

Hai He | Rune Höper | Moritz Dodenhöft | Philippe Marlière |  
Arren Bar-Even

# An optimized methanol assimilation pathway relying on promiscuous formaldehyde-condensing aldolases in *E. coli*

Suggested citation referring to the original publication:

Metabolic Engineering 60 (2020) 1-13

DOI <https://doi.org/10.1016/j.ymben.2020.03.002>

ISSN (print) 1096-7176

ISSN (online) 1096-7184

Postprint archived at the Institutional Repository of the Potsdam University in:

Postprints der Universität Potsdam

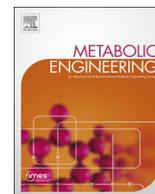
Mathematisch-Naturwissenschaftliche Reihe ; 997

ISSN 1866-8372

<https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-476452>

DOI <https://doi.org/10.25932/publishup-47645>





# An optimized methanol assimilation pathway relying on promiscuous formaldehyde-condensing aldolases in *E. coli*

Hai He<sup>a</sup>, Rune Höper<sup>a</sup>, Moritz Dodenhöft<sup>a</sup>, Philippe Marlière<sup>b,\*\*</sup>, Arren Bar-Even<sup>a,\*</sup>

<sup>a</sup> Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, 14476, Potsdam-Golm, Germany

<sup>b</sup> TESSSI, The European Syndicate of Synthetic Scientists and Industrialists, 81 rue Réaumur, 75002, Paris, France

## ARTICLE INFO

### Keywords:

Pathway design  
Promiscuous enzymes  
Formaldehyde assimilation  
Serine cycle  
Growth selection

## ABSTRACT

Engineering biotechnological microorganisms to use methanol as a feedstock for bioproduction is a major goal for the synthetic metabolism community. Here, we aim to redesign the natural serine cycle for implementation in *E. coli*. We propose the homoserine cycle, relying on two promiscuous formaldehyde aldolase reactions, as a superior pathway design. The homoserine cycle is expected to outperform the serine cycle and its variants with respect to biomass yield, thermodynamic favorability, and integration with host endogenous metabolism. Even as compared to the RuMP cycle, the most efficient naturally occurring methanol assimilation route, the homoserine cycle is expected to support higher yields of a wide array of products. We test the *in vivo* feasibility of the homoserine cycle by constructing several *E. coli* gene deletion strains whose growth is coupled to the activity of different pathway segments. Using this approach, we demonstrate that all required promiscuous enzymes are active enough to enable growth of the auxotrophic strains. Our findings thus identify a novel metabolic solution that opens the way to an optimized methylotrophic platform.

## 1. Introduction

Microbial production of commodity chemicals is limited by feedstock availability and cost. Sugars and starches, despite being commonly used, are not ideal microbial feedstocks as their biotechnological utilization directly competes with human consumption, thus eroding food security (Walker, 2009). Furthermore, the expansion of agricultural cultivation comes at the expense of shrinking natural habitats, hence threatening biodiversity (Fitzherbert et al., 2008). The use of lignocellulosic biomass, while avoiding some of these problems, presents other challenges, including heterogenic composition, difficult processing, and deleterious waste products (Sanderson, 2011). One carbon compounds provide a favorable alternative as they can be produced at high levels without burdening agricultural production and they represent homogenous, easy-to-handle microbial feedstocks (Schrader et al., 2009; Takors et al., 2018; Yishai et al., 2016). Methanol is especially interesting as it is completely water miscible, avoiding mass transfer barriers that constrain the use of gaseous one carbon compounds (e.g., carbon monoxide and methane). Methanol can be produced at low-cost from fossil or renewable methane (Zakaria and Kamarudin, 2016), or it can be produced sustainably and efficiently from CO<sub>2</sub> and electrochemically derived hydrogen (Szima and Cormos,

2018).

There has been much recent progress in the metabolic engineering of microorganisms that naturally grow on methanol, e.g., *Methylobacterium extorquens* (Marx and Lidstrom, 2004; Schada von Borzyskowski et al., 2015). Still, when compared to biotechnological microorganisms, such as *Escherichia coli*, natural methylotrophs are more difficult to manipulate and engineer (Pfeifenschneider et al., 2017). To alleviate this problem, multiple recent efforts have sought to engineer biotechnological hosts for growth on methanol via one of the naturally occurring methanol assimilation pathways (Wang et al., 2020; Whitaker et al., 2015; Zhang et al., 2017): the ribulose monophosphate (RuMP) cycle (Chen et al., 2018; He et al., 2018; Meyer et al., 2018), the dihydroxyacetone (DHA) cycle (Dai et al., 2017), and the serine cycle (Yu and Liao, 2018). However, these natural pathways might not represent optimal solutions. Better pathways, more efficient in the use of cellular resources and/or more metabolically compatible with the host microorganism, could be designed and implemented. For example, a recent study designed and partially implemented a modified serine cycle in *E. coli*, where some of the natural reactions were replaced with others, better fitting the endogenous metabolism of the host (Yu and Liao, 2018). However, the serine cycle and its modified variants are ATP-inefficient, which results in low biomass and product yields

\* Corresponding author.

\*\* Corresponding author.

E-mail addresses: [pmarliere@isthmus.fr](mailto:pmarliere@isthmus.fr) (P. Marlière), [Bar-Even@mpimp-golm.mpg.de](mailto:Bar-Even@mpimp-golm.mpg.de) (A. Bar-Even).

<https://doi.org/10.1016/j.ymben.2020.03.002>

Received 15 January 2020; Received in revised form 26 February 2020; Accepted 6 March 2020

Available online 10 March 2020

1096-7176/ © 2020 The Author(s). Published by Elsevier Inc. on behalf of International Metabolic Engineering Society. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

(Claassens et al., 2019). In other studies, new-to-nature reactions were demonstrated, including the formolase reaction that condenses formaldehyde, the direct product of methanol oxidation, to glycolaldehyde or DHA (Lu et al., 2019; Wang et al., 2017; Yang et al., 2019), or hydroxyacyl-CoA lyase that condenses formyl-CoA with formaldehyde to generate glycolyl-CoA (Chou et al., 2019). However, the condensation rate and/or affinity for formaldehyde of these reactions are too low to be physiologically relevant.

In this study, we aim to upgrade the serine cycle for implementation in *E. coli*. By applying several design principles – including maximization of biomass yield and thermodynamic favorability as well as replacement of CO<sub>2</sub> fixation with formaldehyde assimilation – we devise the homoserine cycle, which relies on two promiscuous formaldehyde-condensing aldolase reactions. We show that this synthetic route is expected to outperform the serine cycle and its modified variants in multiple key metrics. Moreover, even when compared to the RuMP cycle, the most efficient natural methanol assimilation pathway, the homoserine cycle is expected to support higher yields of products that are derived from acetyl-CoA, including ethanol, acetone, butyrate, butanol, citrate, itaconate, 2-ketoglutarate, and levulinic acid. As such, this cycle can outcompete all the natural aerobic methanol assimilation pathways for a wide array of products.

While establishing growth via the complete homoserine cycle is a challenging task that would require extensive further research, here we demonstrate that *E. coli* native enzymes can promiscuously catalyze all non-natural reactions of the homoserine cycle in a physiologically relevant manner, producing essential cellular building blocks. Overall, our work confirms the feasibility of a novel methanol assimilation pathway that could pave the way for future implementation of highly efficient conversion of this one carbon feedstock into commodity chemicals.

## 2. Results

### 2.1. Design of the homoserine cycle

Inspecting the structure of the serine cycle (Fig. 1a and Supplementary Table S1), we identified three key shortcomings: (i) Three of the pathway reactions participate in the central metabolism of the host (the pentose phosphate pathway, glycolysis, and the TCA cycle). This could lead to competition for flux between the pathway reactions and those of central metabolism, thus making regulation highly challenging (Bar-Even, 2016). (ii) The serine cycle is ATP-inefficient, consuming 3 ATP molecules per acetyl-CoA generated. This lowers the biomass and product yields as, rather than being assimilated into biomass, a substantial fraction of the substrate methanol needs to be oxidized to generate the required ATP (Claassens et al., 2019). (iii) For each formaldehyde molecule assimilated by the serine cycle, one CO<sub>2</sub> molecule is fixed. While CO<sub>2</sub> fixation has the advantage of serving as an electron sink for access reducing power (Cotton et al., 2020), it leads to two key problems. First, as carboxylation is thermodynamically unfavorable, it needs to be coupled directly or indirectly to ATP hydrolysis, thus reducing the ATP efficiency of the pathway (Bar-Even et al., 2012). In the case of the serine cycle, glycerate kinase consumes an ATP that is subsequently used to energize carboxylation by phosphoenolpyruvate carboxylase. Second, CO<sub>2</sub> fixation must be accompanied by two reduction steps, in order to bring the carbon to the average oxidation state of carbon in biomass. These extra reduction steps require more of the substrate to be completely oxidized to generate NAD(P)H, thus necessitating a delicate regulation between substrate assimilation and substrate oxidation.

We aimed to design a more applicable version of the serine cycle, which overcomes the shortcomings of the natural pathway. Rather than restricting the design to primary enzymatic reactions, we considered also promiscuous enzyme activities, i.e., side reactions. This dramatically expands the metabolic solution space and thus assists in

identifying optimal pathway structures (Trudeau et al., 2018). To facilitate pathway implementation, we further aimed to employ only *E. coli* enzymes, reduce their overall number, and avoid overlap with central metabolism.

