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 and Arren Bar-Even

Engineering a biotechnological microorganism for growth on one-carbon 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 one-carbon compounds using the reductive glycine pathway. Sequential genomic introgression 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 $\sim 70$ h and growth yield of $\sim 1.5$ g cell dry weight (gCDW) per mol-formate. Short-term evolution decreased doubling time to less than 8 h and improved biomass yield to 2.3 gCDW per mol-formate. Growth on methanol and CO$_2$ was achieved by further expression of a methanol dehydrogenase. Establishing synthetic formatotrophy and methylotrophy, as demonstrated here, paves the way for sustainable bioproduction rooted in CO$_2$ and renewable energy.

Carbon dioxide is a 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, difficulties processing lignocellulosic biomass and, most fundamentally, the low efficiency by which phototrophs use sunlight$^1$. Alternatively, CO$_2$ can be upgraded by purely chemical means; for example, generating syngas$^2$ which can be used to produce complex hydrocarbons$^3$. 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$^4$. Biochemical processes can then convert these C$_1$ compounds into a wide array of chemicals with high specificity under ambient conditions$^5$. Of the possible C$_1$ molecules, formate and methanol are especially interesting. 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 $\geq$40% (ref. $^6$). 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 $\geq$50% (ref. $^7$).

Although 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 (for example, cell concentration and growth rate) and low efficiency of the relevant metabolic pathways$^9$. Adapting a biotechnologically relevant 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 away 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 (rGlyP), 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 (MDH) enables *E. coli* to metabolize methanol to formate, thus supporting growth on this C$_1$ carbon source.

**Results**

The rGlyP. *E. coli*, similar to 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 that could support *E. coli* growth on formate or methanol. Our inspiration came from the anaerobic reductive acetyl-CoA pathway (rAcCoAP)$^{23}$ which assimilates C$_1$ compounds very efficiently.

Max Planck Institute of Molecular Plant Physiology, Potsdam-Golm, Germany. *e-mail: Bar-Even@mpimp-golm.mpg.de*
The rGlyP, as shown in Fig. 1, was designed to be the aerobic twin of the rAcCoAP (ref. 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) and proceed via reduction into a C1-THF intermediate, which is then condensed, within an enzyme complex, with CO₂ to generate a two-carbon (C₂) compound (acetyl-CoA or glycine). The C₂ compound is finally condensed with another C₁ moiety to generate pyruvate as biomass precursor. Importantly, both the rAcCoAP and the rGlyP are characterized by a ‘flat’ thermodynamic profile24,25; 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, although constraining the driving force of the pathway reactions26, indicates very high energetic efficiency, where no energetic input, for example, 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 pyruvate24. Yet, unlike the rAcCoAP, the key enzymatic components of which are highly oxygen sensitive, the rGlyP can operate under aerobic conditions. Hence, the rGlyP represents the most efficient theoretical route—in terms of energy utilization, resource consumption and biomass yield—to assimilate formate in the presence of oxygen24.
A recent study suggests that the complete rGlyP might be naturally occurring 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\(^1,2,3\). However, growth of the bacterium on formate (and CO\(_2\)) has not yet been demonstrated and remains an open challenge.

**Modular engineering establishes growth on formate.** To facilitate the establishment of formotrophic growth, we divided the rGlyP into four metabolic modules (Fig. 1 and Supplementary Fig. 1): (1) a C\(_1\) module (C\(_1\)M), consisting of formate-THF ligase, methenyl-THF cyclohydrolase and methylene-THF dehydrogenase, all from *Methylobacterium extorquens*\(^31\), together converting formate into methylene-THF; (2) a C\(_2\) module (C\(_2\)M), consisting of the endogenous enzymes of the GCS (GcvT, GcvH and GcvP), that condenses methylene-THF with CO\(_2\) and ammonia to give glycine; (3) a three-carbon (C\(_3\)) module (C\(_3\)M), consisting of serine hydroxymethyltransferase (SHMT) and serine deaminase, together condensing glycine with another methylene-THF to generate serine and finally pyruvate; and (4) an energy module (EM), which consists of formate dehydrogenase (FDH) from *Pseudomonas* sp. (strain 101)\(^4\), 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 (ΔserA Δkbl ΔthdA ΔaceA), 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\(^5\). The combined activity of the C\(_1\)M and the C\(_2\)M, together with the native activity of SHMT, should enable the cell to metabolize formate into C\(_1\)-THF, glycine and serine, relieving these auxotrophies (Fig. 2a).

