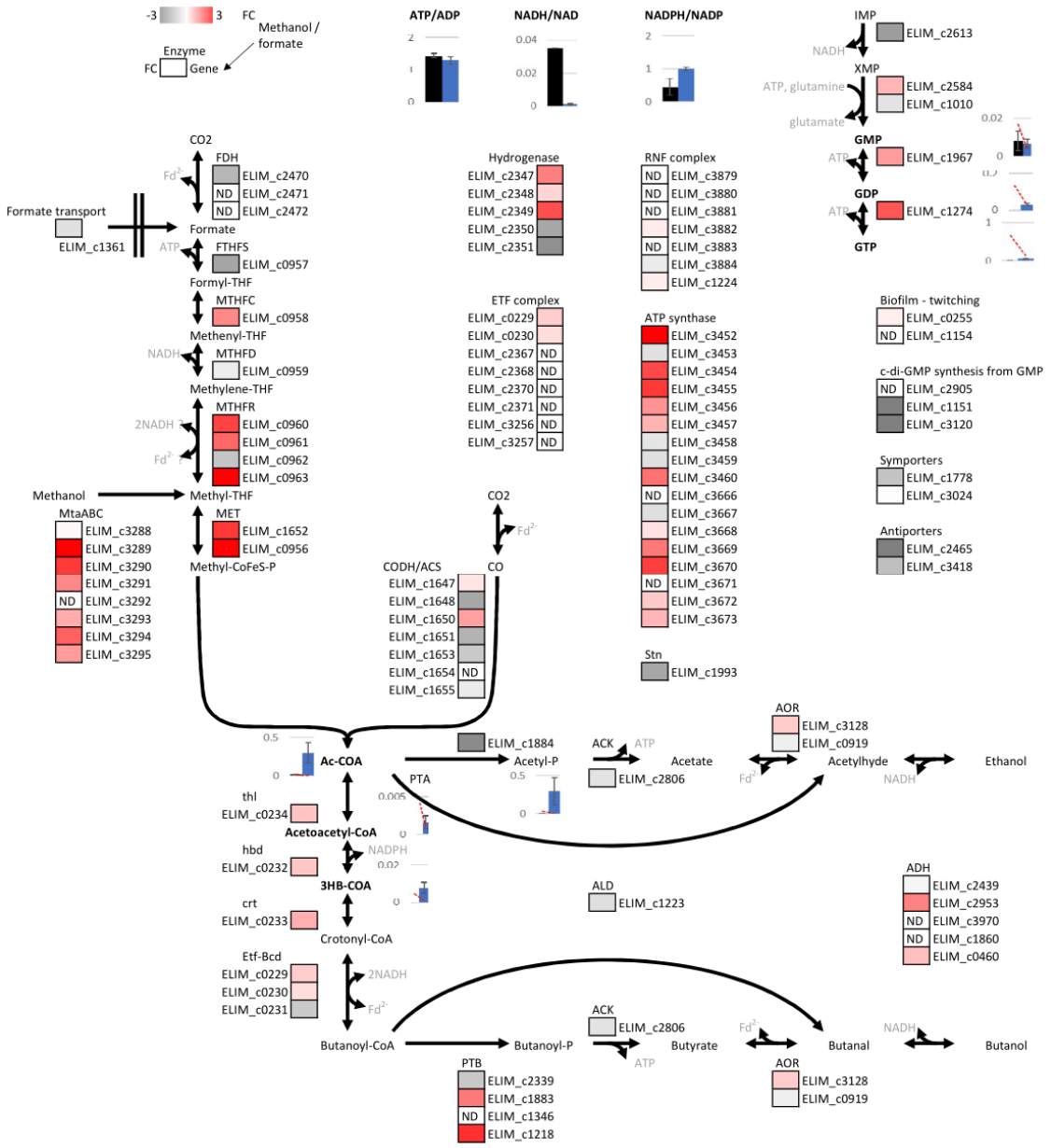


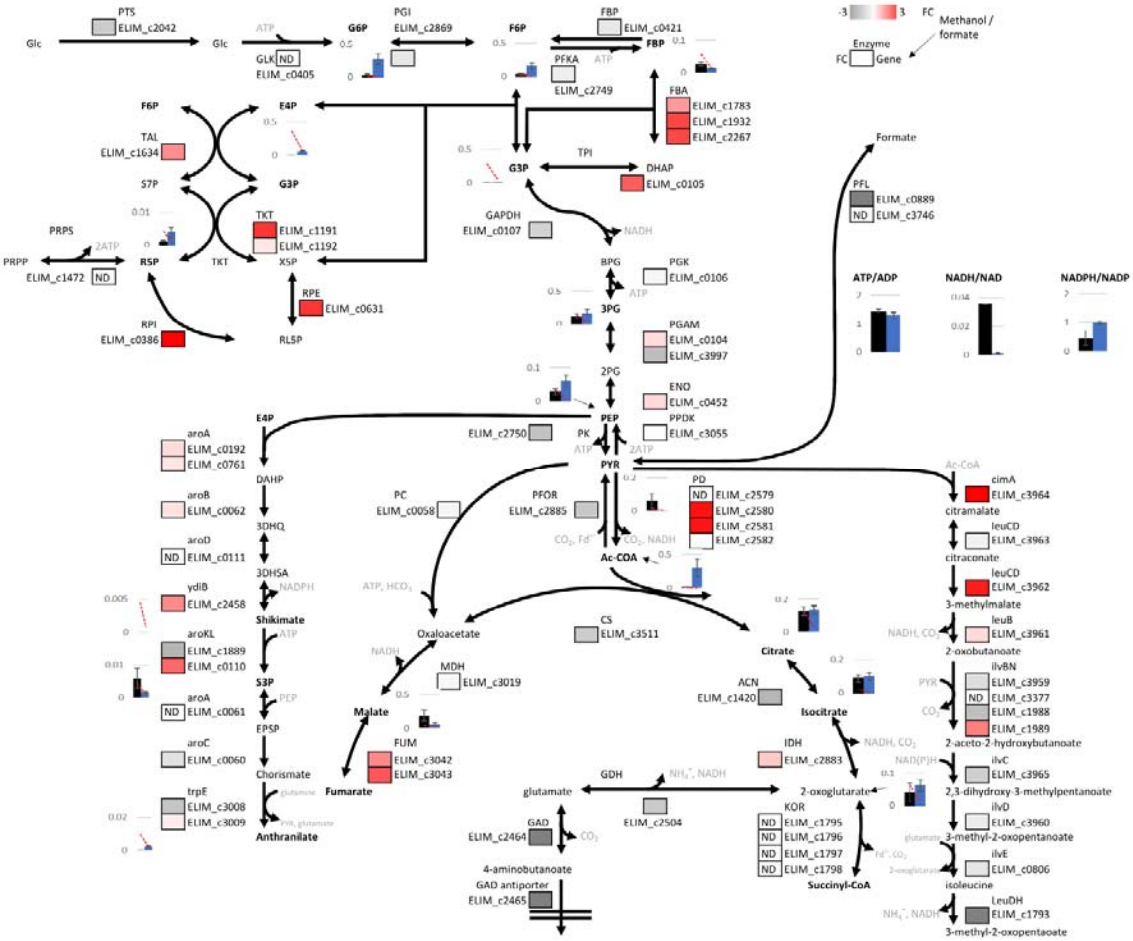
Figure 1 Characteristics of Eubacterium limosum in autotrophic chemostats, with conditions as summarised in Table 1. Methylotrophic data is provided for comparison from (Wood et al. 2022). (A) specific uptake and production rates. (B) carbon distribution of products. Values represent the average ± standard deviation between biological replicate triplicates.



748

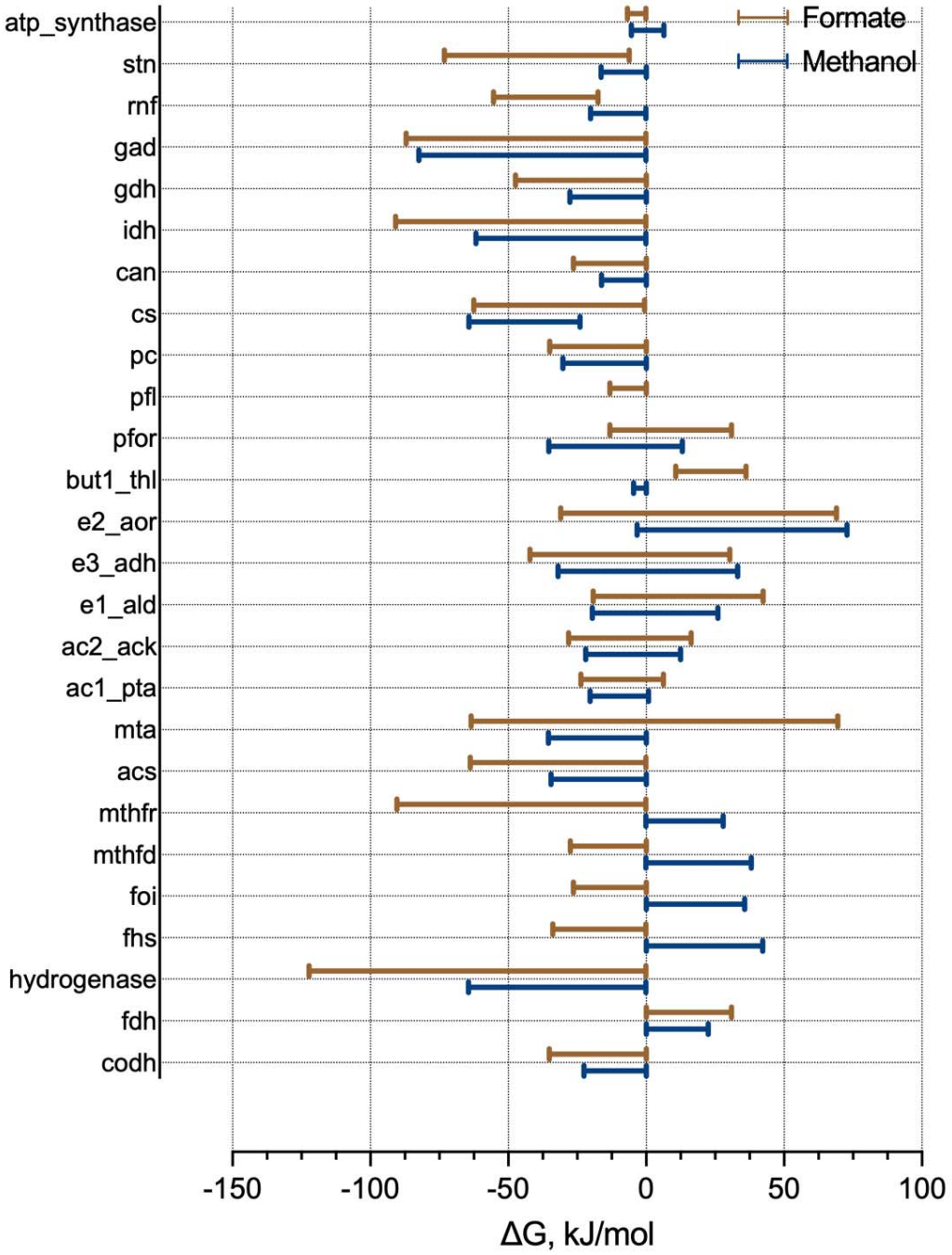
749 *Figure 2 Schematic representation of Wood Ljungdahl Pathway metabolism in E. limosum during C_1 fermentation. Reactions*
750 *are shown with substrates, products and redox mediators, however without stoichiometric balances along pathways.*
751 *Enzymes and associated protein accession numbers (elim_cxxxx) are shown for each reaction, along with log_2 fold changes*
752 *comparing methylotrophic (Wood et al. 2022) and formatotrophic growth. Intracellular metabolites concentrations*
753 *(µmol/gDCW) are shown for select metabolites as the average ± standard deviation of biological replicate triplicates. The*
754 *methanol condition is in blue, and formate condition in black. The lower limit of quantification is shown as a red dashed*
755 *line, which differs between the conditions due to the differences in cell pellet mass.*

756



757

758 *Figure 3 Schematic representation of central carbon metabolism metabolism in E. limosum during C_1 fermentation.*
759 *Reactions are shown with substrates, products and redox mediators, however without stoichiometric balances along*
760 *pathways. Enzymes and associated protein accession numbers (elim_cxxxx) are shown for each reaction, along with log_2*
761 *fold changes comparing methylotrophic (Wood et al. 2022) and formatotrophic growth. Intracellular metabolites*
762 *concentrations (µmol/gDCW) are shown for select metabolites as the average ± standard deviation of biological replicate*
763 *triplicates. The methanol condition is in blue, and formate condition in black. The lower limit of quantification is shown as a*
764 *red dashed line, which differs between the conditions due to the differences in cell pellet mass.*



765

766 *Figure 4 tMFA of E. limosum showing maximum allowable range of Gibbs free energy for respective reactions of the*
767 *methanol/CO_2 (blue) and formate (brown) conditions using thermodynamic variability analysis. Reaction directions are in*
768 *the anabolic direction from CO_2 on Figure 2, Figure 3. Rnf, PFOR, hydrogenase and stn are all calculated in the ferredoxin-*
769 *consuming direction, and ATP synthase in the ATP-forming direction. Pfl is not shown for methylotrophic growth as it was*
770 *not detected in the proteome.*