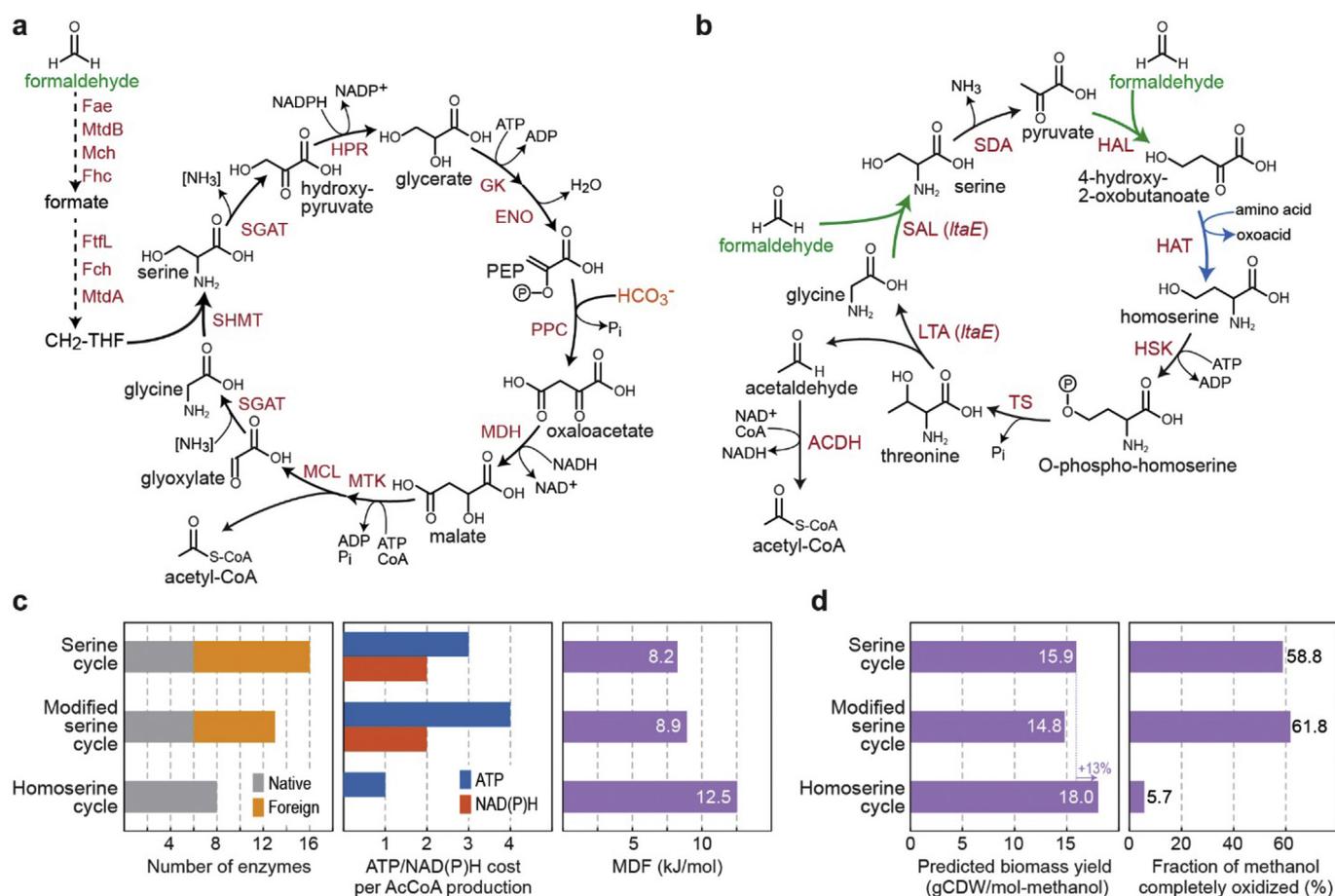
Importantly, we aimed the new pathway to be independent on carboxylation so as to lower the demand for NAD(P)H generation by complete substrate oxidation. A formaldehyde assimilation reaction should ideally replace the existing carboxylation reaction, as, unlike CO<sub>2</sub>, formaldehyde is already at the average oxidation state of cellular carbon. Similarly, consumption of ATP should be minimized in order to reduce substrate oxidation and increase biomass (and product) yield. Despite the reduction of ATP consumption, the pathway should be at least as thermodynamically favorable as the serine cycle (Noor et al., 2014).

Following a comprehensive literature search of formaldehyde condensing reactions catalyzed by *E. coli* native enzymes, we identified the homoserine cycle (Fig. 1b and Supplementary Table S1) as a superior design. In the homoserine cycle, glycine is directly condensed with formaldehyde to generate serine. This serine aldolase (SAL) reaction (Fig. 1b) was previously found to be promiscuously catalyzed (*in vitro*) by threonine aldolase (LtaE) (Contestabile et al., 2001). The SAL reaction bypasses the very long, multi-cofactor-dependent, and ATP-inefficient route for formaldehyde assimilation to 5,10-methylene-tetrahydrofolate (CH<sub>2</sub>-THF) (Crowther et al., 2008) (Fig. 1b). As within the previously proposed modified serine cycles (Bar-Even, 2016; Yu and Liao, 2018), serine is then deaminated to pyruvate by serine deaminase (SDA in Fig. 1b), bypassing a longer route via glycerate, which further involves the highly toxic intermediate hydroxypyruvate (Kim and Copley, 2012) (Fig. 1b). Despite being rather oxygen-sensitive, serine deaminase was shown to support high flux under aerobic conditions (Yu and Liao, 2018).

Pyruvate is then condensed with formaldehyde to generate the non-native metabolite 4-hydroxy-2-oxobutanoate (HOB) (Bouzon et al., 2017), which is subsequently aminated to homoserine. The first of these reactions – HOB aldolase (HAL in Fig. 1b) – was found to be promiscuously catalyzed by *E. coli* 2-keto-3-deoxy-L-rhamnonate aldolase (RhmA) (Hernandez et al., 2017) as well as by similar aldolases (Wang et al., 2019). The latter reaction – HOB amination (HAT in Fig. 1b) – is supported by numerous aminotransferases (Hernandez et al., 2017; Walther et al., 2018; Zhong et al., 2019) as well as amino acid dehydrogenases such as (engineered) glutamate dehydrogenase (Chen et al., 2015). This route effectively replaces a carboxylation reaction (by phosphoenolpyruvate carboxylase) with a formaldehyde assimilation reaction that provides an alternative way to generate a C<sub>4</sub> intermediate. Homoserine is then metabolized by homoserine kinase (ThrB, HSK in Fig. 1b) and threonine synthase (ThrC, TS in Fig. 1b) to produce threonine. Finally, threonine is cleaved by threonine aldolase (LTA in Fig. 1b, catalyzed by the same LtaE that catalyzes the SAL reaction) to regenerate glycine and produce acetaldehyde which can be further oxidized to acetyl-CoA and assimilated to central metabolism.

### 2.2. The homoserine cycle could outperform natural methanol assimilation pathways

We compared the homoserine cycle to the serine cycle and the modified serine cycle according to multiple criteria. Assuming formaldehyde as the substrate of the pathways, we find that: (i) the homoserine cycle requires only eight enzymes, which is half the number of enzymes needed for the serine cycle (16) and ~40% fewer enzymes than the modified serine cycle (13, Fig. 1c); also, unlike the other pathways, the homoserine cycle is not dependent on foreign enzymes. (ii) The homoserine cycle consumes a single ATP molecule for the production of acetyl-CoA, while the serine cycle requires 3 ATP molecules and the modified serine cycle needs 4 ATP equivalents (Fig. 1c). (iii) The homoserine cycle does not lead to the net consumption of NAD(P)H, while the other two pathways consume 2 NAD(P)H per acetyl-



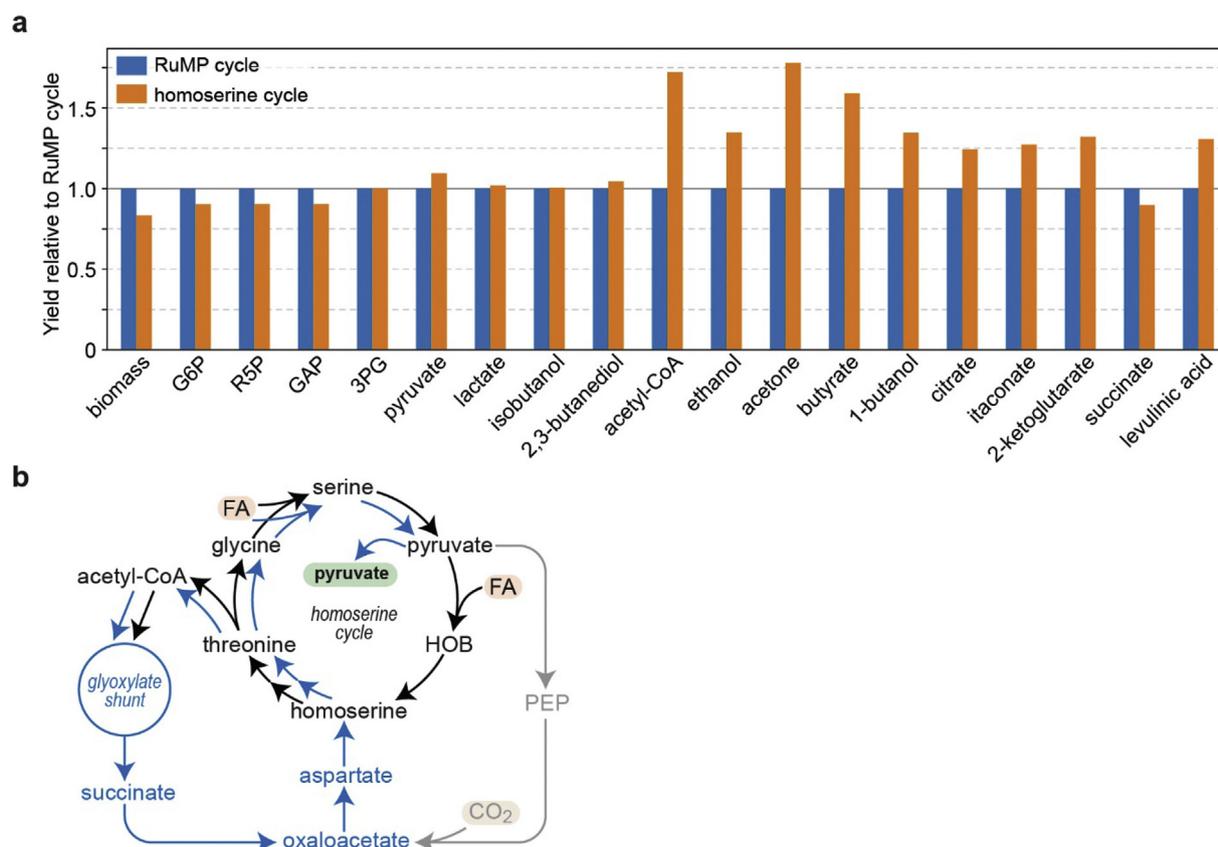
**Fig. 1. Natural and synthetic methanol assimilation pathways.** (a) The natural serine cycle (formaldehyde as feedstock). Pathway enzymes and reactions are listed in Supplementary Table S1 (b) The synthetic homoserine cycle (formaldehyde as feedstock). The homoserine cycle harbors three promiscuous enzyme activities: serine aldolase (SAL), 4-hydroxy-2-oxobutanoate (HOB) aldolase (HAL), and the HOB aminotransferase (HAT). The other pathway reactions include threonine biosynthesis (homoserine kinase, HSK; threonine synthase, TS) and cleavage (threonine aldolase, LTA) as well as serine deaminase (SDA). Note that both SAL and LTA are catalyzed by the same *LtaE* enzyme. Acetylating acetaldehyde dehydrogenase (ACDH) can convert the direct pathway product acetaldehyde to acetyl-CoA. (c) The homoserine cycle outperforms the natural serine cycle and the previously suggested modified serine cycle (Yu and Liao, 2018) in terms of simplicity (small number of reactions), resource consumption efficiency and thermodynamic driving force as indicated by the Max-min Driving Force (MDF) criterion (Noor et al., 2014). These analyses were based on formaldehyde as pathway feedstock (Methods). (d) The homoserine cycle is predicted by Flux Balance Analysis to support higher biomass yield on methanol than the natural serine cycle and modified serine cycle. Furthermore, unlike the other pathways, the homoserine cycle is almost independent on the complete oxidation of methanol to provide the cell with reducing power and energy.

CoA (Fig. 1c); as mentioned above, this difference can be attributed to the fact that the homoserine cycle replaces a carboxylation reaction with another formaldehyde assimilation reaction. (iv) Despite its low resource consumption, the homoserine cycle is more energetically favorable than the other two pathways as its Max-min Driving Force – MDF, representing the minimal thermodynamic driving force via the pathway reactions after optimizing metabolite concentrations within a physiological range (Noor et al., 2014) – is substantially higher (Fig. 1c, Supplementary Fig. S1). The higher thermodynamic driving force could assist in pulling methanol oxidation forward, a reaction which is thermodynamically unfavorable and represents a major constraint for establishing synthetic methylotrophy.

We further used Flux Balance Analysis (FBA) to calculate the maximal biomass yield with each of the three pathways using methanol as a substrate (Methods). We found the expected biomass yield of the homoserine cycle to be 13% higher than with the serine cycle, while the modified serine cycle supported an even lower yield (Fig. 1d). Moreover, while the operation of the two other pathways requires > 50% of the methanol to be completely oxidized to provide NAD(P)H and ATP, only ~6% of methanol is completely oxidized during growth via the homoserine cycle (Fig. 1d). This indicates that the establishment of highly efficient formaldehyde and formate oxidation systems is strictly

required for the operation of the serine cycle and its modified variants, while the homoserine cycle could operate without such systems, in which case the biomass yield drops by less than 1%. The ability to bypass formaldehyde and formate oxidation is key for the establishment of this synthetic route, as it circumvents the need for a delicate fine-tuning of the formaldehyde oxidation flux relative to the formaldehyde assimilation flux.

It is commonly argued that the RuMP cycle is the most efficient naturally occurring route for methanol assimilation. Indeed, among all aerobic methanol assimilation pathways, the RuMP cycle supports the highest biomass yield (Claassens et al., 2019). We used FBA to compare the maximal yields of various central metabolites, precursors, and chemicals of interest using methanol as a substrate and either the RuMP cycle or the homoserine cycle as assimilation routes (Methods). While the RuMP cycle was able to sustain higher yields of biomass and phosphosugars, the homoserine cycle supported substantially higher yields of compounds derived from acetyl-CoA, including ethanol, acetone, butyrate, 1-butanol, citrate, itaconate (Zhao et al., 2018), 2-ketoglutarate, levulinic acid (Cheong et al., 2016), and adipic acid (Yu et al., 2014) (Fig. 2a). Both pathways supported similar yields of pyruvate and its derivatives, including lactate, isobutanol, and 2,3-butanediol (Fig. 2a).



**Fig. 2. The homoserine cycle can outperform the RuMP cycle for the production of various compounds.** (a) We used Flux Balance Analysis to calculate the maximal yields of various central metabolites, precursors, and products using methanol as carbon source and either the RuMP cycle or the homoserine cycle as assimilation route. The values were normalized to the yields via the RuMP cycle. (b) Production of pyruvate, or other  $C_{\geq 3}$  metabolites, via the homoserine cycle involves assimilation of acetyl-CoA via the glyoxylate shunt, followed by conversion of succinate to homoserine, thus reentering the homoserine cycle (blue arrows). Excess reducing power, produced from methanol oxidation, can use  $CO_2$  as an electron sink, where a fraction of the pyruvate molecules is diverted to PEP synthesis and anaplerosis (grey arrows), and subsequently reduced to homoserine. FA corresponds to formaldehyde. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Why does homoserine cycle outcompete the RuMP cycle for so many products? Methanol is a highly reduced carbon source, more reduced than most products. Excess electrons generated during methanol bioconversion are channeled to the respiratory chain and dissipated wastefully. The more electrons are dissipated in this way, the greater the loss in potential yield (Cotton et al., 2020). The RuMP cycle especially suffers from electron overflow as it does not use an electron sink such as  $CO_2$  and further releases  $CO_2$  during the oxidative biosynthesis of acetyl-CoA (Bogorad et al., 2014). The homoserine cycle, on the other hand, generates acetyl-CoA directly without the release of  $CO_2$ . Moreover, while the homoserine cycle is not strictly dependent on carboxylation, it can channel a fraction of its flux via a carboxylation route that serves as sink for excess reducing power: some of the pyruvate, instead of reacting with formaldehyde, can be converted to homoserine via the carboxylation-dependent anaplerotic route, which serves as an electron sink (grey arrows in Fig. 2b). The flexibility of the homoserine cycle in terms of  $CO_2$  use – being generally independent of carboxylation while able to moderately fix  $CO_2$  when necessary – is the main reason of the high yields it can support. Furthermore, conversion of acetyl-CoA, the product of the homoserine cycle, into pyruvate and other  $C_{\geq 3}$  compounds does not involve wasteful decarboxylation: succinate, the product of the glyoxylate shunt, can be reintegrated directly into the homoserine cycle to produce pyruvate (blue arrow in Fig. 2b).

Overall, the homoserine cycle seems to be preferable to its counterparts, with the potential to outperform both the serine cycle and the RuMP cycle for the production of a wide array of value-added chemicals. However, as the synthetic pathway relies on promiscuous enzyme

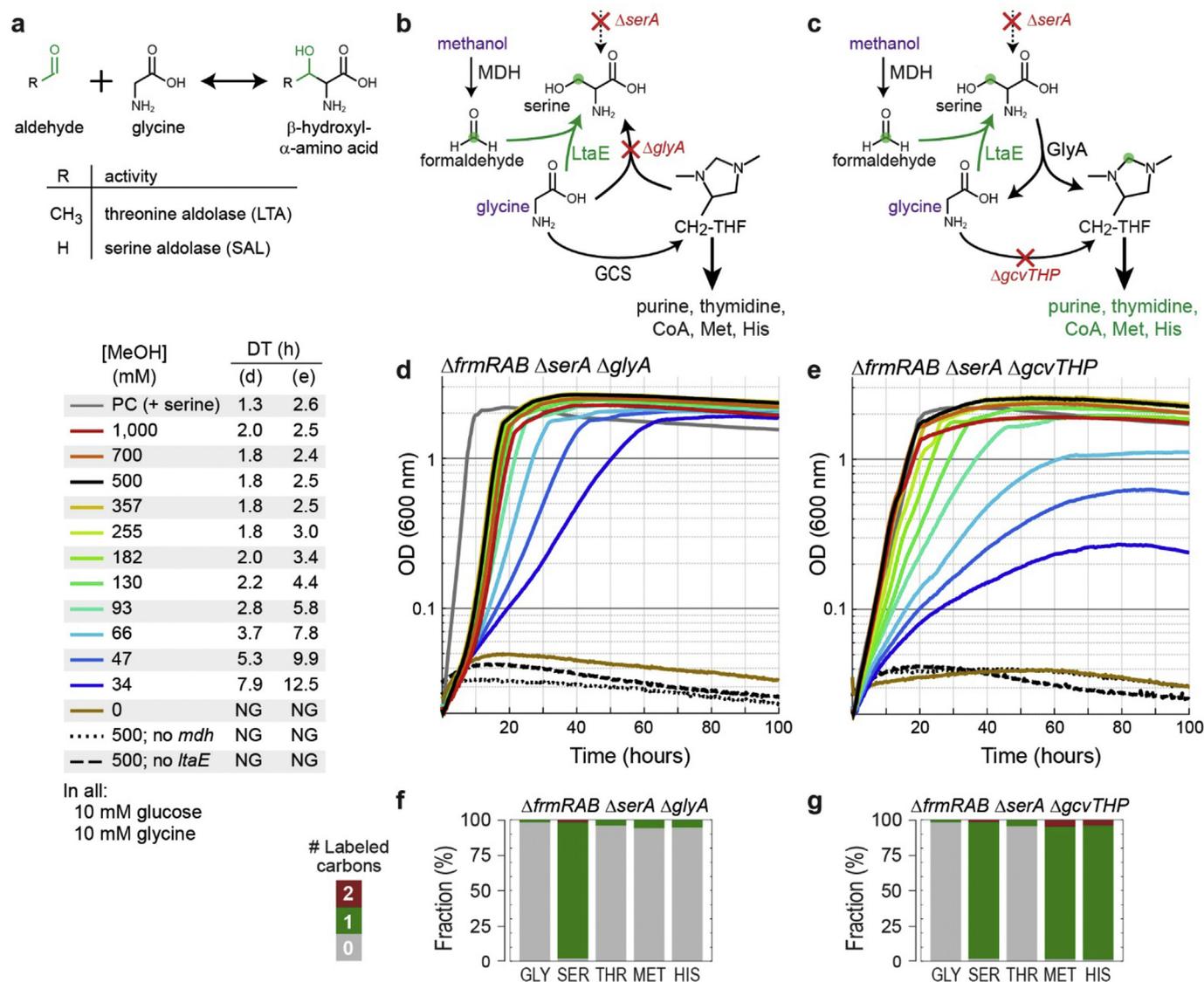
activities, the question remains whether these activities are high enough as to be of physiological relevance. In the rest of the paper, we explore the *in vivo* activities of these reactions.

### 2.3. Demonstration of the *in vivo* activity of serine aldolase

Most of the reactions of the homoserine cycle correspond to the primary activities of their catalyzing enzymes and thus are less likely to constrain pathway flux. For example, overexpression of serine deaminase, either on a plasmid or from the genome, enabled *E. coli* to use serine as a sole carbon and energy source (Supplementary Fig. S2). However, three pathway reactions – SAL, HAL, and HAT – correspond to promiscuous activities that, while characterized *in vitro*, might not be able to support physiologically relevant fluxes. Hence, we aimed to test each of these promiscuous activities *in vivo* within dedicated gene deletion strains, the growth of which is dependent on the activity of these reactions.

We started by testing the ability of LtaE to catalyze the SAL reaction *in vivo* (Fig. 3a). Towards this aim we constructed two strains auxotrophic for glycine and serine. In both strains the gene encoding for 3-phosphoglycerate dehydrogenase (*DserA*) was deleted. In one strain the gene encoding serine hydroxymethyltransferase was also deleted ( $\Delta$ *glyA*) while in the other strain the genes of the glycine cleavage system were deleted ( $\Delta$ *gcvTHP*). The growth of these strains required the addition of glycine and serine, as the cellular interconversion of these compounds is blocked (Fig. 3b,c).

We reasoned that if the SAL reaction indeed supports



**Fig. 3. *In vivo* serine aldolase activity catalyzed by LtaE.** (a) LtaE catalyzes aldol condensation between glycine and different aldehydes. Threonine aldolase (LTA) is its primary function, while serine aldolase is a promiscuous activity (Contestabile et al., 2001). (b), (c) Two selection schemes for the *in vivo* activity of the SAL reaction. Carbon sources are shown in purple (glucose not shown) while the formaldehyde moiety is shown in green. (b)  $\Delta frmRAB \Delta serA \Delta glyA$  strain in which methanol assimilation is required for the biosynthesis of serine. (c)  $\Delta frmRAB \Delta serA \Delta gcvTHP$  strain, in which methanol assimilation is required for the biosynthesis of serine and the cellular C<sub>1</sub> moieties. (d), (e) Growth with different concentrations of methanol confirm the activity of the SAL reaction of the strains shown in (b) and (c), respectively. In all cases, 10 mM glucose and 10 mM glycine were added to the medium. Each growth curve represents the average of three replicates, which differ from each other by less than 5%. “PC” corresponds to positive control. “DT” corresponds to doubling time. (f), (g) Labeling pattern of proteinogenic glycine (GLY), serine (SER), threonine (THR), methionine (MET) and histidine (HIS), within the strains shown in (b) and (c), respectively, upon feeding with <sup>13</sup>C-methanol as well as unlabeled glucose and glycine. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

physiologically relevant flux, both strains should be able to grow when methanol dehydrogenase (MDH) and LtaE are overexpressed and serine is replaced with methanol in the medium. In the  $\Delta serA \Delta glyA$  strain, the SAL reaction would be responsible for the production of serine (Fig. 3b), which accounts for ~3% of the carbon in biomass (Neidhardt et al., 1990). In the  $\Delta serA \Delta gcvTHP$  strain, the SAL reaction would be responsible for the production of both serine and the cellular C<sub>1</sub> moieties (Fig. 3c), together accounting for ~6% of the carbon in biomass (Neidhardt et al., 1990). To avoid formaldehyde oxidation to formate, which might deplete its intracellular pool and constrain its assimilation, the genes encoding for the glutathione-dependent formaldehyde oxidation system were also deleted ( $\Delta frmRAB$ ).

Upon overexpression of MDH and LtaE, we observed growth of both selection strains with glucose as the main carbon source and glycine and methanol as precursors of serine (Fig. 3d,e). This indicates that the

SAL reaction can operate *in vivo* at a physiologically significant rate. Expression of only MDH or only LtaE failed to sustain growth, indicating that the native expression of genomic *ltaE* is too low to support the SAL reaction. The observed growth rate and yield were dependent on the concentration of methanol, where the  $\Delta serA \Delta glyA \Delta frmRAB$  strain had higher rates and yields than the  $\Delta serA \Delta gcvTHP \Delta frmRAB$  strain on low methanol concentrations. This corresponds to our prediction that the latter strain depends on the SAL reaction to provide a higher fraction of the cellular carbons. Methanol concentrations in the range of 200–1000 mM seem to be optimal, supporting growth rates similar to that of the positive control (in which serine was added to the medium). In all experiments, we added 50  $\mu$ M MnCl<sub>2</sub>, as Mn<sup>2+</sup> is a known cofactor of LtaE (Fesko, 2016). Without the additional supplementation of MnCl<sub>2</sub>, we observed lower growth rates and yields (Supplementary Fig. S3).

To confirm the activity of the SAL reaction, we conducted  $^{13}\text{C}$ -labeling experiments. We cultivated both strains with  $^{13}\text{C}$ -methanol as well as unlabeled glucose and glycine. In the  $\Delta serA \Delta glyA \Delta frmRAB$  strain we found serine to be entirely singly labeled as expected, while the other amino acids were unlabeled (beyond the natural abundance of  $^{13}\text{C}$ , Fig. 3f). As threonine and methionine are derived from oxaloacetate, and hence carry a carbon that originates from  $\text{CO}_2$  (i.e., anaplerotic reactions), their lack of labeling indicates that formaldehyde oxidation to  $\text{CO}_2$  is negligible, as expected by the deletion of *frmRAB*. In the  $\Delta serA \Delta gcvTHP \Delta frmRAB$  strain we found serine, methionine, and histidine to be entirely singly labeled (Fig. 3g). Unlike threonine, both methionine and histidine harbor a carbon derived from THF carrying a  $\text{C}_1$  unit – methyl-THF in the case of methionine and formyl-THF in the case of histidine. The labeling of these amino acids thus indicates that all cellular  $\text{C}_1$  moieties are derived from methanol. Overall, the labeling results confirm that the SAL reaction provides the sole source of serine in both strains and sole source of  $\text{C}_1$  moieties in the  $\Delta serA \Delta gcvTHP \Delta frmRAB$  strain.

Next, we tested whether it is possible to omit glycine from the medium, such that it will be produced endogenously via LtaE-dependent threonine cleavage (Fig. 4a). Towards this aim, we deleted, in both selection strains, the genes encoding for threonine dehydrogenase and 2-amino-3-ketobutyrate CoA-ligase ( $\Delta kbl\text{-}tdh$ ), thus blocking the LtaE-independent route of threonine degradation (Fig. 4a) (Yishai et al., 2017). We found that replacing glycine with threonine did not alter the growth of either selection strain (Fig. 4b, note that growth was still strictly dependent on methanol). To enable this growth, the overexpressed LtaE catalyzes two subsequent reactions, first cleaving threonine to glycine and acetaldehyde and then reacting glycine with formaldehyde to produce serine. Hence, LtaE can be regarded as a glycytransferase – transferring a glycine moiety from one small aldehyde (acetaldehyde) to another (formaldehyde).

We wondered whether it was also possible to omit threonine from the medium and rely on native threonine biosynthesis to provide this amino acid as a precursor for glycine and serine (Fig. 4a). Indeed, despite showing reduced growth rates, both selection strains were able to grow with only glucose (as main carbon source) and methanol without the addition of glycine or threonine (green lines in Fig. 4b). This indicates that half of the homoserine cycle is active: homoserine (generated natively from aspartate) is metabolized to threonine via HSK and TS, and LtaE then cleaves threonine to glycine (and acetaldehyde) and condenses glycine with formaldehyde to produce serine (Fig. 4a).

To confirm that, also in the absence of externally provided glycine or threonine, all cellular serine is produced from glycine condensation with formaldehyde, we conducted  $^{13}\text{C}$ -labeling experiments. The strains were cultured in the presence of labeled or unlabeled methanol as well as glucose labeled at different carbons (glucose-1- $^{13}\text{C}$ , glucose-2- $^{13}\text{C}$ , and glucose-3- $^{13}\text{C}$ ). While the labeling pattern of glycine changed according to the labeled carbon of glucose, cultivation with  $^{13}\text{C}$ -methanol always resulted in exactly one more labeled carbon in serine than in glycine (Fig. 4c). This unequivocally confirms the methanol-dependent production of serine from glycine when the latter compound is produced internally from homoserine metabolism.

Our results demonstrate the capability of LtaE to convert threonine to serine *in vivo* by releasing acetaldehyde and assimilating formaldehyde. These findings further confirm the physiologically relevant activity of half of the homoserine cycle, where homoserine metabolism to glycine and serine provided all the biomass requirement of these amino acids as well as cellular  $\text{C}_1$  moieties.

#### 2.4. Demonstration of the *in vivo* activity of HOB aldolase and transaminase

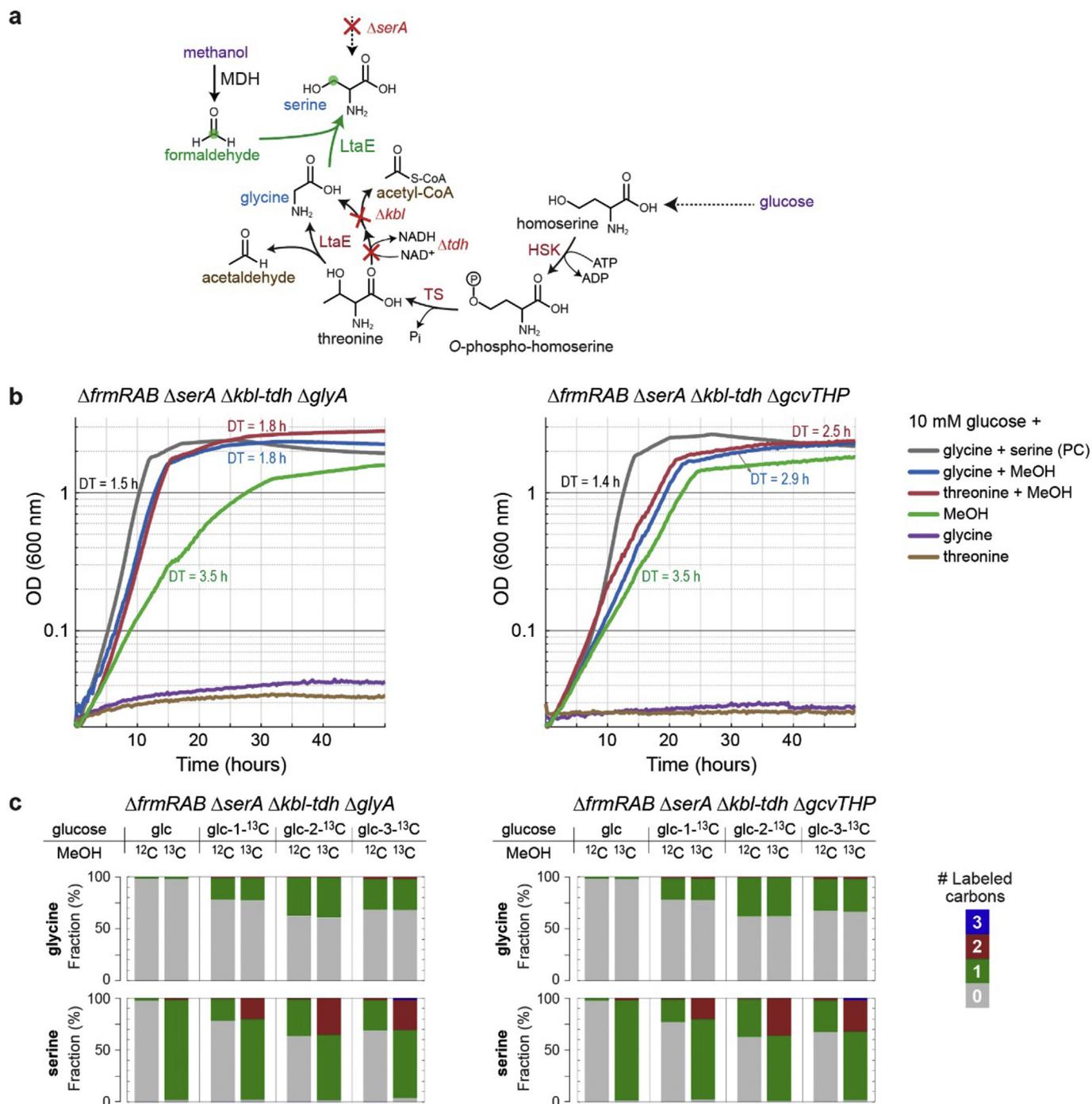
After demonstrating the methanol-dependent conversion of homoserine to serine, we aimed to demonstrate methanol-dependent conversion of pyruvate to homoserine. To select for the *in vivo* conversion

of pyruvate to homoserine and threonine via HOB production and amination, we constructed a homoserine auxotroph strain: a deletion of the gene encoding for aspartate-semialdehyde dehydrogenase ( $\Delta asd$ ) resulted in a strain capable of growing only when homoserine and diamino-pimelate (DAP) were added to the medium (Cardineau and Curtiss, 1987). In this strain, homoserine is metabolized to methionine, threonine, and isoleucine, while DAP is metabolized to lysine and peptidoglycans. (We note that despite being formally reversible, homoserine dehydrogenase was not able to oxidize homoserine to aspartate-semialdehyde, the precursor of DAP, and hence the addition of the latter intermediate to the medium was required).

We reasoned that, in the presence of methanol and methanol dehydrogenase, the combined activities of HAL and HAT should enable the  $\Delta asd \Delta frmRAB$  strain to grow without the addition of homoserine to the medium (Fig. 5a). While only few enzymes were previously shown to catalyze the HAL reaction (Hernandez et al., 2017; Wang et al., 2019), we speculated that similar aldolases could also support this activity, maybe even outperforming those identified before. Hence, we searched for all *E. coli* enzymes (strain MG1655, using EcoCyc (Keseler et al., 2005)) that are known to catalyze an aldolase reaction with pyruvate as a donor (and which might be able to use formaldehyde as an acceptor). Besides RhmA itself (Hernandez et al., 2017), we found six candidate aldolases: GarL, YagE, YjhH, Eda, DgoA, and MhpE (Fig. 5b). RhmA and GarL belong to the structural family of HpcH while mhpE belongs to DmpG family. Both of these families are Type II pyruvate aldolases which use a divalent metal cation for donor binding and enolization (Fang et al., 2019). YagE and YjhH belong to the structural family of DHDPs while Eda and DgoA belong to KDPG family. These families are Type I pyruvate aldolases, using a lysine residue to form a Schiff base with the donor substrate (Fang et al., 2019). We decided to test all of the seven aldolases for their ability to catalyze the HAL reaction *in vivo*. As the HAT reaction is known to be supported by the native aspartate aminotransferase (AspC) (Zhong et al., 2019), a highly expressed protein (Li et al., 2014), and might be further catalyzed by other highly expressed, promiscuous aminotransferases, we hypothesized that no dedicated enzyme overexpression would be required to enable this key reaction.

We found that overexpression of *mdh* together with *rhmA*, *garL*, *yagE*, or *yjhH* enabled growth of the  $\Delta asd \Delta frmRAB$  strain when homoserine was replaced with methanol (Fig. 5c). These four aldolase enzymes supported roughly the same growth rates. No growth was observed without methanol, or when methanol dehydrogenase or the aldolase enzymes were overexpressed alone. Growth was not observed when *mdh* was expressed together with *eda*, *dgoA*, or *mhpE*. The relative sequence similarity between RhmA, GarL, YagE, and YjhH (Fig. 5b and Supplementary Fig. S4) might explain why these enzymes, and not the others, were able to support the HAL reaction. Indeed, the structures of RhmA and GarL are almost identical (Supplementary Fig. S5). As anticipated, growth was possible without dedicated overexpression of an aminotransferase enzyme, presumably since the native high expression level of these enzymes is sufficient to support the promiscuous HAT reaction. Genomic overexpression of *aspC* or a mutated version of alanine aminotransferase (*alaC\**, the protein product of which was previously shown to catalyze the HAT reaction (Bouzon et al., 2017)) did not alter growth substantially (Supplementary Fig. S6). Similarly, genomic overexpression of *thrBC* did not consistently assist growth (Supplementary Fig. S6). This indicates that the HAT, HSK and TS reactions do not constrain the flux from pyruvate to threonine.

To confirm that homoserine and its downstream products threonine and methionine are produced from pyruvate and methanol via the HAL and HAT reactions, we performed  $^{13}\text{C}$ -labeling experiments. Upon cultivation with unlabeled glucose and  $^{13}\text{C}$ -methanol, we found threonine and methionine to be completely once labeled, where lysine and aspartate (serving as control) were fully unlabeled (beyond the natural abundance of  $^{13}\text{C}$ , Fig. 5d). This confirms that homoserine and threonine are completely derived from pyruvate and methanol.



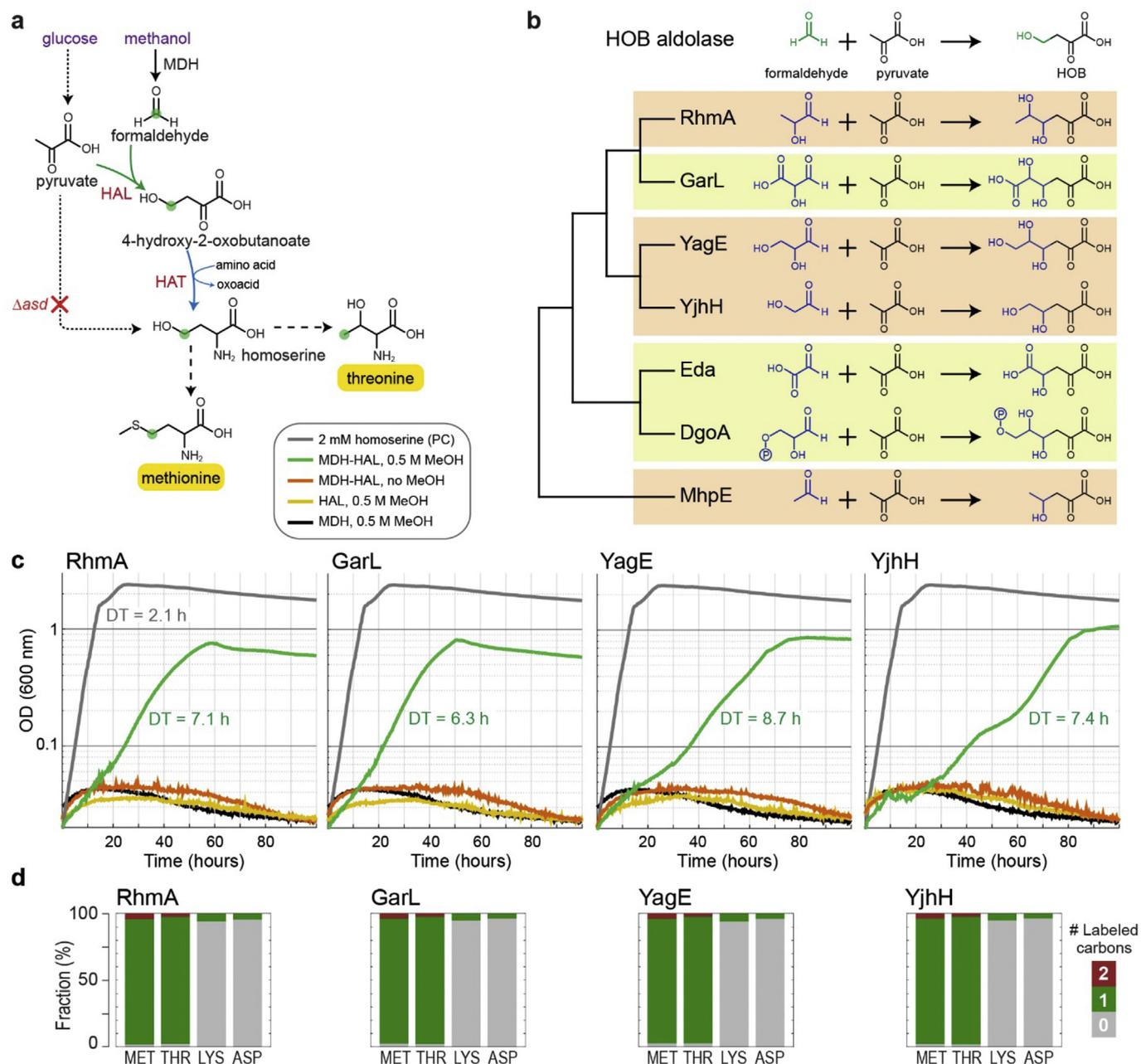
**Fig. 4.** LtaE can operate as threonine aldolase and serine aldolase simultaneously. (a) Selection scheme for serine and glycine production from threonine and homoserine. (b) The two selection strains, deleted in the LtaE-independent threonine cleavage system ( $\Delta kbl\text{-}tdh$ ), can grow with methanol as serine precursor. Glycine is either provided externally (10 mM, blue lines) or produced internally, either from externally added threonine (10 mM, red lines) or from the internal pool of threonine (green lines). The latter growth confirms that LtaE can catalyze the LTA and SAL reactions simultaneously. In all cases, 10 mM glucose and 500 mM methanol were added. Each growth curve represents the average of three replicates, which differ from each other by less than 5%. “PC” corresponds to positive control. “DT” corresponds to doubling time. (c) Labeling pattern of proteinogenic glycine and serine within the strains shown in (b) upon feeding with glucose (glc) labeled at different carbon as well as labeled or unlabeled methanol. This labeling confirms that all cellular serine is produced from glycine and methanol even when glycine is produced internally from threonine biosynthesis and degradation. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Our experiments thus demonstrate the capacity of several *E. coli* endogenous aldolases to generate HOB *in vivo* and the capacity of the native aminotransferase network to convert HOB into homoserine. These findings further confirm the physiologically relevant activity of the “other half” of the homoserine cycle, where pyruvate metabolism to homoserine and threonine provided the biomass requirement of these

amino acids.

### 3. Discussion

Different metabolic engineering strategies can be used to establish novel modes of growth and production. In the simplest approach, a



**Fig. 5.** *In vivo* HOB aldolase and transaminase activities. (a) A selection scheme for the *in vivo* activity of the HAL and the HAT reactions. Carbon sources are shown in purple while the formaldehyde moiety is shown in green. (b) Several *E. coli* enzymes are known to catalyze an aldolase reaction with pyruvate as a donor, and might be able to react with formaldehyde as an acceptor. The sequence similarity of these enzymes is indicated by the schematic tree to the left. YagE and YjhH can catalyze each others reactions. Similarly, Eda and DgoA can catalyze each others reactions. (c) Four aldolases, once overexpressed together with methanol dehydrogenase, support growth of the selection strain. Glucose was added at 10 mM, methanol at 500 mM, diaminopimelate at 0.25 mM, and isoleucine at 1 mM. Each growth curve represents the average of three replicates, which differ from each other by less than 5%. “PC” corresponds to positive control. “DT” corresponds to doubling time. (d) Labeling pattern of proteinogenic methionine (MET), threonine (THR), lysine (LYS), and aspartate (ASP) within the strains shown in (c) upon feeding with unlabeled glucose, diaminopimelate and isoleucine as well as  $^{13}\text{C}$ -methanol. The results confirm that all cellular threonine and methionine are derived from HOB aldolase and transaminase reactions. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

metabolic pathway is “copied” from one organism and “pasted” in another, where pathway enzymes are expressed in the new host (Erb et al., 2017). Most studies that try to establish synthetic methylotrophy in heterotrophic hosts follow this strategy by expressing the enzymes of the known methanol assimilation pathways in *E. coli* (Chen et al., 2018; He et al., 2018; Meyer et al., 2018) and other biotechnological organisms (Lessmeier et al., 2015; Withhoff et al., 2015). However, this approach is limited to a small number of possible routes, which might not represent an optimal solution (e.g., in terms of resource consumption

and biomass yield) and might not interact favorably with the endogenous metabolism of the host (e.g., disrupt key cellular fluxes, necessitate complex regulation, or introduce reactive metabolic intermediates).

Alternatively, a mix-and-match of existing enzymes could expand the space of possible pathways: while all the components are known, the combinatorics of their integration can be quite extensive (Erb et al., 2017). The design of the modified serine cycle exemplifies this approach, where key reactions of the natural pathway were replaced with

other known reactions in order to optimize pathway activity in the host (Yu and Liao, 2018). While this approach is considerably more flexible than “copy-paste”, it is still limited by the repertoire of known enzymes and reactions.

In order to fully harness the power of synthetic biology and identify the most promising pathways, it is beneficial to consider also reactions which, while chemically plausible, are not known to be enzymatically catalyzed (Bouzon et al., 2017; Erb et al., 2017). A recent study has demonstrated this approach for the design and *in vitro* realization of synthetic photorespiration pathways which do not release CO<sub>2</sub> and could thus boost photosynthetic carbon fixation (Trudeau et al., 2018). Following such guidelines, previous studies designed novel methanol assimilation pathways, based on direct condensation of formaldehyde molecules (Chou et al., 2019; Lu et al., 2019; Wang et al., 2017; Yang et al., 2019); however, the rates of most of the key reactions remain too low to be of physiological significance.

In this study, we use a similar strategy to identify an optimal methanol assimilation design based on the general structure of the serine cycle. Instead of relying on completely novel reactions, the establishment of which could be highly challenging, we decided to follow a middle-ground approach, where besides known enzymatic reactions we considered characterized promiscuous activities for which at least basic supporting evidence exists.

We identified the homoserine cycle as a superior pathway design. It is short and relies solely on *E. coli* native enzymes. Still, its reactions are “simple” enough to be catalyzed by native enzymes of many different hosts. The pathway does not overlap with central metabolism, thus bypassing possible regulation difficulties. As its consumption of ATP and NAD(P)H is low, it is expected to support high biomass yield. Importantly, by replacing a carboxylation reaction with another formaldehyde condensation reaction, the homoserine cycle should be able to operate effectively without oxidizing formaldehyde, thus avoiding the difficulty of regulating a flux bifurcation point in which formaldehyde (and formate) can either be fully oxidized or assimilated to biomass. We further showed that the homoserine cycle has the potential to outperform even the RuMP cycle for the production of a wide array of chemicals which are derived from acetyl-CoA. Moreover, whereas the synthetic pathway is highly thermodynamically favorable (MDF > 12 kJ/mol), the RuMP cycle is only moderately favorable (MDF ~ 4 kJ/mol). Overall, the homoserine cycle seems to display clear advantages when compared to all of its natural counterparts.

By using dedicated auxotrophic strains we were able to demonstrate that the promiscuous enzyme activities required for the homoserine cycle are sufficiently high to enable the production of biomass building blocks. Quite surprisingly, we identified four native aldolase enzymes that can catalyze the HAL reaction at a sufficient rate to enable growth of the auxotrophic strain. While dedicated overexpression of the promiscuous aldolase enzymes was necessary to establish the required activity, the needed aminotransferase reaction was supported by endogenously expressed aminotransferases.

The findings of this study provide a firm basis for future experiments towards the establishment of methylotrophic growth via the full homoserine cycle. Protein engineering and evolution can be applied to improve the activity of the promiscuous enzymes and thus enhance pathway kinetics. Specifically, while LtaE activity, which generates 10% of cellular carbon, supports a relatively high growth rate, all aldolases catalyzing the HAL reaction support a relatively low growth rate despite providing only 4% of cellular carbon. Hence, engineering these aldolases for better kinetics with formaldehyde would be highly useful.

The establishment of the activity of the full homoserine cycle would probably require long term adaptive evolution under selective conditions (Gresham and Hong, 2015). The auxotrophic strains used in this study would be useful for such evolution, as formaldehyde assimilation via the aldolase reactions must avoid the biosynthesis of serine and homoserine via other routes which would thermodynamically and

kinetically “push” the aldolases in the reverse direction (i.e., formaldehyde-producing). For example, a strain auxotrophic to both serine and homoserine, overexpressing methanol dehydrogenase and both aldolases, could be first cultivated on a limiting amount of glycine and saturating amount of pyruvate and methanol. This would select for the emergence of a strain that could grow without glycine, thus combining the activities of both aldolases. The resulting strain can be then cultivated on a limiting amount of pyruvate and saturating amount of methanol, until a strain capable of growing on methanol as a sole carbon source emerges. For such growth on methanol to arise, a delicate balance between the metabolic fluxes within the homoserine cycle and those that converge to and diverge from the pathway would have to evolve. For example, the fluxes that generate pyruvate – from serine deamination and acetyl-CoA assimilation – would have to balance the fluxes that consume pyruvate, both within the homoserine cycle as well as towards other cellular routes – for example, pyruvate oxidation to acetyl-CoA. Such an evolutionary approach was previously shown to be successful for the implementation of the Calvin Cycle in *E. coli*, where the fluxes that diverge from the cycle were downregulated to match the fluxes within the cycle (Antonovsky et al., 2016; Gleizer et al., 2019). While achieving *E. coli* growth via the homoserine cycle is undoubtedly a challenging task, it holds the promise of vast new production opportunities for the bioindustry.

## 4. Methods

### 4.1. Max-min driving force (MDF) analysis

MDF analysis (Noor et al., 2014) was applied to evaluate the thermodynamic feasibility of the homoserine cycle for acetyl-CoA production from formaldehyde. The natural serine cycle and the modified serine cycle (Yu and Liao, 2018) were also analyzed for comparison. eQuilibrator-API (Noor et al., 2013) was used for these analyses. Metabolite concentrations were constrained as described before (Noor et al., 2014) with two changes: (i) formaldehyde upper bound was set to 0.5 mM according to highest concentration *E. coli* can tolerate (He et al., 2018); (ii) glutamate and 2-oxoglutarate were set as cofactors with concentrations of 100 mM and 0.5 mM, respectively (Bennett et al., 2009). pH was assumed to be 7.5 (as *E. coli* cytoplasm) and ionic strength was assumed to be 0.25 M (as recommended (Alberly et al., 2011)). Since energetic calculation using C<sub>1</sub>-tetrahydromethanopterin intermediates are not supported by eQuilibrator-API, we used formate as substrate for serine cycle and the modified serine cycle. The scripts and details can be found at <https://gitlab.com/hi-he/wenzhangfujian> within the “2020\_Promiscuous aldolases” directory.

### 4.2. Flux balance analysis

Flux balance analysis was conducted in Python with COBRApy (Ebrahim et al., 2013). New reactions were added to the most updated *E. coli* genome-scale metabolic network iML1515 (Monk et al., 2017) with several curations and changes: (i) transhydrogenase (THD2pp) translocates one proton instead of two (Bizouarn et al., 2005); (ii) homoserine dehydrogenase (HSDy) produces homoserine from aspartate-semialdehyde irreversibly (as we found experimentally, see main text); (iii) GLYCK (glycerate-3P producing glycerate kinase) and POR5 (pyruvate synthase) were removed from the model as their existence in *E. coli* is highly disputable; (iv) since we introduced NAD-dependent formate dehydrogenase, the two quinone-dependent formate dehydrogenase, FDH4pp and FDH5pp, were removed from the model; (v) pyruvate formate lyase (PFL) and 2-oxobutanoate formate lyase (OBTF) were removed from the model as they operate only under anaerobic condition. We further removed the ATP maintenance reaction (ATPM) as, rather than estimating growth rate, we used FBA to estimate the maximal yield. Biomass yield was calculated by the predicted maximal biomass reaction flux divided by the flux of methanol

**Table 1**  
List of *E. coli* strains and plasmids used in the study.

Strain	Genotype	Source
MG1655	K-12 F <sup>-</sup> λ <sup>-</sup> <i>ibvG<sup>-</sup> rfb-50 rph-1</i>	Lab collection
SIJ488	MG1655 Tn7::para-exo-beta-gam; prha-FLP; xylSpm-IsceI	(Jensen et al., 2015)
DH5α	F <sup>-</sup> <i>endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20 φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(r<sub>K</sub> m<sub>K</sub><sup>+</sup>), λ</i>	Lab collection
	SIJ488 Δ <i>frmRAB ΔserA ΔglyA</i>	This study
	SIJ488 Δ <i>frmRAB ΔserA ΔgcvTHP</i>	This study
	SIJ488 Δ <i>frmRAB ΔserA ΔglyA Δkbl-tdh</i>	This study
	SIJ488 Δ <i>frmRAB ΔserA ΔgcvTHP Δkbl-tdh</i>	This study
	SIJ488 Δ <i>frmRAB Δasd</i>	This study
PthrBC_D	SIJ488 Δ <i>P<sub>thrABC</sub>::CAP-P<sub>pgi-20</sub>-thrBC</i>	This study
PalaC_D	SIJ488 Δ <i>P<sub>alaC</sub>::CAP-P<sub>pgi-20</sub>-alaC*</i>	This study
PaspC_D	SIJ488 Δ <i>P<sub>aspC</sub>::CAP-P<sub>pgi-20</sub>-aspC</i>	This study
	SIJ488 Δ <i>frmRAB Δasd ΔP<sub>thrABC</sub>::P<sub>pgi-20</sub>-thrBC</i>	This study
	SIJ488 Δ <i>frmRAB Δasd ΔP<sub>alaC</sub>::P<sub>pgi-20</sub>-alaC*</i>	This study
	SIJ488 Δ <i>frmRAB Δasd ΔP<sub>aspC</sub>::P<sub>pgi-20</sub>-aspC</i>	This study
	SIJ488 Δ <i>frmRAB Δasd ΔP<sub>thrABC</sub>::P<sub>pgi-20</sub>-thrBC ΔP<sub>alaC</sub>::P<sub>pgi-20</sub>-alaC*</i>	This study
	SIJ488 Δ <i>frmRAB Δasd ΔP<sub>thrABC</sub>::P<sub>pgi-20</sub>-thrBC ΔP<sub>aspC</sub>::P<sub>pgi-20</sub>-aspC</i>	This study
Plasmid	Genes	Source
pZASS	p15A ori; Strep <sup>R</sup> ; P <sub>pgi-20</sub>	Lab collection
pZASS-mdh	pZASS::C <i>gadhA</i>	(He et al., 2018)
pZASS-ltaE-mdh	pZASS::ltaE, C <i>gadhA</i>	This study
pZASS-ltaE-rhmA	pZASS::ltaE, rhmA	This study
pZASS-mdh-rhmA	pZASS::C <i>gadhA</i> , rhmA	This study
pZASS-garL	pZASS::garL	This study
pZASS-mdh-yagE	pZASS::C <i>gadhA</i> , yagE	This study
pZASS-yagE	pZASS::yagE	This study
pZASS-mdh-yjhH	pZASS::C <i>gadhA</i> , yjhH	This study
pZASS-yjhH	pZASS::yjhH	This study
pZASS-mdh-eda	pZASS::C <i>gadhA</i> , eda	This study
pZASS-eda	pZASS::eda	This study
pZASS-mdh-dgoA	pZASS::C <i>gadhA</i> , dgoA	This study
pZASS-dgoA	pZASS::dgoA	This study
pZASS-mdh-mhpE	pZASS::C <i>gadhA</i> , mhpE	This study
pZASS-mhpE	pZASS::mhpE	This study
pZASS-sdaA	pZASS::sdaA	This study

uptake. While modeling acetyl-CoA production, two unrealistic pathways were blocked by deleting reactions DRPA and PAI2T respectively. The full code, including changes to the model, reactions specific to the methanol assimilation pathways, and the reactions of each production route can be found at <https://gitlab.com/hi-he/wenzhangfujian> within the “2020\_Promiscuous aldolases” directory.

#### 4.3. Strains and genomic manipulation

Strains used in this study are listed in Table 1. An *E. coli* MG1655 derived strain SIJ488 (Jensen et al., 2015) was used as the parental strain for genomic modifications. Iterative rounds of λ-Red recombineering (Jensen et al., 2015) or P1 phage transduction (Thomason et al., 2007) were used for gene deletions. For the recombineering, selectable resistance cassettes were generated via PCR – primers 50 bp homologous arms as in (Baba et al., 2006) – using the FRT-PGK-gb2-neo-FRT (Km) cassette (Gene Bridges, Germany) for

kanamycin resistance (Km) and the pKD3 plasmid (GenBank: AY048742) (Datsenko and Wanner, 2000) as a template for chloramphenicol resistance cassettes (CAP). The procedures of the deletion, verification and antibiotic cassette removal are detailed in (Wenk et al., 2018).

A similar strategy was applied to exchange the genomic promoter of target genes. A constitutive strong promoter *pgi-20* (*P<sub>pgi-20</sub>*) (Braatsch et al., 2008) and a ribosome binding site “C” (AAGTTAAGAGGCAAGA) (Zelcbuch et al., 2013) were constructed downstream of the CAP cassette using primers listed shown in Supplementary Table S2. The synthetic promoter was first introduced to the SIJ488 strain by the recombineering method; P1 transduction was then used to transfer the synthetic promoter into the selection strains. *thrB* (encoding homoserine kinase, HSK) and *thrC* (encoding threonine synthase, TS) are on the same operon with *thrL* and *thrA*. Since *thrL* encodes regulatory peptide and *thrA* is redundant in the Δ*asd* selection strains, *thrLA* was deleted during the promoter exchange of *thrBC*. The point mutations A142P Y275D (Bouzon et al., 2017) were introduced along with the promoter exchange of *alaC* (In this case, the recombineering cassette has the mutated gene downstream the CAP cassette and synthetic promoter). Promoter exchanges were confirmed by sequencing the promoter regions.

#### 4.4. Plasmids construction

All cloning procedures were carried out in *E. coli* DH5α strain. *E. coli* native genes *ltaE*, *rhmA*, *garL*, *yagE*, *yjhH*, *eda*, *dgoA*, *mhpE*, and *SdaA* were cloned from *E. coli* MG1655 genome with the primers shown in Supplementary Table S2. NAD-dependent methanol dehydrogenase (CgAdhA) was taken from *Corynebacterium glutamicum* R after codon optimization (He et al., 2018). Genes were inserted into a pNivC vector downstream of a ribosome binding site “C” (AAGTTAAGAGGCAAGA) (Zelcbuch et al., 2013). The genes were assembled into one operon using BioBrick enzymes: *BcuI*, *Sall*, *NheI* and *XhoI* (FastDigest, Thermo Scientific) (Zelcbuch et al., 2013). Using *EcoRI* and *PstI*, the synthetic operon was then inserted into an overexpression pZASS vector (Wenk et al., 2018) under a constitutive strong promoter *pgi-20* (Braatsch et al., 2008). The final plasmids are listed in Table 1.

#### 4.5. Growth media

LB medium (0.5% yeast extract, 1% tryptone, 1% NaCl) was used for strain engineering and recombinant plasmids cloning. Antibiotics were used at the following concentrations: kanamycin, 50 μg/mL; ampicillin, 100 μg/mL; streptomycin, 100 μg/mL; chloramphenicol, 30 μg/mL. Growth experiments were performed in M9 minimal media (47.8 mM Na<sub>2</sub>HPO<sub>4</sub>, 22 mM KH<sub>2</sub>PO<sub>4</sub>, 8.6 mM NaCl, 18.7 mM NH<sub>4</sub>Cl, 2 mM MgSO<sub>4</sub> and 100 μM CaCl<sub>2</sub>), supplemented with trace elements (134 μM EDTA, 31 μM FeCl<sub>3</sub>, 6.2 μM ZnCl<sub>2</sub>, 0.76 μM CuCl<sub>2</sub>, 0.42 μM CoCl<sub>2</sub>, 1.62 μM H<sub>3</sub>BO<sub>3</sub>, 0.081 μM MnCl<sub>2</sub>). Additional 50 μM MnCl<sub>2</sub> was added for all experiments since it improves *in vivo* activity of LtaE (Supplementary Fig. S3). Carbon sources were added according to the strain and the specific experiment (if not stated otherwise in the text): 10 mM glucose, 500 mM methanol, 10 mM glycine, 10 mM serine, 10 mM threonine, 2 mM homoserine, and 1 mM isoleucine. 0.25 mM diaminopimelate (DAP) was supplemented in all media used to cultivate the Δ*asd* strain (Cardineau and Curtiss, 1987).

#### 4.6. Growth experiments

Strains were precultured in 4 mL M9 medium with proper carbon sources and streptomycin. The precultures were harvested and washed three times in M9 medium, then inoculated in M9 media with suitable carbon sources, with a starting OD<sub>600</sub> of 0.02. 150 μL of culture were added to each well of 96-well microplates (Nunclon Delta Surface, Thermo Scientific). Further 50 μL mineral oil (Sigma-Aldrich) was

added to each well to avoid evaporation (while enabling gas diffusion). The 96-well microplates were incubated at 37°C in microplate reader (BioTek EPOCH 2). The shaking program cycle (controlled by Gen5 v3) had 4 shaking phases, lasting 60 s each: linear shaking followed by orbital shaking, both at an amplitude of 3 mm, then linear shaking followed by orbital shaking both at an amplitude of 2 mm. The absorbance (OD<sub>600</sub>) in each well was monitored and recorded after every three shaking cycles (~16.5 min). Raw data from the plate reader were calibrated to normal cuvette measured OD<sub>600</sub> values according to OD<sub>cuvette</sub> = OD<sub>plate</sub>/0.23. Growth parameters were calculated using MATLAB (MathWorks) based on three technical triplicates – the average values were used to generate the growth curves. Checked in MATLAB, in all cases variability between triplicates measurements were less than 5%.

#### 4.7. Stable isotopic labeling

<sup>13</sup>C-Methanol, glucose-1-<sup>13</sup>C, glucose-2-<sup>13</sup>C, glucose-3-<sup>13</sup>C were purchased from Sigma-Aldrich. Strains were cultivated aerobically in glass tubes on M9 minimal media with the appropriate carbon sources at the concentrations mentioned above (section 4.5: Growth media). Experiments were performed in duplicates which in all cases showed identical results (± 5%). Cells were harvested at the late exponential phase. The equivalent volume of 1 mL of culture at OD<sub>600</sub> of 1 was harvested and washed by centrifugation. Protein biomass was hydrolyzed with 6 M HCl, at 95°C for 24 h (You et al., 2012). The samples were completely dried under a stream of air at 95°C. Hydrolyzed amino acids were analyzed with UPLC–ESI–MS as previously described (Gialvalisco et al., 2011). Chromatography was performed with a Waters Acquity UPLC system (Waters), using an HSS T3 C<sub>18</sub> reversed phase column (100 mm × 2.1 mm, 1.8 μm; Waters). 0.1% formic acid in H<sub>2</sub>O (A) and 0.1% formic acid in acetonitrile (B) were the mobile phases. The flow rate was 0.4 mL/min and the gradient was: 0 to 1 min – 99% A; 1 to 5 min – linear gradient from 99% A to 82%; 5 to 6 min – linear gradient from 82% A to 1% A; 6 to 8 min – kept at 1% A; 8–8.5 min – linear gradient to 99% A; 8.5–11 min – re-equilibrate. Mass spectra were acquired using an Exactive mass spectrometer (Thermo Scientific) in positive ionization mode, with a scan range of 50.0 to 300.0 m/z. The spectra were recorded during the first 5 min of the LC gradients. Data analysis was performed using Xcalibur (Thermo Scientific). The identification of amino acids was based on retention times and m/z, which were determined by analyzing amino acid standards (Sigma-Aldrich) under the same conditions.

#### 4.8. Molecular phylogenetic analysis

The protein sequences of the aldolases predicted to catalyze the HAL reaction were obtained from UniProt: RhmA/YfaU P76469, GarL P23522, YagE P75682, YjhH P39359, Eda P0A955, DgoA Q6BF16 and MhpE P51020. MAFFT v7 (Katoh and Standley, 2013) was used for multiple sequence alignment with default parameters. The aligned sequences were used by MEGA X (Kumar et al., 2018) with Maximum Likelihood method to construct a phylogenetic tree. The bootstrap consensus tree was generated with the setting No. of bootstrap replications to 1000.

#### CRediT authorship contribution statement

**Hai He:** Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing - original draft, Writing - review & editing. **Rune Höper:** Investigation, Methodology. **Moritz Dodenhöft:** Investigation, Methodology. **Philippe Marlière:** Conceptualization, Writing - review & editing. **Arren Bar-Even:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Project administration, Resources, Supervision, Validation, Writing - original draft, Writing - review & editing.

#### Declaration of competing interest

A.B.E. is co-founder of b.fab, aiming to commercialize microbial growth of C<sub>1</sub> compounds. b.fab was not involved in this study and did not fund it.

#### Acknowledgements

The authors thanks Anne Michaelis and Pauline Kuschel for experimental assistance, Elad Noor for help in the MDF analysis, and Charlie Cotton, Seohyoung Kim, Steffen Lindner, Nico Claassens, and Andrew Hanson for critical reading of the manuscript. This study was funded by the Max Planck Society. Hai He is funded by the China Scholarship Council.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ymben.2020.03.002>.

#### References

- Alberly, R.A., Cornish-Bowden, A., Goldberg, R.N., Hammes, G.G., Tipton, K., Westerhoff, H.V., 2011. Recommendations for terminology and databases for biochemical thermodynamics. *Biophys. Chem.* 155, 89–103. <https://doi.org/10.1016/j.bpc.2011.03.007>.
- Antonovsky, N., Gleizer, S., Noor, E., Zohar, Y., Herz, E., Barenholz, U., Zelcbuch, L., Amram, S., Wides, A., Tepper, N., Davidi, D., Bar-On, Y., Bareia, T., Wernick, D.G., Shani, I., Malitsky, S., Jona, G., