Growth of *E. coli* on formate and methanol via the reductive glycine pathway

Seohyoung Kim, Steffen N. Lindner, Selçuk Aslan, Oren Yishai, Sebastian Wenk, Karin Schann, Arren Bar-Even*

Max Planck Institute of Molecular Plant Physiology, Potsdam-Golm, Germany

* Corresponding author: Bar-Even@mpimp-golm.mpg.de

Abstract

Engineering a biotechnological microorganism for growth on one-carbon (C_1) intermediates, produced from the abiotic activation of CO_2, is a key synthetic biology step towards the valorization of this greenhouse gas to commodity chemicals. Here, we redesign the central carbon metabolism of the model bacterium *Escherichia coli* for growth on C_1 compounds using the reductive glycine pathway. Sequential genomic introduction of the four metabolic modules of the synthetic pathway resulted in a strain capable of growth on formate and CO_2 with a doubling time of ~70 hours and growth yield of ~1.5 gCDW / mol-formate. Short-term evolution decreased doubling time to less than 8 hours and improved biomass yield to 2.3 gCDW / mol-formate. Growth on methanol and CO_2 was achieved by the expression of methanol dehydrogenase in the evolved strain. Establishing synthetic formatotrophy and methylotrophy, as demonstrated here, paves the way for sustainable bioproduction rooted in CO_2 and renewable energy.

Introduction

Carbon dioxide is the focal point of many of our societal challenges and opportunities. The anthropogenic release of CO_2 threatens the balance of the planetary climate and could lead to a calamitous increase in global temperatures. On the other hand, CO_2 has the potential to replace fossil carbons as the primary feedstock for production of carbon-based value-added chemicals, including fuels, plastics, solvents, feed, and food. Yet, valorization of carbon dioxide remains an open challenge. Biological fixation of CO_2 by plants and algae takes place naturally on a massive scale. However, photosynthetic carbon fixation is challenging to harness due to multiple constraints, including competition for agricultural resources which erodes food security, land use which jeopardizes biodiversity, difficult processing of lignocellulosic biomass, and, most fundamentally, the low efficiency by which phototrophs use sunlight [1]. Alternatively, CO_2 can be upgraded by purely chemical means, e.g., generating syngas [2,3] which can be used to produce complex hydrocarbons [4]. However, such processes rely on extreme conditions and suffer from limited operational flexibility, narrow product spectrum, and low product selectivity.

An emerging solution is to integrate abiotic and biotic processes, harnessing their respective advantages while avoiding their specific drawbacks. Physicochemical methods excel in both capturing renewable energy and using it to activate CO_2 into energized small molecules. Specifically, one carbon (C_1) compounds can be derived from CO_2 and renewable energy with high efficiency [5]. Biochemical processes can then convert these C_1 compounds into a wide array of chemicals with high specificity under ambient conditions [6]. Of the possible C_1 molecules, formate and methanol are especially interesting, as, unlike gases such as carbon monoxide and methane, they are miscible in water, thus avoiding mass transfer limitations. Formate can be produced by the direct electrochemical reduction of CO_2 with an energetic efficiency of >40% [5]. Methanol can be produced in a two-step process, where electrolysis first generates hydrogen which is then reacted with CO_2; the overall energetic efficiency of this process was demonstrated to be >50% [7].

While anaerobic acetogens and methanogens can consume formate or methanol at very high efficiency, their product spectrum is very limited [8]. Aerobic cultivation, while associated with lower bioconversion efficiency, is generally much more flexible in terms of production capability. Despite considerable progress in developing better genetic tools for engineering natural aerobic formatotrophs and methylotrophs, their biotechnological application is still limited. This is in part due to unfavorable cultivation parameters (e.g., cell concentration and growth rate) and low efficiency of the relevant metabolic pathways [9]. Adapting a model biotechnological microorganism for growth on formate or methanol has therefore been a key goal of the synthetic biology community in the last decade [10-21]. However, so far, the success of these efforts has been limited. This could be partially explained by the complexity of the natural pathways – the Calvin Cycle, the Serine Cycle, and the Ribulose Monophosphate Cycle [22] – the cyclic activity of which strongly overlaps with central metabolism and requires complex regulation of the fluxes that converge into and diverge from the pathway.

Here, we use a modular-engineering approach to enable *E. coli* to grow on formate and methanol. Instead of attempting to engineer a cyclic pathway, we focus on the reductive glycine pathway, a linear route that directly assimilates formate and CO_2 into central metabolism. We divide the pathway into four modules and show how their sequential expression from the genome enables the bacterium to grow on formate. We then cultivate the engineered *E. coli* strain on formate for several generations and isolate a mutant with substantially higher growth rate and yield. We identify two genes, the overexpression of which explains the enhanced growth. Further expression of methanol dehydrogenase enables *E. coli* to metabolize methanol to formate, thus supporting growth on this C_1 carbon source. Our study represents the first case in which a synthetic pathway was successfully engineered to enable the growth of a microorganism on formate and methanol.

Results

The reductive glycine pathway

Escherichia coli, as most other key biotechnological microorganisms, cannot naturally grow on C_1 feedstocks. In this study, we aimed to design and engineer a simple, linear synthetic pathway which could support *E. coli* growth on formate or methanol as sole carbon source. Our inspiration came from the anaerobic reductive acetyl-CoA pathway (rAcCoAP) [23] which assimilates C_1 compounds very efficiently. The reductive glycine pathway (rGlyP), as shown in Figure 1, was designed to be the aerobic twin of the rAcCoAP [24]. Both are linear routes with limited overlap with central metabolism, minimizing the need for regulatory optimization. Both pathways start with the ligation of formate and tetrahydrofolate (THF), proceed via reduction into a C_1-THF intermediate, which is then condensed, within an enzyme complex, with CO_2 to generate a C_2 compound (acetyl-CoA or glycine). The C_2 compound is finally condensed with another C_1 moiety and metabolized to generate pyruvate as biomass precursor. Importantly, both the rAcCoAP and the rGlyP are characterized by a 'flat' thermodynamic profile [24,25], that is, both are mostly reversible such that the direction of the metabolic flux they carry is determined mainly by the concentrations of their substrates and products. This thermodynamic profile, while constraining the driving force of the pathway reactions [26], indicates very high energetic efficiency, where no energetic input, e.g., in the form of ATP hydrolysis, is wasted. Indeed, both pathways are associated with a very low ATP cost: only 1-2 ATP molecules are invested in the metabolism of formate to pyruvate [24]. Yet, unlike the rAcCoAP, the key enzymatic components of which are highly oxygen sensitive, the rGlyP can operate under full aerobic conditions. Hence, the rGlyP represents the most efficient theoretical route – in terms of energy utilization, resources consumption, and biomass yield – to assimilate formate in the presence of oxygen [24].

A recent study suggests that the complete rGlyP might be naturally operating in a phosphite-oxidizing microbe [27]. Moreover, the key enzymatic conversion of the rGlyP, catalyzed by the glycine cleavage system (GCS), was shown to be fully reversible in many organisms [28-30]. Previous studies demonstrated that the GCS can support glycine and serine biosynthesis from formate in an engineered *E. coli* strain at

elevated CO_2 concentration [31,32]. However, growth of the microorganism on formate (and CO_2) has not yet been demonstrated and remains an open challenge.

Modular-engineering establishes grow on formate

To facilitate the establishment of formatotrophic growth, we divided the rGlyP into four metabolic modules (Figure 1 and Supplementary Figure 1): (i) a C_1 Module (C_1M), consisting of formate-THF ligase, methenyl-THF cyclohydrolase, and methylene-THF dehydrogenase, all from *Methylobacterium extorquens* [33], together converting formate into methylene-THF; (ii) a C_2 Module (C_2M), consisting of the endogenous enzymes of the GCS (GcvT, GcvH, and GcvP) which condenses methylene-THF with CO_2 and ammonia to give glycine; (iii) a C_3 Module (C_3M), consisting of serine hydroxymethyltransferase (SHMT) and serine deaminase, together condensing glycine with another methylene-THF to generate serine and finally pyruvate; and (iv) an Energy Module (EM), which consists of formate dehydrogenase (FDH) from *Pseudomonas sp.* (strain 101) [34], generating reducing power and energy from this C_1 feedstock.

Our strategy was to establish the activities of the different modules in consecutive steps, integrating subsequent modules and selecting for their combined activity. We started with an *E. coli* strain that is auxotrophic for serine, glycine, and C_1 moieties – Δ*serA* Δ*kbl* Δ*ltaE* Δ*aceA* – where the first deletion abolishes native serine biosynthesis, the second and the third abolish threonine cleavage to glycine, and the final deletion prevents the formation of glyoxylate that could potentially be aminated to glycine [32]. The combined activity of the C_1M and the C_2M, together with the native activity of SHMT, should enable the cell to metabolize formate into C_1-THF, glycine, and serine, relieving these auxotrophies (Figure 2A).

Into the serine auxotroph strain, we introduced the enzymes of the C_1M and the C_2M, either on plasmid or in the genome (Supplementary Figure 1). For genome integration of C_1M, we combined all relevant enzymes into one operon, under the regulation of a strong constitutive promoter [35], which was inserted into a genomic 'safe spot', SS9 [36]. In the case of the C_2M, we replaced the native promoter of the GCS with a strong constitutive one (Supplementary Figure 1), increasing transcript levels 20-50 fold (Supplementary Figure 2). As expected, growth with formate was observed upon overexpression of both modules (Figure 2B) and was dependent upon high CO_2 concentration (10% in the headspace) which thermodynamically and kinetically supports the reductive activity of the GCS. While genomic integration of the enzymes of the C_1M (gC_1M) did not improve growth compared to plasmid expression (pC_1M), replacing plasmid borne expression of the enzymes of the C_2M (pC_2M) with genomic overexpression (gC_2M) supported a higher growth rate (Figure 2B).

Next, we aimed to establish formate as the primary carbon source, which requires high expression of the enzymes of the C_3M to convert glycine into the central metabolism intermediate pyruvate (Figure 2C). To enable formate assimilation to biomass, an energy source is required, which at this stage we chose to be acetate. The TCA cycle can fully oxidize acetate to generate reducing power and energy, while the deletion of isocitrate lyase (Δ*aceA*) abolishes the activity of the glyoxylate shunt, thus preventing the cell from using this molecule as a carbon source. Growth should thus be dependent on formate assimilation

via the rGlyP for biomass generation and acetate oxidation for the production of reducing power and energy (Figure 2C).

The enzymes of the C_3M were either overexpressed on a plasmid (pC_3M) or in the genome (gC_3M) (Supplementary Figure 1); in the latter case, the native *glyA* and *sdaA* were deleted and a synthetic operon harboring both genes under the regulation of a strong constitutive promoter was introduced into another genomic 'safe spot', SS7 [36]. Overexpression of the enzymes of the C_3M, within a strain that genomically expresses the enzymes of the C_1M and the C_2M, resulted on growth on formate and acetate (at 10% CO_2) (Figure 2D). Genomic expression of C_3M supported more robust growth compared to the C_3M expressed from plasmid. To confirm that the expression level of C_3M does not constrain the growth rate, we tested a strain in which the expression of *glyA* and *sdaA* is controlled by a stronger ribosome binding site (RBS-A instead of RBS-C [35]). We found that this strain to grow rather poorly (Supplementary Figure 3), indicating that higher expression of these genes is deleterious.

Finally, we aimed to introduce the EM such that formate can serve as sole carbon and energy source (Figure 2E). Overexpression of FDH on a plasmid (Supplementary Figure 1), in the strain carrying the genes of the C_1M, C_2M and C_3M in the genome, enables growth on formate (Supplementary Figure 4). However, when we introduced FDH into yet another genomic 'safe spot', SS10 [36], we failed to establish growth (Supplementary Figure 4), suggesting that the expression level of FDH was too low. We therefore tested a strain in which the genomic expression of FDH was controlled by a stronger ribosome binding site (RBS-A instead of RBS-C [35], Supplementary Figure 1). This strain, carrying no plasmid, was able to grow on formate as a sole carbon and energy source (Figure 2F and Supplementary Figure 4). This is the first case in which growth on formate was made possible in a microorganism that cannot assimilate C_1 compounds natively.

Short-term evolution improves growth on formate

To improve growth on formate we decided to conduct a short term evolution experiment in fed batch mode. We cultivated the engineered strain in test tubes, where formate was added every 3-6 days, increasing the concentration in the medium by 30 mM (red arrows in Figure 3A). Once cell turbidity reached an OD_{600} of 0.4, we diluted the cells to OD_{600} of 0.03-0.05 and started a new cycle of cultivation (Figure 3A shows six typical cycles).

Within 13 cultivation cycles (≤40 generations), growth rate on formate was substantially improved (Figure 3A), with the doubling time dropping from 65-80 h in the first two cycles to less than 10 h in the last cycle (Figure 3B). This growth rate is at least double that of a recently reported *E. coli* strain growing on formate via an engineered Calvin Cycle [37]. The short-term evolution also improved the growth yield on formate, from ≈1.5 gCDW / mol-formate in the first cycle to 2.3 ± 0.2 gCDW / mol-formate. This yield is similar to that of microorganisms growing autotrophically on formate via the Calvin cycle (3.2 ± 1.1 gCDW / mol-formate [38]). The growth of the evolved bacterium on formate was directly coupled to a decrease in the concentration of the feedstock in the medium (Figure 3C). Furthermore, as formatotrophy consumes protons (net oxidation and net assimilation both consume formic acid rather

than formate), we observed a direct correlation between cell density and the pH of the medium (Supplementary Figure 5).

To better characterize growth on formate, we conducted growth experiments in 96-well plates, automatically measuring OD_{600} every ~10 minutes. We found that maximal cell density increased monotonically with increasing formate concentration from 10 mM to 150 mM (Figure 3D). Similarly, the doubling time decreased monotonically with increasing formate concentration: from 17 hours with 10 mM formate to less than 8 hours at formate concentrations higher than 100 mM (Figure 3D). The cellular toxicity of formate, which is attributed to inhibition of cytochrome *c* oxidase [39] and dissipation of the proton motive force [40], probably explains the increased lag time at formate concentrations of 109 mM and 153 mM, and the failure to grow at higher concentrations.

Adaptive laboratory evolution usually requires hundreds of generation to improve the fitness of *E. coli* in a substantial way [41,42]. Our strain required less than 40 generations, presumably as the growth of the parent strain was so poor that a small number of mutations were sufficient to drastically improve fitness. To check whether this is indeed the case, we isolated multiple colonies of the evolved strain and sequenced their genomes. We found two mutations which occurred in all sequenced colonies (Supplementary Figure 6). The first was a single base-pair substitution in the 5'-UTR of the newly introduced FDH gene, which increased the level of transcript 2.5-fold (Supplementary Figure 7) and resulted in a 7.4-fold increase in formate oxidation activity in cell extract assays (Supplementary Figure 8). The second mutation was a single base-pair substitution in the promoter region of *pntAB*, which encodes for the membrane-bound transhydrogenase. This mutation increased transcript level by more than 13-fold (Supplementary Figure 7). The beneficial effect of these two mutations is to be expected, as the first increases energy supply to the cell from formate and the second increases the availability of NAPDH, a key cofactor for the activity of the rGlyP (consumed by methylene-THF dehydrogenase), the supply of which could limit pathway activity.

To confirm that the two mutations suffice to support the improved growth on formate, we used Multiplex Automated Genomic Engineering (MAGE [43]) to introduce these mutations into a non-evolved strain. We found that while the parent strain could hardly grow in 96-well plates, the strain in which the two mutations were present displayed a growth profile almost identical to that of the evolved strain (Supplementary Figure 9). We therefore conclude that overexpression of FDH and PntAB were sufficient to enable the observed improved growth on formate. By further optimizing cultivation conditions, we found that addition of 100 mM sodium bicarbonate to the medium enabled the evolved strain, as well as the reconstructed strain, to grow at higher formate concentrations, tolerating even 300 mM (Supplementary Figure 10). The increased tolerance to formate might be attributed to a higher buffer capacity of the medium containing bicarbonate, possibly decreasing fluctuations in local pH due to formate consumption.

Carbon labeling sheds light on cellular fluxes

To confirm that growth on formate indeed proceeds via the rGlyP, we performed carbon labeling experiments. We fed the cultures with ^{13}C-formate/$^{12}CO_2$, ^{12}C-formate/$^{13}CO_2$, and ^{13}C-formate/$^{13}CO_2$, and measured the labeling pattern of proteinogenic amino-acids using liquid chromatography–mass spectrometry. We focused on 7 amino-acids – glycine, serine, alanine, valine, proline, threonine, and histidine – which either directly relate to the activity of the rGlyP or originate from different parts of central metabolism, thus providing an indication of key metabolic fluxes.

As shown in Figure 4, the amino acid labeling confirms the activity of the rGlyP. Specifically, feeding ^{13}C-formate/$^{12}CO_2$ resulted in single labeled glycine and double labeled serine and pyruvate (as indicated by the labeling of alanine). As valine – derived from two pyruvate molecules, one of which loses its carboxylic acid carbon – is mostly quadruple labeled, we deduce that pyruvate is labeled in its two non-carboxylic carbons, as predicted for growth via the rGlyP (Supplementary Figure 11). Conversely, feeding ^{12}C-formate/$^{13}CO_2$ resulted, as expected, in single labeled glycine, serine and pyruvate. As valine is also single labeled, we deduce that pyruvate is labeled in its carboxylic carbon, again confirming the activity of the rGlyP (Supplementary Figure 11). Upon feeding ^{13}C-formate/$^{13}CO_2$, all amino-acids were nearly-completely labeled, where the overall fraction of labeled carbon (marked above the bars in Figure 4 in italics) is 97-98%, as expected by feeding with 99% ^{13}C-labeled formate and 99% ^{13}C-labeled CO_2.

The labeling of threonine (derived from oxaloacetate) and proline (derived from 2-ketoglutarate) sheds light on the flux via the anaplerotic reactions and the TCA cycle. Specifically, if cyclic flux via the TCA cycle would predominate over anaplerotic flux, threonine and proline would be expected to be almost fully labeled upon feeding with ^{13}C-formate and almost fully unlabeled when feeding with $^{13}CO_2$ (Supplementary Figure 11). Conversely, if anaplerotic flux and non-cyclic flux would predominate over the cyclic flux, then threonine would be expected to be mostly double labeled on either ^{13}C-formate or $^{13}CO_2$ and proline would be expected to be mostly quadruple labeled on ^{13}C-formate and single labeled on $^{13}CO_2$ (Supplementary Figure 11). The results shown in Figure 4 are thus consistent with high anaplerotic flux and low cyclic flux. This indicates that the cell obtains sufficient reducing power and energy from formate oxidation via FDH, and hence does not need to wastefully oxidize the assimilated carbons within pyruvate and acetyl-CoA (i.e., investing cellular resources for C_1 assimilation, only to completely oxidize the assimilated product).

Engineered growth of *E.coli* on methanol

Next, we aimed to use the rGlyP for methanol assimilation. A single enzyme, methanol dehydrogenase (MDH), can convert methanol to formaldehyde, which can be oxidized to formate by the endogenous glutathione system [44] (Figure 5A). The expression of MDH can thus be regarded as the introduction of another module – a Methanol Module (MM) – that serves to metabolize methanol to formate, while providing the cells with reducing power (Figure 5B). We tested NAD-dependent MDH from several organisms: *Bacillus stearothermophilus* (*Bs*MDH) [19], *Corynebacterium glutamicum* (*Cg*MDH) [45], and *Cupriavidus necator* N-1 (*Cn*MDH, WT *mdh2*) [46], as well as two MDHs from *Bacillus methanolicus* (*Bm*MDH2 and *Bm*MDH3) [10,47] and an improved variant (*Bm*MDH2*, carrying Q5L A363L modifications)

[47]. These MDH variants were expressed on plasmids in three genetic backgrounds: the parent strain (gC_1M gC_2M gC_3M gEM), the evolved strain, and the parent strain to which the mutation within the promoter of the *pntAB* was introduced via MAGE. Overexpression of *Bs*MDH supported growth on 600 mM methanol, which was most efficient in the latter strain (Figure 5C) and somewhat poorer in the other strains (Figure 5D). The other MDH variants failed to support growth (Figure 5D, final OD_{600} not higher than inoculation, as indicated by the brown dashed line).

To confirm that growth on methanol indeed depends on formaldehyde oxidation via the glutathione system, we deleted the endogenous genes encoding for S-(hydroxymethyl)glutathione dehydrogenase (Δ*frmA*) in the above strains. We found the deletion to completely abolish growth on methanol (Figure 5D), confirming the essentiality of the glutathione system to the observed growth. Moreover, overexpression of NAD-dependent formaldehyde dehydrogenase from *Pseudomonas putida* (PpFADH), as demonstrated in a previous study [12], or from *Pseudomonas aeruginosa* (PaFADH [48]) did not improve growth on methanol (Figure 5D), indicating that the endogenous glutathione system is sufficiently fast and that the rate limiting step lies in methanol oxidation.

To confirm that growth on methanol indeed proceed via the rGlyP, we performed a carbon labeling experiment. We fed the cultures with ^{13}C-methanol/$^{12}CO_2$ and measured the labeling pattern of the proteinogenic amino-acids described above. The labeling pattern we measured (Figure 5E) was essentially identical to that observed with ^{13}C-formate/$^{12}CO_2$ (Figure 4), confirming that growth on methanol takes place via the synthetic route.

Notably, the growth rate on methanol was considerably lower than that on formate – doubling time of 54 ± 5.5 h. This can be attributed to the slow rate of methanol oxidation. The observed biomass yield was 4.2 ± 0.17 gCDW / mole methanol, considerably lower than that of microorganisms naturally growing on methanol (7.2 ± 1.2 gCDW / mol-methanol via the Calvin cycle, 12 ± 1.6 gCDW / mol-methanol via the serine cycle, and 15.6 ± 2.7 gCDW / mol-methanol via the Ribulose Monophosphate Cycle [38]). We speculate that the low yield is also related to the slow rate of methanol oxidation: a low growth rate increases the proportional consumption of energy for cell maintenance, thus lowering biomass yield. Addition of 100 mM sodium bicarbonate significantly increased the final OD_{600}, but the growth parameters did not improve: doubling time of 55 ± 1 h and biomass yield of 4.2 ± 0.1 gCDW / mol-methanol (Supplementary Figure 12, also showing methanol consumption during growth).

Discussion

This study provides the first demonstration of synthetic formatotrophy and methylotrophy. We show that rational design alone can suffice to achieve such a goal, but that short-term evolution can provide useful fine-tuning to improve growth characteristics. Further improvement of growth on formate and methanol can be achieved via long term evolution or via the introduction of metabolic routes that bypass limiting reactions. For example, replacing NAD-dependent MDH with methanol oxidase might reduce biomass yield (as this enzyme dissipates reducing power) but could support a much higher growth rate, as it replaces a thermodynamically- and kinetically-limited reaction with a favorable and fast one.

We recently used computational analysis to compare different C_1 assimilation pathways according to the biomass and product yields they are expected to support on formate and methanol [49]. For formate assimilation, we found that the rGlyP has the potential to outperform its natural and synthetic counterparts in terms of both biomass and product yields. With regards to methanol assimilation, the Ribulose Monophosphate Cycle supports the highest biomass yield. However, this pathway is outperformed by the rGlyP for the production of the key metabolic precursors acetyl-CoA and pyruvate. This is attributed to the overflow of reducing power in the Ribulose Monophosphate Cycle, while the rGlyP pathway uses CO_2 as an electron sink [49]. Overall, the rGlyP seems to be the most flexible C_1 assimilation pathway, with the potential to support the highest yields of acetyl-CoA and pyruvate using either formate or methanol as feedstocks [49]. However, reaching the full potential of the rGlyP would require considerable growth optimization via rational design and adaptive laboratory evolution.

The C_1 assimilating strains can be further engineered for the production of value-added chemicals. Especially interesting are chemicals that can be derived directly from the rGlyP intermediates or product, and can thus be produced with high yield and productivity. For example, lactate and isobutanol, both of which are derived from pyruvate, should be produced with high yield. Similarly, cysteine, which is derived from serine, a key pathway intermediate, might be an ideal product. Coupling the abiotic synthesis of formate and methanol with their microbial conversion to chemicals of interest will enable an integrated process for the valorization of CO_2 into renewable commodities.

Acknowledgements

The authors thank Charlie Cotton, Nico Claassens, Hai He, Ron Milo, Elad Noor, Niv Antonovsky, Avi Flamholz, Yinon Bar-On, William Newell, Tobi Erb, and Madeleine Bouzon for critical reading of the manuscript and helpful suggestions. This work was funded by the Max Planck Society, by the German Ministry of Education and Research grant FormatPlant (part of BioEconomy 2030, Plant Breeding Research for the Bioeconomy), and by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 763911 (Project eForFuel).

Author contributions

A.B.-E. designed and supervised the research and wrote the paper;
S.K., S.N.L, S.A., and O.Y. genetically engineered *E. coli* for growth on formate and methanol, and performed the growth experiments;
S.K. and S.N.L measured biomass yield on formate and methanol.
S.A. performed the qPCR experiments;
S.W., and K.S. cloned the methanol dehydrogenase and formaldehyde dehydrogenase genes, and assisted in the growth experiments on methanol;
S.K., S.N.L., S.A., O.Y., S.W., K.S., and A.B.-E. analyzed the data;

Competing interests statement

A.B.-E. is cofounder of b.fab, exploring the commercialization of microbial bioproduction using formate as feedstock. The company was not involved in any way in performing or funding this study.

ORCID

Seohyoung Kim: 0000-0002-5873-5451

Steffen Lindner: 0000-0003-3226-3043

Selçuk Aslan: 0000-0003-4378-1575

Oren Yishai: 0000-0001-8712-3959

Sebastian Wenk: 0000-0001-9404-3535

Karin Schann: 0000-0003-1548-7643

Arren Bar-Even: 0000-0002-1039-4328

References

1 Blankenship, R. E. *et al.* Comparing photosynthetic and photovoltaic efficiencies and recognizing the potential for improvement. *Science* **332**, 805-809, doi:10.1126/science.1200165 (2011).

2 Scheffe, J. R. & Steinfeld, A. Oxygen exchange materials for solar thermochemical splitting of H2O and CO2: a review. *Materials Today* **17**, 341-348 (2014).

3 Snoeckx, R. & Bogaerts, A. Plasma technology - a novel solution for CO2 conversion? *Chem Soc Rev* **46**, 5805-5863, doi:10.1039/c6cs00066e (2017).

4 Zhang, Q., Kang, J. & Wang, Y. Development of novel catalysts for Fischer–Tropsch synthesis: tuning the product selectivity. *ChemCatChem* **2**, 1030-1058 (2010).

5 Jouny, M., Luc, W. & Jiao, F. General techno-economic analysis of CO2 electrolysis systems. *Ind Eng Chem Res* **57**, 2165-2177 (2018).

6 Yishai, O., Lindner, S. N., Gonzalez de la Cruz, J., Tenenboim, H. & Bar-Even, A. The formate bio-economy. *Curr Opin Chem Biol* **35**, 1-9, doi:10.1016/j.cbpa.2016.07.005 (2016).

7 Szima, S. & Cormos, C. C. Improving methanol synthesis from carbon-free H2 and captured CO2: A techno-economic and environmental evaluation. *J CO2 Util* **24**, 555-563 (2018).

8 Bertsch, J. & Muller, V. Bioenergetic constraints for conversion of syngas to biofuels in acetogenic bacteria. *Biotechnology for biofuels* **8**, 210, doi:10.1186/s13068-015-0393-x (2015).

9 Bennett, R. K., Steinberg, L. M., Chen, W. & Papoutsakis, E. T. Engineering the bioconversion of methane and methanol to fuels and chemicals in native and synthetic methylotrophs. *Curr Opin Biotechnol* **50**, 81-93, doi:10.1016/j.copbio.2017.11.010 (2017).

10 Muller, J. E. *et al.* Engineering Escherichia coli for methanol conversion. *Metab Eng* **28**, 190-201, doi:10.1016/j.ymben.2014.12.008 (2015).

11 Dai, Z. *et al.* Metabolic construction strategies for direct methanol utilization in Saccharomyces cerevisiae. *Bioresour Technol* **245**, 1407-1412, doi:10.1016/j.biortech.2017.05.100 (2017).

12 Yu, H. & Liao, J. C. A modified serine cycle in Escherichia coli coverts methanol and CO2 to two-carbon compounds. *Nature communications* **9**, 3992, doi:10.1038/s41467-018-06496-4 (2018).

13 Meyer, F. *et al.* Methanol-essential growth of Escherichia coli. *Nature communications* **9**, 1508, doi:10.1038/s41467-018-03937-y (2018).

14 Woolston, B. M., King, J. R., Reiter, M., Van Hove, B. & Stephanopoulos, G. Improving formaldehyde consumption drives methanol assimilation in engineered E. coli. *Nature communications* **9**, 2387, doi:10.1038/s41467-018-04795-4 (2018).

15 Bennett, R. K., Gonzalez, J. E., Whitaker, W. B., Antoniewicz, M. R. & Papoutsakis, E. T. Expression of heterologous non-oxidative pentose phosphate pathway from Bacillus methanolicus and phosphoglucose isomerase deletion improves methanol assimilation and metabolite production by a synthetic Escherichia coli methylotroph. *Metab Eng* **45**, 75-85, doi:10.1016/j.ymben.2017.11.016 (2017).

16 Gonzalez, J., Bennett, R. K., Papoutsakis, E. T. & Antoniewicz, M. R. Methanol assimilation in Escherichia coli is improved by co-utilization of threonine and deletion of leucine-responsive regulatory protein. *Metab Eng* **45**, 67-74, doi:10.1016/j.ymben.2017.11.015 (2017).
17 Rohlhill, J., Sandoval, N. R. & Papoutsakis, E. T. Sort-Seq Approach to Engineering a Formaldehyde-Inducible Promoter for Dynamically Regulated Escherichia coli Growth on Methanol. *ACS synthetic biology* **6**, 1584-1595, doi:10.1021/acssynbio.7b00114 (2017).
18 Woolston, B. M., Roth, T., Kohale, I., Liu, D. R. & Stephanopoulos, G. Development of a formaldehyde biosensor with application to synthetic methylotrophy. *Biotechnol Bioeng* **115**, 206-215, doi:10.1002/bit.26455 (2018).
19 Whitaker, W. B. *et al.* Engineering the biological conversion of methanol to specialty chemicals in Escherichia coli. *Metab Eng* **39**, 49-59, doi:10.1016/j.ymben.2016.10.015 (2017).
20 Lu, X. *et al.* Constructing a synthetic pathway for acetyl-coenzyme A from one-carbon through enzyme design. *Nature communications* **10**, 1378, doi:10.1038/s41467-019-09095-z (2019).
21 Wang, X. *et al.* Biological conversion of methanol by evolved Escherichia coli carrying a linear methanol assimilation pathway. *Bioresour Bioprocess* **4**, 41-46 (2017).
22 Anthony, C. *The Biochemistry of Methylotrophs*. (Academic Press, 1982).
23 Drake, H. L., Kirsten, K. & Matthies, C. in *The Prokaryotes* 354-420 (Springer New York, 2006).
24 Bar-Even, A., Noor, E., Flamholz, A. & Milo, R. Design and analysis of metabolic pathways supporting formatotrophic growth for electricity-dependent cultivation of microbes. *Biochim Biophys Acta* **1827**, 1039-1047 (2013).
25 Bar-Even, A. Does acetogenesis really require especially low reduction potential? *Biochim Biophys Acta* **1827**, 395-400, doi:10.1016/j.bbabio.2012.10.007 (2013).
26 Noor, E. *et al.* Pathway thermodynamics highlights kinetic obstacles in central metabolism. *PLoS Comput Biol* **10**, e1003483, doi:10.1371/journal.pcbi.1003483 (2014).
27 Figueroa, I. A. *et al.* Metagenomics-guided analysis of microbial chemolithoautotrophic phosphite oxidation yields evidence of a seventh natural CO2 fixation pathway. *Proc Natl Acad Sci U S A* **115**, E92-E101, doi:10.1073/pnas.1715549114 (2018).
28 Kawasaki, H., Sato, T. & Kikuchi, G. A new reaction for glycine biosynthesis. *Biochem Biophys Res Commun* **23**, 227-233 (1966).
29 Motokawa, Y. & Kikuchi, G. Glycine metabolism by rat liver mitochondria. Reconstruction of the reversible glycine cleavage system with partially purified protein components. *Arch Biochem Biophys* **164**, 624-633 (1974).
30 Pasternack, L. B., Laude, D. A., Jr. & Appling, D. R. 13C NMR detection of folate-mediated serine and glycine synthesis in vivo in *Saccharomyces cerevisiae*. *Biochemistry* **31**, 8713-8719 (1992).
31 Tashiro, Y., Hirano, S., Matson, M. M., Atsumi, S. & Kondo, A. Electrical-biological hybrid system for CO2 reduction. *Metab Eng* **47**, 211-218, doi:10.1016/j.ymben.2018.03.015 (2018).
32 Yishai, O., Bouzon, M., Doring, V. & Bar-Even, A. In Vivo Assimilation of One-Carbon via a Synthetic Reductive Glycine Pathway in Escherichia coli. *ACS synthetic biology* **7**, 2023-2028, doi:10.1021/acssynbio.8b00131 (2018).
33 Crowther, G. J., Kosaly, G. & Lidstrom, M. E. Formate as the main branch point for methylotrophic metabolism in *Methylobacterium extorquens* AM1. *J Bacteriol* **190**, 5057-5062 (2008).
34 Tishkov, V. I. & Popov, V. O. Catalytic mechanism and application of formate dehydrogenase. *Biochemistry (Mosc)* **69**, 1252-1267, doi:BCM69111537 [pii] (2004).
35 Wenk, S., Yishai, O., Lindner, S. N. & Bar-Even, A. An Engineering Approach for Rewiring Microbial Metabolism. *Methods Enzymol* **608**, 329-367, doi:10.1016/bs.mie.2018.04.026 (2018).
36 Bassalo, M. C. *et al.* Rapid and Efficient One-Step Metabolic Pathway Integration in E. coli. *ACS synthetic biology* **5**, 561-568, doi:10.1021/acssynbio.5b00187 (2016).
37 Gleizer, S. *et al.* Conversion of Escherichia coli to Generate All Biomass Carbon from CO2. *Cell* **179**, 1255-1263 e1212, doi:10.1016/j.cell.2019.11.009 (2019).
38 Claassens, N. J., Cotton, C. A., Kopljar, D. & Bar-Even, A. Making quantitative sense of electromicrobial production. *Nature Catalysis* **2**, 437 (2019).
39 Nicholls, P. Formate as an inhibitor of cytochrome c oxidase. *Biochem Biophys Res Commun* **67**, 610-616 (1975).
40 Warnecke, T. & Gill, R. T. Organic acid toxicity, tolerance, and production in Escherichia coli biorefining applications. *Microb Cell Fact* **4**, 25, doi:10.1186/1475-2859-4-25 (2005).
41 Dragosits, M. & Mattanovich, D. Adaptive laboratory evolution -- principles and applications for biotechnology. *Microb Cell Fact* **12**, 64, doi:10.1186/1475-2859-12-64 (2013).

42 Wytock, T. P. *et al.* Experimental evolution of diverse Escherichia coli metabolic mutants identifies genetic loci for convergent adaptation of growth rate. *PLoS Genet* **14**, e1007284, doi:10.1371/journal.pgen.1007284 (2018).
43 Wang, H. H. *et al.* Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894-898, doi:nature08187 [pii] 10.1038/nature08187 (2009).
44 Gutheil, W. G., Kasimoglu, E. & Nicholson, P. C. Induction of glutathione-dependent formaldehyde dehydrogenase activity in Escherichia coli and Hemophilus influenza. *Biochem Biophys Res Commun* **238**, 693-696 (1997).
45 Kotrbova-Kozak, A., Kotrba, P., Inui, M., Sajdok, J. & Yukawa, H. Transcriptionally regulated adhA gene encodes alcohol dehydrogenase required for ethanol and n-propanol utilization in Corynebacterium glutamicum R. *Appl Microbiol Biotechnol* **76**, 1347-1356, doi:10.1007/s00253-007-1094-6 (2007).
46 Wu, T. Y. *et al.* Characterization and evolution of an activator-independent methanol dehydrogenase from Cupriavidus necator N-1. *Appl Microbiol Biotechnol* **100**, 4969-4983, doi:10.1007/s00253-016-7320-3 (2016).
47 Roth, T. B., Woolston, B. M., Stephanopoulos, G. & Liu, D. R. Phage-Assisted Evolution of Bacillus methanolicus Methanol Dehydrogenase 2. *ACS synthetic biology* **8**, 796-806, doi:10.1021/acssynbio.8b00481 (2019).
48 Zhang, W. *et al.* Expression, purification, and characterization of formaldehyde dehydrogenase from Pseudomonas aeruginosa. *Protein Expr Purif* **92**, 208-213, doi:10.1016/j.pep.2013.09.017 (2013).
49 Cotton, C. A., Claassens, N. J., Benito-Vaquerizo, S. & Bar-Even, A. Renewable methanol and formate as microbial feedstocks. *Curr Opin Biotechnol* **62**, 168-180 (2020).

Figure legends

Figure 1

The synthetic reductive glycine pathway is similar in structure to the reductive acetyl-CoA pathway. Yet, while the latter pathway is restricted to anaerobic conditions, the former can operate under aerobic conditions. Both pathways are highly ATP-efficient, as only 1-2 ATP molecules are consumed in the conversion of formate to pyruvate (e.g., instead of 7 by the Calvin Cycle). Molecular structure in brown corresponds to a sub-structure of tetrahydrofolate. Enzymes of the reductive glycine pathway, as implemented in this study, are indicated in purple (Lpd, unlike the other enzymes of the glycine cleavage system, was not overexpressed). 'Me' corresponds to *Methylobacterium extorquens* and 'Ec' corresponds to *Escherichia coli*. Division of the pathway into modules, as explained in the text, is shown in light brown to the right of the figure.

Figure 2

Modular establishment of the reductive glycine pathway. (A) Selection scheme of C_1M and C_2M for the biosynthesis of C_1-moieties, glycine, and serine. (B) Overexpression of C_1M and C_2M enabled growth with formate (and CO_2) as sole source of C_1-moieties, glycine, and serine. (C) Selection scheme of C_1M, C_2M, and C_3M to generate biomass building blocks, where acetate oxidation provides reducing power and energy. Deletion of *aceA* prevents acetate from being used as a carbon source. (D) Overexpression of C_1M, C_2M, and C_3M enabled growth with formate as source of biomass and acetate as an energy source. Genomic integration of C_3M was performed in strain in which the endogenous *glyA* and *sdaA* were deleted. (E) Selection scheme of C_1M, C_2M, C_3M, and EM to use formate as sole carbon and energy source. (F) Growth on formate is demonstrated only when all four modules are overexpressed.

Genomic overexpression is indicated by 'g', while overexpression from a plasmid is indicated by 'p'. Experiments were conducted at 10% CO_2 within 96-well plates and were performed in triplicates, which displayed identical growth curves (±5%), and hence were averaged. Doubling times (DT) shown in the figure. All experiments (in triplicates) were repeated five times, which showed highly similar growth behavior.

Figure 3

Short term evolution improves growth on formate. (A) Test-tube cultivation on formate as sole carbon source. The vertical small red arrows correspond to the addition of formate, increasing the concentration in the medium by 30 mM. Upon reaching an OD_{600} of 0.4, cells were reinoculated into a new test-tube with an initial OD_{600} of 0.03-0.05. Dots corresponds to the measured OD_{600} values in duplicate experiments (using biologically independent strains) and the line to their average. Six exemplifying cycles of cultivation are shown. (B) Doubling time decreased with cultivation cycle. Dots corresponds to the doubling time in duplicate experiments and the line to their average. (C) Growth of the evolved strain (in test-tube) is directly coupled to a decrease in formate concentration. Dots corresponds to measured values in duplicate experiments and the lines to their averages. (D) Cultivation of the evolved strain on formate as a sole carbon source within a 96-well plate. Experiments were conducted at 10% CO_2. Plate reader experiments were performed in triplicates, which displayed identical growth curves (±5%), and hence were averaged. The experiment (in triplicates) was repeated three times, which showed highly similar growth behavior. Doubling times (DT) are shown in the figure. DT were considerably shorter in the plate reader than in test-tube as the measurements in were more accurate (taken every 10 minutes rather than once per day) and since the conditions are different (e.g., more stable cultivation environment in the plate reader).

Figure 4

Labeling pattern of proteinogenic amino acids confirms the activity of the reductive glycine pathway. As elaborated in Supplementary Figure 11, the labeling pattern is consistent with the assimilation of formate and CO_2 via the synthetic pathway, and indicates low cyclic flux via the TCA cycle. Numbers written in italics above the bars correspond to the overall fraction of labeled carbons.

Figure 5

Engineered growth on methanol. (A) Methanol can be assimilated via the activity of methanol dehydrogenase (MDH), where formaldehyde is oxidized to formate via the native activity of the glutathione system. (B) The Methanol Module (MM) converts methanol to formate and provides the cell with reducing power and energy. (C) Overexpression of MDH from *Bacillus stearothermophilus* (*Bs*MDH) within the gC_1M gC_2M gC_3M gEM strain, carrying a mutation in the promoter of the *pntAB* operon (Supplementary Figure 6), enabled growth on methanol within test-tubes. Experiments were conducted at 10% CO_2. Dots corresponds to the doubling times in triplicate experiments (using biologically independent strains) and the line to their average. (D) Comparison of growth on methanol (shown are final cell densities) with different expressed enzymes and at different genetic backgrounds. We tested NAD-dependent MDH from several organisms: *Bacillus stearothermophilus* (*Bs*MDH), *Corynebacterium*

glutamicum (*Cg*MDH), and *Cupriavidus necator* N-1 (*Cn*MDH), as well as two MDHs from *B. methanolicus* (*Bm*MDH2 and *Bm*MDH3) and an improved variant (*Bm*MDH2*, carrying Q5L A363L modifications). We further tested formaldehyde dehydrogenases from *Pseudomonas putida* (PpFADH) and *Pseudomonas aeruginosa* (PaFADH). Experiments were conducted in (biologically independent) duplicates; dots show the measured OD_{600} values and bars corresponds to the average. (E) Labeling pattern of proteinogenic amino acids upon feeding with ^{13}C-methanol/^{12}-CO_2 is identical that with ^{13}C-formate/^{12}-CO_2 (Figure 4), confirming the activity of the reductive glycine pathway. Numbers written in italics above the bars correspond to the overall fraction of labeled carbons.

Online Methods

Chemicals and reagents

Primers were synthesized by Integrated DNA Technologies (IDT, Leuven, Belgium). PCR reactions were carried out either using Phusion High-Fidelity DNA Polymerase or Dream Taq. Restrictions and ligations were performed using FastDigest enzymes and T4 DNA ligase, respectively, all purchased from Thermo Fisher Scientific (Dreieich, Germany). Glycine, sodium formate, sodium formate-^{13}C, methanol-^{13}C were ordered from Sigma-Aldrich (Steinheim, Germany). $^{13}CO_2$ was ordered from Cambridge Isotope Laboratories, Inc. (Andover, MA, USA).

Bacterial Strains

Wild type *Escherichia coli* strain MG1655 (F^- λ^- *ilv*G^- *rfb*-50 *rph*-1) was used as the host for all genetic modifications. *E. coli* strain DH5α (F^-, λ^-,Φ80*lac*ZΔM15, Δ(*lac*ZYA-*arg*F)U169, *deo*R, *rec*A1, *end*A1, *hsd*R17(rK^- mK^+), *pho*A, *sup*E44, *thi*-1, *gyr*A96, *rel*A1) and *E. coli* strain ST18 (*pro thi hsdR*$^+$ Tp^r Sm^r; chromosome::RP4-2 Tc::Mu-Kan::Tn7λpir Δ*hemA*) [50] were used for cloning and conjugation procedures, respectively.

Genome engineering

Gene knockouts were introduced in MG1655 by P1 phage transduction [51]. Single gene knockout mutants from the National BioResource Project (NIG, Japan) [52] were used as donors of specific mutations. For the recycling of selection marker (as the multiple gene deletions and integrations were required) all the antibiotic cassettes integrated into genome were flanked by FRT (Flippase Recongnition Target) sites. Cells were transformed with a flippase recombinase helper plasmid (FLPe, replicating at 30°C, Gene Bridges), which carries a gene encoding FLP which recombines at the FRT sites and removes the antibiotic cassette. Elevated temperature (37°C) was subsequently used to cure the cell from the FLPe plasmid.

Exchange of *E. coli* native promoter with a synthetic one was performed by using PCR-mediated λ-Red recombination method. The synthetic promoter fused with FRT-flanked kanamycin resistance gene was cloned into the pZ vector and the DNA fragment was obtained by PCR amplification with primers containing 50 base pair homology for recombination. Recombinant *E. coli* MG1655 harboring λ-Red recombinase (pRed/ET, Gene Bridges) was cultivated at 30 °C, and the expression of λ-Red recombinase was induced by the addition of 10 mM L-arabinose. Electro-competent cells were prepared by washing three times with ddH_2O. The PCR product was introduced into *E. coli* expressing the λ-Red recombinase via electroporation. Mutants with exchanged promoter occurred via homologous recombination, selected on the LB agar plate containing 50 μg ml^{-1} kanamycin, and subsequently screened by colony PCR.

To enable genomic overexpression from a synthetic operon, conjugation based genetic recombination methods was adapted as previously described [35]. The synthetic operons were digested with BcuI and NotI, and ligated by T4 ligase into previously digested with the same enzyme pDM4 (with oriR6K)

genome integration vector. This vector has two 600bp homology region compatible with target spot, chloramphenicol resistance gene (*cam*R), a levansucrase gene (*sac*B), and the conjugation gene *tra*JI for the transfer of the plasmid. The resulting ligation products were used to transform chemically competent *E. coli* ST18 strains. Positive clones growing on chloramphenicol medium supplemented with 5-aminolevulinic acid (50 µg ml^{-1}) were identified by colony PCR, and the confirmed recombinant ST18 strain was used as donor strain for the conjugation. Chloramphenicol resisting recipient *E. coli* strains were screened as positive strains for the first round of recombination. Subsequently, sucrose counter selection and kanamycin resistance tests were carried out to isolate recombinant *E. coli* strains with the correct synthetic operon integration into chromosome. All constructs were verified via PCR and sequencing.

Introducing point mutations on genome – to establish the mutation shown in Supplementary Figure 6 – was achieved by using multiplex automated genome engineering (MAGE) [43,53]. A single colony of desired strain(s) transformed with pORTMAGE [53] (Addgene catalog no. 72680) was incubated in LB medium supplemented with 100 mg l^{-1} of ampicillin at 30 °C in a shacking incubator. To start the MAGE cycle, overnight cultures were diluted by 100 times in the same medium and cultivated to an optical density of 0.4-0.5 at 600 nm. 1 ml of each culture was transferred to sterile microcentrifuge tubes, and then transferred to 42 °C thermomixer (Thermomixer C, Eppendorf) to express λ-Red genes by heat shock for 15 min at 1000 rpm. After induction, cells were quickly chilled on ice for at least 15 min, and then made electrocompetent by washing three times with ice-cold ddH_2O. 40 µl of electrocompetent cell was mixed with 2 µl of 50 µM of oligomer stock solution and the final volume of the suspension was adjusted to 50 µl. The oligomers used for MAGE were: 5´- T*T*T TTG GCG CTA GAT CAC AGG CAT AAT TTT CAG TAC GTT ATA GGG tGT TTG TTA CTA ATT TAT TTT AAC GGA GTA ACA TTT AGC TCG T*A*C -3´ (*pntAB_MAGE*), 5'-T*A*A AGT TAA ACA AAA TTA TTT CTA TTA ACT AGT GAA TTC GGT CAt TGC GTC CTG CGC ATA TTA TAT GTG AAT CAC AGT GAT ATG TCA A*G*T-3' (*fdh*_MAGE) where the asterisk (*) indicates phosphorothiolated bond. Electroporation was done on Gene Pulser XCell (Bio-Rad) set to 1.8 kV, 25 µF capacitance, and 200 Ω resistance for 1 mm gap cuvette. Immediately after electroporation, 1 ml of LB was added to cuvette and the electroporation mixes in LB was transferred to sterile culture tubes and cultured with shaking at 30 °C, 240 rpm for 1 hour to allow for recovery. After recovery, 2 ml of LB medium supplemented with ampicillin was added and then further incubated in the same condition. When the culture reached an OD_{600} of 0.4-0.5, cells were either subjected to additional MAGE cycle or analyzed for genotype via PCR and sequencing. We have performed 8 consecutive MAGE cycles before analyzing the genotype to identify strains carrying the required mutations.

All strains used are shown in Supplementary Table 1.

Synthetic-Operon construction

Protein sequences of formate-tetrahydrofolate ligase (*ftf*L, UniProt: Q83WS0), 5,10-methenyl-tetrahydrofolate cyclohydrolase (*fch*A, UniProt: Q49135), and 5,10-methylene-tetrahydrofolate dehydrogenase (*mtd*A, UniProt: P55818) were taken from *Methylobacterium extorquens* AM1. Formate

dehydrogenase (*fdh*, UniProt: P33160) was taken from *Pseudomonas* sp. Formaldehyde dehydrogenase were obtained from *Pseudomonas aeruginosa* (*fdh*A, UnitProt: Q9HTE3) and *Pseudomonas putida* (*fdh*A, UnitProt: P46154). Methanol dehydrogenases were prepared from *Bacillus stearothermophilus* (*adh*, UniProt: P42327), *Corynebacterium glutamicum* (cgR_2695, UniProt: A4QHJ5), *Cupriavidus necator* (*mdh*2, UniProt: F8GNE5), and *Bacillus methanolicus* (UnitProt: I3E2P9 and I3E949, as well as en engineered MDH, as reported in [47]). These genes were codon optimized for *E. coli* K-12 and synthesized (Baseclear, Netherlands). Gene sequences are listed in the Supplementary Note.

Genes native to *E. coli* – that is, serine hydroxymethyltransferase (*gly*A) and serine deaminase (*sda*A) – were prepared via PCR-amplification from *E. coli* MG1655 genome. Genes were integrated into a high copy number cloning vector pNiv to construct synthetic operons using the method described previously [35,54]. Plasmid-based gene overexpression was achieved by cloning the desired synthetic operon into the pZ vector (15A origin of replication, streptomycin marker [35]) digested with *EcoR*I and *Pst*I utilizing T4 DNA ligase. All molecular biology techniques were performed with standard methods [55] or following manufacturer protocol.

Promoters and ribosome binding sites were used as described previously [35,54,56]. Briefly, we used either a medium strength constitutive promoter ('PGI-10' [56]) or a strong constitutive promoter ('PGI-20' [56]), as indicated in the text and in Supplementary Figure 1. We further used either medium strength ribosome binding site (RBS_C [54]) or a strong ribosome binding site (RBS_A [54]), as indicated in the text and in Supplementary Figure 1.

All plasmid used are shown in Supplementary Table 1.

Growth medium and Conditions

Luria Bertani medium (1% NaCl, 0.5% yeast extract, and 1% tryptone) was used for strain propagation. Further cultivation was done in M9 minimal media (50 mM Na_2HPO_4,20 mM KH_2PO_4, 1 mM NaCl, 20 mM NH_4Cl, 2 mM $MgSO_4$, and 100 µM $CaCl_2$), with trace elements (134 µM EDTA, 13 µM $FeCl_3 \cdot 6H_2O$, 6.2 µM $ZnCl_2$, 0.76 µM $CuCl_2 \cdot 2H_2O$, 0.42 µM $CoCl_2 \cdot 2H_2O$, 1.62 µM H_3BO_3, 0.081 µM $MnCl_2 \cdot 4H_2O$). For the cell growth test, overnight cultures in LB medium were used to inoculate a pre-culture at an optical density (600 nm, OD_{600}) of 0.02 in 4 ml fresh M9 medium containing 10 mM glucose, 1 mM glycine and 30 mM formate in 10 ml glass test tube. Cell were then cultivated at 37 °C and shaking of 240 rpm. Cell cultures were harvested by centrifugation (18,407 x g, 3 min, 4 °C) and washed twice with fresh M9 medium and used to inoculate the main culture, conducted aerobically either in 10 ml glass tube or Nunc 96-well microplates (Thermo Fisher Scientific) with appropriate carbon sources according to strain and specific experiment: 10 mM glucose, 20 mM acetate, 30 mM formate, 600 mM methanol, and/or 10% CO_2 (90% air). In the microplates cultivation, each well containing 150 µl culture covered with 50 µl mineral oil (Sigma-Aldrich) to avoid evaporation (note that small gaseous molecules such as O_2 and CO_2 can freely diffuse via this oil coverage). Growth experiments were conducted (either 100% air or 90% air / 10% CO_2) in a BioTek Epoch 2 plate reader (BioTek Instrument, USA) at 37 °C. Growth (OD_{600}) was measured after a kinetic cycle of 12 shaking steps, which alternated between linear and orbital (1 mm

amplitude), and were each 60 s long. OD values measured in the plate reader were calibrated to represent OD values in standard cuvettes according to ODcuvette =ODplate/0.23. Glass tube culture was carried out in 4 ml of working volume, at 37 °C and shaking of 240 rpm. Volume loss due to evaporation was compensated by adding the appropriate amount of sterile double distilled water (ddH_2O) to culture tube every two days. All growth experiments were performed in triplicate, and the growth curves shown represent the average of these triplicates.

^{13}C labeling of Proteinogenic Amino acids

For stationary isotope tracing of proteinogenic amino acids, cells were cultured in 4 ml of M9 media supplemented with either labeled or unlabeled carbon sources, that is, ^{13}C-formate, ^{13}C-methanol and/or $^{13}CO_2$, under conditions as described above. A 6 L vacuum desiccator (Lab Companion, South Korea) was used for cultures grown in $^{13}CO_2$, where the original gas was expelled by using vacuum pump followed by refilling with 90% air and 10% $^{13}CO_2$. The cell was harvested by centrifugation for 3 min at 18,407 x g when the stationary growth phase was reached. Biomass was hydrolyzed by incubation with 1 ml of 6 N hydrochloric acid for a duration of 24 h in 95°C. Samples were dried via heating at 95°C and re-dissolved in 1 ml of ddH_2O. Hydrolyzed amino acids were separated using ultra performance liquid chromatography (Acquity, Waters, Milford, MA, USA) using a C18-reversed-phase column (Waters) as previously described [57]. Mass spectra were acquired using an Exactive mass spectrometer (Thermo Fisher). Data analysis was performed using Xcalibur (Thermo Fisher). Prior to analysis, amino-acid standards (Sigma-Aldrich) were analyzed under the same conditions in order to determine typical retention times.

Dry weight analysis

To determine dry cell weight of *E. coli* grown formate or methanol, pre-cultures prepared as described above were inoculated to at a final OD_{600} of 0.01 into fresh M9 medium containing either formate (30 mM) or methanol (600 mM) in 125 ml pyrex Erlenmeyer flask and grown at 37 °C with agitation at 240 rpm. Up to 50 ml of cell culture, growing in shake-flasks, were harvested by centrifugation (3,220 x g, 20 min). To remove residual medium compounds cells were washed be three cycles of centrifugation (7,000 x g, 5 min) and resuspension in 2 ml ddH_2O. Cell-solutions were transferred to pre-weighted and pre-dried aluminum dish, dried at 90 °C for 16 h, and weight of the dried cells in the dish was determined and subtracted by the weight of the empty dish.

CDW of *E. coli* strains was measured during exponential growth phase (OD_{600} of 0.3-0.4) in the presence of 10% CO_2 on 30 mM formate (at OD_{600} of 0.2, 0.37, and 0.41) and on 600 mM methanol (at OD_{600} of 0.21, 0.22, and 0.24). As a control, CDW of *E. coli* strain growing either on formate or methanol was determined during exponential growth phase in the presence of 10% CO_2 and 30 mM formate and either 10 mM glucose (at OD_{600} of 1.26), 20 mM pyruvate (at OD_{600} of 0.78), or 20 mM succinate (at OD_{600} of 0.37). To determine CDW of *E. coli* WT, cells were grown in the presence of 10% CO_2 on 10 mM glucose and CDW was determined during exponential growth phase (at OD_{600} of 0.78).

Enzymes and chemical assays

Absorbance changes for all assays were monitored in a BioTek Epoch 2 plate reader. We confirmed working at the measurement linear range in all assays. Results represent averages of at least three cell preparations. To determine the activity of formate dehydrogenase, 1.5 ml of OD_{600} 1.0 cell culture grown in M9 minimal medium and supplemented with glucose and formate from glass test tubes were washed twice with 9 g l^{-1} sodium chloride. Cells were lysed by adding CelLytic Reagent (Sigma) and allowed to sit for 20 min at the room temperature. After cell disruption, cellular debris was removed by centrifugation (18,407 x g, 4 °C, 10 min) and the supernatant used for crude assays without further purification. Formate dehydrogenase assay performed in the presence of 10 mM 2-mercaptoenthanol, 100 mM sodium formate, 200 mM sodium phosphate buffer pH 7.0, and 2 mM NAD^+ in a total volume of 200 µl at 37 °C [58]. The increase in NADH concentration resulting from formate oxidation was monitored at 340 nm. Protein concentration was measured using the Bradford Reagent (Sigma) with bovine serum albumin as a standard. Formate and methanol in the culture were quantified by a colorimetric assay using formate assay kit (Sigma-Aldrich) and methanol assay kit (BioVision) respectively. All samples were diluted to ensure the reading are within the standard curve range according to the manufacturer's instructions.

Quantitative polymerase chain reaction

Total RNA was extracted from 1 ml of overnight culture at an OD_{600} 0.5 using the RNeasy Mini Kit (Qiagen, Hilden, Germany), and following the protocol of the supplier. All RNA samples were treated with DNase I (Sigma-Aldrich, St. Louis, MO, US) to remove any residual DNA. First-strand cDNA was synthesized using a qScript cDNA Synthesis kit following the manufacturer instructions (Quanta Biosciences, Gaithersburg, MD, US), and 1 µg of total RNA was used as a template in 20 µl reaction volume. Quantitative reverse-transcription-polymerase chain reactions (qRT-PCR) were made using a Maxima™ SYBR Green qPCR Master Mix (ThermoFisher Scientific, Darmstadt, Germany) supplemented with 5 µM primers and 5 µl cDNA template, which was diluted up to 200 µl after synthesis. The primers used for QPCR were: 5´- GCC AAT CTG CAA CAG TGC TC -3´ (*pntA_forward*), 5´- TTT TTG GCT GGA TGG CAA GC -3´ (*pntA reverse*), 5´- CGT GAC GAA TAC CTG ATC GTT -3´ (*fdh* forward), 5´- GGT AGC GTT ACC TTT AGA GTA AGA GTG -3´ (fdh reverse). PCR was performed in 96-well optical reaction plates (ThermoFisher Scientific, Darmstadt, Germany) as follows: 10 min at 50 °C, 5 min at 95 °C, and 40 cycles of 10 s at 95 and 30 s at 60 °C, and finally 1 min at 95 °C. The specificity of the reactions, and the amplicon identities were verified by melting curve analysis. Reaction mixtures without cDNA were used as a negative control. Data were evaluated using the CT method [59] and with correction for the PCR efficiency, which was determined based on the slope of standard curves. Normalization of gene expression levels were carried on by using the *rrs*A gene [60], and eventually the fold-differences in the transcript levels and mean standard error were calculated as described before [59].

Data Availability Statement

Complete information on the experimental setup as well as detailed results are available from the corresponding author upon reasonable request.

Code Availability Statement

MATLAB code used for the analysis of the experiments is available from the corresponding author upon request.

Methods-only References

50 Thoma, S. & Schobert, M. An improved Escherichia coli donor strain for diparental mating. *FEMS Microbiol Lett* **294**, 127-132, doi:10.1111/j.1574-6968.2009.01556.x (2009).

51 Thomason, L. C., Costantino, N. & Court, D. L. E. coli genome manipulation by P1 transduction. *Curr Protoc Mol Biol* **Chapter 1**, Unit 1 17, doi:10.1002/0471142727.mb0117s79 (2007).

52 Baba, T. *et al.* Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**, 2006-2008, doi:10.1038/msb4100050 (2006).

53 Nyerges, A. *et al.* A highly precise and portable genome engineering method allows comparison of mutational effects across bacterial species. *Proc Natl Acad Sci U S A* **113**, 2502-2507, doi:10.1073/pnas.1520040113 (2016).

54 Zelcbuch, L. *et al.* Spanning high-dimensional expression space using ribosome-binding site combinatorics. *Nucleic Acids Res* **41**, e98, doi:gkt151 [pii] 10.1093/nar/gkt151 (2013).

55 Sambrook, J. & Russell, D. W. *Molecular cloning: a laboratory manual*. 3rd edn, (Cold Spring Harbor Laboratory Press, 2001).

56 Braatsch, S., Helmark, S., Kranz, H., Koebmann, B. & Jensen, P. R. Escherichia coli strains with promoter libraries constructed by Red/ET recombination pave the way for transcriptional fine-tuning. *Biotechniques* **45**, 335-337, doi:000112907 [pii] 10.2144/000112907 (2008).

57 Giavalisco, P. *et al.* Elemental formula annotation of polar and lipophilic metabolites using 13C, 15N and 34S isotope labelling, in combination with high-resolution mass spectrometry. *Plant J* **68**, 364-376 (2011).

58 Liu, A., Feng, R. & Liang, B. Microbial surface displaying formate dehydrogenase and its application in optical detection of formate. *Enzyme Microb Technol* **91**, 59-65, doi:10.1016/j.enzmictec.2016.06.002 (2016).

59 Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. *methods* **25**, 402-408 (2001).

60 Zhou, K. *et al.* Novel reference genes for quantifying transcriptional responses of Escherichia coli to protein overexpression by quantitative PCR. *BMC Mol Biol* **12**, 18, doi:10.1186/1471-2199-12-18 (2011).