Into the serine auxotroph strain, we introduced the enzymes of the C\(_1\)M and the C\(_2\)M, either on plasmids or in the genome (Supplementary Fig. 1). For genome integration of C\(_2\)M, we combined all relevant enzymes into one operon, under the regulation of a strong constitutive promoter\(^6\), which was inserted into a genomic ‘safe spot’, SS9 (ref. \(^6\)). In the case of the C\(_1\)M, we replaced the native promoter of the GCS with a strong constitutive one (Supplementary Fig. 1), increasing transcript levels 20–50-fold (Supplementary Fig. 2). As expected, growth with formate was observed upon overexpression of both modules (Fig. 2b) and was dependent on 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\(_2\)M (gC\(_2\)M) did not improve growth compared with plasmid expression (pC\(_2\)M), replacing plasmid-borne expression of the enzymes of the C\(_2\)M (pC\(_2\)M) with genomic overexpression (gC\(_2\)M) supported a higher growth rate (Fig. 2b).

Next, we aimed to establish formate as the primary carbon source, which requires high expression of the enzymes of the C\(_1\)M to convert glycine into the central metabolism intermediate pyruvate (Fig. 2c). To enable formate assimilation to biomass, an energy source is required, which at this stage we chose to be acetate. The tricarboxylic acid (TCA) cycle can fully oxidize acetate to generate reducing power and energy, while the deletion of isocitrate lyase (ΔaceA) abolishes the activity of the glyoxylate shunt, thus preventing the cell from using acetate as a carbon source. Growth should therefore depend on formate assimilation via the rGlyP for biomass generation and acetate oxidation for the production of reducing power and energy (Fig. 2c).

The enzymes of the C\(_2\)M were either overexpressed on a plasmid (pC\(_2\)M) or in the genome (gC\(_2\)M) (Supplementary Fig. 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 (ref. \(^6\)). Overexpression of the enzymes of the C\(_2\)M, within a strain that genomically expresses the enzymes of the C\(_1\)M and the C\(_2\)M, resulted in growth on formate and acetate (at 10% CO\(_2\)) (Fig. 2d). Genomic expression of C\(_2\)M supported more robust growth compared with the C\(_2\)M expressed from a plasmid. To confirm that the expression level of C\(_2\)M 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 (ref. \(^6\))). We found this strain to grow rather poorly (Supplementary Fig. 3), indicating that higher expression of these genes is deleterious.

Finally, we introduced the EM such that formate can serve as sole carbon and energy source (Fig. 2e). Overexpression of FDH on a plasmid (Supplementary Fig. 1), in the strain carrying the genes of the C\(_1\)M, C\(_2\)M and C\(_2\)M in the genome, enables growth on formate (Supplementary Fig. 4). However, when we introduced FDH into yet another genomic ‘safe spot’, SS10 (ref. \(^6\)), we failed to establish growth (Supplementary Fig. 4), suggesting that the expression level of FDH was too low. Therefore, we tested a strain in which the genomic expression of FDH was controlled by a stronger ribosome binding site (RBS-A instead of RBS-C (ref. \(^6\))); Supplementary Fig. 1). This strain, carrying no plasmid, was able to grow on formate as a sole carbon and energy source (Fig. 2f and Supplementary Fig. 4).

**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 d, increasing the concentration in the medium by 30 mM (Fig. 3a). Once cell turbidity reached an optical density at 600 nm (OD\(_{600}\)) of 0.4, we diluted the cells to OD\(_{600}\) 0.03–0.05 and started a new cycle of cultivation (Fig. 3a shows six typical cycles).

Within 13 cultivation cycles (≤40 generations), growth rate on formate was substantially improved (Fig. 3a), with the doubling time dropping from 65–80 h in the first two cycles to less than 10 h in the last cycle (Fig. 3b). This growth rate is at least double that of a recently reported *E. coli* strain growing on formate via an engineered Calvin cycle\(^7\). The short-term evolution also improved the growth yield on formate, from ~1.5 gCDW per mol-formate in the first cycle to 2.3 ± 0.2 gCDW per mol-formate. This yield is similar to that of microorganisms growing autotrophically on formate via the Calvin cycle (3.2 ± 1.1 gCDW per mol-formate (ref. \(^8\))). The growth of the evolved bacterium on formate was directly coupled to a decrease in the concentration of the feedstock in the medium (Fig. 3c). Furthermore, as formotrophy 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 Fig. 5).

To better characterize growth on formate, we conducted growth experiments in 96-well plates, automatically measuring OD\(_{600}\) every ~10 min. We found that maximal cell density increased monotonically with increasing formate concentration from 10 mM to 150 mM (Fig. 3d). Similarly, the doubling time decreased monotonically with increasing formate concentration: from 17 h with 10 mM formate to less than 8 h at formate concentrations higher than 100 mM (Fig. 3d). The cellular toxicity of formate, which is attributed to inhibition of respiratory proteins\(^9\) and dissipation of the proton motive force\(^10\), 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 generations 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 was indeed the case, we isolated multiple colonies of the evolved strain and sequenced their genomes. We found two mutations that occurred in all sequenced colonies (Supplementary Fig. 6). The first was a single base-pair substitution in the 5' untranslated region (UTR) of the newly introduced FDH gene, which increased the level of transcript 2.5-fold (Supplementary Fig. 7) and resulted in a 7.4-fold increase in formate oxidation activity in cell extract assays (Supplementary Fig. 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 Fig. 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 apparently limits pathway activity.

To confirm that the two mutations suffice to support the improved growth on formate, we used multiplex automated genome engineering (MAGE) 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 Fig. 9). We therefore concluded that overexpression of FDH and PntAB was 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 Fig. 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/$^{14}$CO$_2$, and measured the labeling pattern of proteinogenic amino acids using liquid chromatography–mass spectrometry. We focused on seven amino acids—glycine, serine, alanine, valine, proline, threonine and histidine—that either directly relate to the activity of the rGlyP or...
Fig. 4 | Labeling pattern of proteinogenic amino acids confirms the activity of the rGlyP. As elaborated in Supplementary Fig. 11, the labeling pattern is consistent with the assimilation of formate and CO₂ 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.

originate from different parts of central metabolism, thus providing an indication of key metabolic fluxes.

As shown in Fig. 4, the amino-acid labeling confirms the activity of the rGlyP. Specifically, feeding [¹³C]-formate/[¹²CO₂] resulted in singly labeled glycine and doubly 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 quadruply labeled, we deduce that pyruvate is labeled in its two noncarboxylic carbons, as predicted for growth via the rGlyP (Supplementary Fig. 11). Conversely, feeding [¹³C]-formate/[¹⁴CO₂] resulted, as expected, in singly labeled glycine, serine and pyruvate. As valine is also singly labeled in this condition, we deduce that pyruvate is labeled in its carbonyl carbon, again confirming the activity of the rGlyP (Supplementary Fig. 11). On feeding [¹³C]-formate/[¹⁴CO₂], all seven amino acids were nearly completely labeled, where the overall fraction of labeled carbon (marked in italics above the bars in Fig. 4) is 97–98%, as expected by feeding with 99% [¹³C]-labeled formate and 99% [¹⁴C]-labeled CO₂.

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 predominates over anaplerotic flux, threonine and proline would be expected to be almost fully labeled on feeding with [¹³C]-formate and almost fully unlabeled when feeding with [¹⁴CO₂] (Supplementary Fig. 11). Conversely, if anaplerotic flux and noncyclic flux predominate over the cyclic flux, then threonine would be expected to be mostly doubly labeled on either [¹³C]-formate or [¹⁴CO₂], and proline would be expected to be mostly quadruply labeled on [¹³C]-formate and singly labeled on [¹⁴CO₂] (Supplementary Fig. 11). The results shown in Fig. 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 wastefully oxidize the assimilated carbons within pyruvate and acetyl-CoA (that is, investing cellular resources for C₄ 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, MDH, can convert methanol to formaldehyde, which can be oxidized to formate by the endogenous glutathione system (44) (Fig. 5a). The expression of MDH can thus be regarded as the introduction of another module—a methanol module—that serves to assimilate methanol to formate, while providing the cells with reducing power (Fig. 5b). We tested NAD-dependent MDHs from several organisms: *Bacillus stearothermophilus* (BsMDH) (45), *Corynebacterium glutamicum* (CgMDH) (46) and *Capriavidus necator* N-1 (NmMDH, wild type mdh2) (47), as well as two MDHs from *Bacillus methanolicus* (BmMDH2) and *BmMDH3* (48) and an improved variant (BmMDH2*, carrying Q5L A363L modifications) (49). These MDH variants were expressed on plasmids in three genetic backgrounds: the parent strain (gCM gCg M gCM gEM), the evolved strain and the parent strain to which the mutation within the promoter of *pntAB* was introduced via MAGE. Overexpression of BsMDH supported growth on 600mM methanol, which was most efficient in the latter strain (Fig. 5c) and somewhat poorer in the other strains (Fig. 5d). The other MDH variants failed to support growth (Fig. 5d, final OD₆₀₀ 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 gene encoding for S-(hydroxymethyl)glutathione dehydrogenase (ΔfrmA) in the above strains. We found that this deletion completely abolished growth on methanol (Fig. 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 (49), or from *Pseudomonas aeruginosa* (PαFADH (ref. 43)) did not improve growth on methanol (Fig. 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 proceeds via the rGlyP, we performed a carbon-labeling experiment. We fed the cultures with [¹³C]-methanol/[¹²CO₂] and measured the labeling pattern of the proteinogenic amino acids glycine, serine, alanine, valine, proline, threonine and histidine. The labeling pattern we measured (Fig. 5e) was essentially identical to that observed with [¹³C]-formate/[¹⁴CO₂] (Fig. 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 ± 5.6h. This can be attributed to the slow rate of methanol oxidation. The observed biomass yield was 4.2 ± 0.17gCDW per mol-methanol, considerably lower than that of microorganisms naturally growing on methanol (7.2 ± 1.2gCDW per mol-methanol via the Calvin cycle, 12 ± 1.6gCDW per mol-methanol via the serine cycle and 15.6 ± 2.7gCDW per mol-methanol via the ribulose monophosphate cycle (48)). 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 100mM sodium bicarbonate increased the final OD₆₀₀ by 30%, but the growth parameters did not improve: doubling time of 55 ± 1h and biomass yield of 4.2 ± 0.1gCDW per mol-methanol (Supplementary Fig. 12, also showing methanol consumption during growth).

Discussion

This study demonstrates synthetic formatotrophy and methylytroph. 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 C1 assimilation pathways according to the biomass and product yields they are expected to support on formate and methanol. 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 regard 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 CO2 as an electron sink. Overall, the rGlyP seems to be the most flexible C1 assimilation pathway, with the potential to support the highest yields of acetyl-CoA and pyruvate using either formate or methanol as feedstocks. However, reaching the full potential of the rGlyP would require considerable growth optimization via rational design and adaptive laboratory evolution.

The C1 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 feedstocks that can be 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 CO2 into renewable commodities.

Online content
Any Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review...
Articles Nature Chemical Biology

24. Bar-Even, A., Noor, E., Flamholz, A. & Milo, R. Design and analysis of metabolic pathways supporting formatargeted growth for electricity-dependent cultivation of microbes. Biochem. Biophys. Acta 1827, 1039–1047 (2013).

25. Bar-Even, A. Does acetogenesis really require especially low reduction potential? Biochim. Biophys. Acta 1827, 397–400 (2015).

26. Noor, E. et al. Pathway thermodynamics highlights kinetic obstacles in central metabolism. PLoS Comput. Biol. 10, e1003483 (2014).

27. Figueroa, I. A. et al. Metagenomics-guided analysis of microbial chemolithoautotrophic phosphate oxidation yields evidence of a seventh natural CO2 fixation pathway. Proc. Natl Acad. Sci. USA 115, E92–E101 (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. Pasterнак, 1. B., Laude, D. A. Jr. & Appl, 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., Matsum, M. M., Atsumi, S. & Kondo, A. Electrical-biological hybrid system for CO2 reduction. Metab. Enzym. 47, 211–218 (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 Synth. Biol. 7, 2023–2028 (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, 5657–5662 (2008).

34. Tishkov, V. I. & Popov, V. O. Catalytic mechanism and application of formate dehydrogenase. Biochem. (Mosc.) 69, 1252–1267 (2004).

35. Wenk, S., Yishai, O., Lindner, S. N. & Bar-Even, A. An engineering approach for rewiring microbial metabolism. Methods Enzymol. 608, 329–367 (2018).

36. Bassalo, M. C. et al. Rapid and efficient one-step metabolic pathway integration in E. coli. ACS Synth. Biol. 5, 561–568 (2016).

37. Gleizer, S. et al. Conversion of Escherichia coli to generate all biomass carbon from CO2. Cell 179, 1255–1263.e12 (2019).

38. Claassens, N. J., Cotton, C. A., Kopjar, D. & Bar-Even, A. Making quantitative sense of microbiolcular production. Nat. Catal. 2, 437 (2019).

39. Nicholls, P. Formate as an inhibitor of cytochrome c oxidase. Biochem. Biophys. Res. Commun. 67, 610–616 (1975).

40. Warmecke, T. & Gill, R. T. Organic acid toxicity, tolerance, and production in Escherichia coli biorefining applications. Micro. Cell Fact. 4, 25 (2005).

41. Dragostis, M. & Mattanovich, D. Adaptive laboratory evolution—principles and applications for biotechnology. Micro. Cell Fact. 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, e1007294 (2018).

43. Wang, H. H. et al. Programming cells by multiplex genome engineering and accelerated evolution. Nature 460, 894–898 (2009).

44. Guthiel, W. G., Kasimoglu, E. & Nicholson, P. C. Induction of glutathione-dependent formaldehyde dehydrogenase activity in Escherichia coli and Hemophilus influenza. Biocenm. Biophys. Res. Commun. 238, 693–696 (1997).

45. Kotbova-Kozak, A., Kotba, P., Inui, M., Sajdok, I. & Yukawa, H. Transcriptionally regulated adaA gene encodes alcohol dehydrogenase required for ethanol and x-propanol utilization in Corynebacterium glutamicum gltA. Appl. Microbiol. Biotechnol. 76, 1347–1356 (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 (2016).

47. Roth, T. B., Woloson, B. M., Stepantovich, G. & Liu, D. R. Phage-assisted evolution of Bacillus methanolicus methanol dehydrogenase 2. ACS Synth. Biol. 8, 796–806 (2019).

48. Zhang, W. et al. Expression, purification, and characterization of formaldehyde dehydrogenase from Pseudomonas aeruginosa. Protein Expr. Purif. 92, 208–213 (2013).

49. Cotton, C. A., Claassens, N. J., Benito-Vaquero, S. & Bar-Even, A. Renewable methanol and formate as microbial feedstocks. Curr. Opin. Biotechnol. 62, 168–180 (2020).

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

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2020
Methods

Chemicals and reagents. Primers were synthesized by Integrated DNA Technologies. 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 Thermofisher Scientific. Glycine, sodium formate, sodium formate-13C and methanol-13C were ordered from Cambridge Isotope Laboratories.

Bacterial strains. Wild-type E. coli strain MG1655 (F−λ−irpV-RF−rF−50-rF−90) were used for cloning and pir-Red recombinase via electroporation. Mutants with exchanged λ strain DH5α (F−λ−irpV-RF−rF−50) were used for selection and conjugation procedures, respectively.

Genome engineering. Gene knockouts were introduced in MG1655 by P1 phage transduction. Single-gene knockout mutants from the National BioResource Project (National Institute of Genetics, Japan) were used as donors of specific mutations. For the recycling of selection marker (as multiple gene deletions and integrations were required) all of the antibiotic cassettes integrated into the genome were flanked by lphspase recognition target (FRT) sites. Cells were transformed with a lphspase recombinase helper plasmid (FLP, 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 FLP plasmid.

Exchange of E. coli native promoter with a synthetic one was performed by using pRed/ET (Gene Bridges) and the conjugation gene (pDM4 (with oriR6K) genome integration vector previously digested with the original promoter (as multiple gene deletions and integrations were required) all of the antibiotic cassettes integrated into the genome were flanked by lphspase recognition target (FRT) sites. Cells were transformed with a lphspase recombinase helper plasmid (FLP, 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 FLP plasmid.

To enable genomic overexpression from a synthetic operon, conjugation-based genetic modification methods were adapted as previously described. The synthetic operons were digested with Bcsl and Notl, and ligated by T4 ligase into pDM4 (with oriR6K) genome integration vector previously digested with the same enzyme. This vector has two 600-base-pair homology regions complementary to the target gene, chloramphenicol resistance cassette (cmar), a leucinase gene (sucB) and the conjugation gene traJ 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 ampicillin were selected for cloning and conjugation procedures, respectively.

Introducing point mutations on the genome—to establish the mutation shown in Supplementary Fig. 6—was achieved by using MAGEASS. A single colony of desired strain(s) transformed with PORTMAGE (Addgene catalog no. 72680) was incubated in LB medium supplemented with 100 mg l−1 ampicillin at 30°C in a shaking incubator. To start the MAGE cycle, overnight cultures were diluted by 100 times in the same medium and cultivated to an OD600 of 0.4–0.5. Then, 1 ml of each culture was transferred to sterile microcentrifuge tubes, and transferred to the 42°C thermomixer (Thermomixer C, Eppendorf) to express λ-Red genes by heat shock for 15 min at 1,000 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 ddH2O. Next, 40 μl of electrocompetent cells were mixed with 2 μl of 50 μM oligomer stock solution and the final volume of the suspension was adjusted to 50 μl. The oligomers used for MAGE were 5′-T′*A*A AGT TAA ACA AAA TTA TTT CTA TTA ACT AGT GAA TTC GGT GAT CAC AGG CAT AAT TTT CAG TTA ATC ATT TTG GGC GGA GTA ACA TTT AGC TGC TAC-3′ (protAB_MAGE), 5′-T*A*A AGA TTT TAC ACT TTA ATC CGT GAC CAT TGC GTC CTG CGC ATG TTA TAT GAT CAC GAT GAT ATC AAG-3′ (lbd_MAGE), where the asterisk (*) indicates a phosphorothioated modification. Electroporation was done on a Gene Pulser Xcell (Bio-Rad) set to 1.8 kV, 25 μF capacitance and 200 Ω resistance for a 1-mm gap cuvette. Immediately after electroporation, 1 ml of electrocompetent cells was added to the cuvette and the electroporation mix in LB was transferred to sterile culture tubes and cultured with shaking at 30°C, 240 rpm for 1 h to allow for recovery. After recovery, 2 ml of LB medium supplemented with ampicillin was added and then incubated in the same conditions. When the culture reached an OD600 of 0.4–0.5, cells were either subjected to additional MAGE cycles or analyzed for genotype by PCR and sequencing. We performed eight consecutive MAGE cycles before analyzing the genome to identify the required mutations.

All strains used are shown in Supplementary Table 1.

Synthetic operon construction. Protein sequences of formate-TFH ligase (jfl, UniProt:Q835W0), 5,10-methenyl-TFH cytochrome (fda, UniProt:Q89135) and 5,10-methylene-TFH dehydrogenase (mda, UniProt:Q55818) were taken from M. extorquens AM1. FDH (jfl, UniProt: P31630) was taken from Pseudomonas sp. Formaldehyde dehydrogenase were obtained from P. aeruginosa (fda, UniProt: Q9HTE3) and P. putida (jfl, UniProt: P46154). MDHs were prepared from B. stearothermophilus (SdhB, UniProt: P42327), C. glutamicum (C4A, UniProt: A4QJH4), C. asaccharolyticum (C4A, UniProt: A4QJH4), and P. stutzeri (SdhB, UniProt: P42327). Methyl-coenzyme M reductase (mcr, UniProt: P49054) was ordered from Thermo Fisher Scientific. Glycine, sodium formate, sodium formate-13C and 13CO2 were ordered from Sigma-Aldrich. 13CO2 was ordered from Sigma-Aldrich.

Promoters and ribosome binding sites were used as described previously. Briefly, we used either a medium-strength constitutive promoter (‘PGI-10’ (ref. 18)) or a strong constitutive promoter (‘PGI-20’ (ref. 19)), as indicated in Supplementary Fig. 1. We further used either a medium-strength ribosome binding site (RBS, (ref. 20)) or a strong ribosome binding site (RBS, (ref. 21)), as indicated in Supplementary Fig. 1.

All plasmids used are shown in Supplementary Table 1.

Growth medium and conditions. LB medium (1% NaCl, 0.5% yeast extract and 1% tryptone) was used for strain propagation. Further cultivation was done in M9 minimal medium (50 mM NaPO4, 20 mM KH2PO4, 1 mM NaCl, 20 mM NH4Cl, 50 mM MgSO4, and 100 mM CaCl2) with trace elements (134 mM EDTA, 13 μM FeCl3·6H2O, 6.2 μM ZnCl2, 0.76 μM CuCl2, 20 μM MgCl2·2H2O, 1.62 μM H3BO3, 0.081 μM MnCl2·4H2O). For the cell growth test, overnight cultures in LB medium were used to inoculate a pre-culture at an OD600 of 0.02 in 4 ml of fresh M9 medium containing 10 mM glucose, 1 mM glycine and 30 mM formate in 10 ml glass test tubes. Cells were then cultivated at 37°C with shaking at 240 rpm. Cell cultures were collected by centrifugation (18,407 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 tubes 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% CO2 (90% air). In the microplate cultivation, each well contained 150 μl of culture covered with 50 μl of mineral oil (Sigma-Aldrich) to avoid evaporation (note that small gaseous molecules such as CO2 and CO can freely diffuse through this oil cover). Growth experiments were conducted (either 100% air or 90% air/10% CO2) using a BioTek Epoch 2 plate reader (BioTek Instruments) at 37°C. Growth (OD600) 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 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 at 240 rpm. Volume loss due to evaporation was compensated by adding the appropriate amount of sterile ddH2O to culture tubes every 2 d. All growth experiments were performed in triplicate, and the growth curves shown represent the average of these triplicates.

13C labeling of proteinogenic amino acids. For stationary isotope tracing of proteinogenic amino acids, cells were cultured in 4 ml of M9 medium supplemented with either labeled or unlabeled carbon sources; that is, 13C-formate, 13C-methanol and/or 13C-O2. A 6-l vacuum desiccator (Lab Companion) was used for cultures grown in 13C2O2, where the original gas was expelled by vacuum pump followed by refilling with 90% air and 10% CO2. The cell was collected by centrifugation for 5 min at 18,407 g when the stationary growth phase was reached. Biomass was hydrolyzed by incubation with 1 ml of 6N hydrochloric acid for a duration of 24 h at 95°C. Samples were dried via heating at 95°C and re-dissolved in 1 ml of ddH2O. Hydrolyzed amino acids were separated using ultra-performance liquid chromatography (Acquity, Waters) using a C18-reversed phase column (Waters) mass spectrometric spectra were used for analysis by a high-resolution mass spectrometer (Thermo Fisher). Data analysis was performed using Xcalibur (Thermo Fisher). Before analysis, amino-acid standards (Sigma-Aldrich) were analyzed under the same conditions to determine typical retention times.

Dry weight analysis. To determine the dry cell weight of E. coli grown on formate or methanol, preculures were inoculated at a final OD600 of 0.01 into fresh M9 medium containing either formate (30 mM) or methanol (600 mM) in 125 ml
Pyrex Erlenmeyer flasks and grown at 37°C with agitation at 240 r.p.m. Up to 50 ml of cell cultures, growing in shake flasks, were collected by centrifugation (3,220 g, 20 min). To remove residual medium compounds, cells were washed by three cycles of centrifugation (7,000 g, 5 min) and resuspension in 2 ml ddH2O. Cell solutions were transferred to a preweighed and predried aluminum dish and dried at 90°C for 16 h, and then the weight of the dried cells in the dish was determined by subtracting the weight of the empty dish.

The cell dry weight (CDW) of *E. coli* strains was measured during exponential growth phase (OD600, 0.3–0.4) in the presence of 10% CO2, 30 mM formate (at OD600 of 0.2, 0.37 and 0.41) and on 600 mM methanol (at OD600 of 0.21, 0.22 and 0.24). As a control, the CDW of *E. coli* strains growing either on formate or methanol was determined during exponential growth phase in the presence of 10% CO2, 30 mM formate and 10 mM glucose (at OD600 of 1.26), 20 mM pyruvate (at OD600 of 0.78) or 20 mM succinate (at OD600 of 0.37). To determine the CDW of *E. coli* wild type, cells were grown in the presence of 10% CO2, on 10 mM glucose and CDW was determined during exponential growth phase (at OD600 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 FDH, 1.5 ml of OD600 1.0 cell culture grown in M9 minimal medium and supplemented with glucose and formate from glass test tubes was washed twice with 9 g/l sodium chloride. Cells were lysed by adding Celllytic Reagent (Sigma) and allowed to sit for 20 min at room temperature. After cell disruption, cellular debris was removed by centrifugation (18,407 g, 4 °C, 10 min) and the supernatant was used for crude assays without further purification. FDH assay was performed in the presence of 10 mM 2-mercaptoethanol, 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 (ref. 37). The increase in NADH concentration resulting from formate oxidation was monitored at 340 nm. Protein concentration was measured using the Bradford Reagent (Sigma) with BSA as a standard. Formate and methanol in the culture were quantified by a colorimetric assay using a formate assay kit (Sigma-Aldrich) and a methanol assay kit (BioVision), respectively. All samples were diluted to ensure the readings were within the standard curve range according to the manufacturer’s instructions.

Quantitative PCR. Total RNA was extracted from 1 ml of overnight culture at an OD600 of 0.5 using the RNeasy Mini Kit (Qiagen), and following the protocol of the supplier. All RNA samples were treated with DNase I (Sigma-Aldrich) to remove any residual DNA. First-strand complementary DNA was synthesized using a qScript cDNA Synthesis kit following the manufacturer instructions (Quanta Biosciences). 1 μl of total RNA was used as a template in a 20-μl reaction volume. Quantitative PCR with reverse transcription was performed using a Maxima SYBR Green qPCR Master Mix (Thermo Fisher Scientific) supplemented with 5 μM primers and 5 μl of cDNA template, which was diluted up to 200 μl after synthesis. The primers used for quantitative PCR were: 5’-GCC AAT CTG CAA CAG TGC TC-3’ (pntA forward), 5’-TTT TGT GCT GTA TGG CAA GC-3’ (pntA reverse), 5’-CGT GAC GAA TAC CTG ATC GTT-3’ (fdh forward), 5’-GGT AGG GTC ATT ACC TTT AGA GTA ATG-3’ (fdh reverse). PCR was performed in 96-well optical reaction plates (Thermo Fisher Scientific) as follows: 10 min at 50°C, 5 min at 95°C, 40 cycles of 10 s at 95°C 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 delta-delta Ct method and with correction for the PCR efficiency, which was determined based on the slope of standard curves. Normalization of gene expression levels was carried out using the *rrsA* gene, and eventually the fold-differences in the transcript levels and mean standard error were calculated as described before.

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

Data availability

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

Code availability

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

References

50. Thoma, S. & Schobert, M. An improved *Escherichia coli* donor strain for diparental mating. *FEMS Microbiol. Lett.* 294, 127–132 (2009).
51. Thomason, L. C., Costantino, N. & Court, D. L. *E. coli* genome manipulation by PI transduction. *Curr. Protoc. Mol. Biol.* https://doi.org/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 (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. USA* 113, 2502–2507 (2016).
54. Zelburch, L. et al. Spanning high-dimensional expression space using ribosome-binding site combinatorics. *Nucleic Acids Res.* 41, e6 (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 (2008).
57. Giavalisco, P. et al. Elemental formula annotation of polar and lipophilic metabolites using 13C, 15N and 3H 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. *Enzym. Microb. Technol.* 91, 59–65 (2016).
59. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2⁰ΔΔCT 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 (2011).

Acknowledgements

We thank C. Cotton, N. Claessens, H. He, R. Milo, E. Noor, N. Antonovsky, A. Flamholz, Y. Bar-On, W. Newell, D. Bonder, A. Satonowski, T. Erd and M. Bouzon for helpful discussions and 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 eCoFuel).

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

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

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41589-020-0473-5.

Correspondence and requests for materials should be addressed to A.B.-E.

Reprints and permissions information is available at www.nature.com/reprints.
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

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

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

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

| Data collection       | Biotek Gen5 v3.04 (plate reader) |
|-----------------------|----------------------------------|
| Data analysis         | MATLAB R2015a (Mathworks)        |

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

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The bacterial strains developed in this study, as well as information regarding their growth, are available from the corresponding author upon reasonable request.

Field-specific reporting

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

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

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-list.pdf
Life sciences study design

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

Sample size: This study was conducted using single microorganism in the controlled laboratory environment and focused on its performance. All growth experiments were performed in triplicates.

Data exclusions: No data were excluded from the analysis.

Replication: Reproducibility was verified by conducting the same experiments multiple times with freshly prepared cells from glycerol stock. All attempts to repeat the experiment were successful.

Randomization: Since this study was conducted using single microorganism in the controlled laboratory environment and focused on its performance, the allocation is not relevant to this study.

Blinding: Since this study was conducted using single microorganism in the controlled laboratory environment and focused on its performance, the blinding is not relevant to this study.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

| n/a | Involved in the study |
|-----|-----------------------|
| ☒   | Antibodies            |
| ☒   | Eukaryotic cell lines  |
| ☒   | Paleontology          |
| ☒   | Animals and other organisms |
| ☒   | Human research participants |
| ☒   | Clinical data         |

Methods

| n/a | Involved in the study |
|-----|-----------------------|
| ☒   | ChiP-seq              |
| ☒   | Flow cytometry        |
| ☒   | MRI-based neuroimaging |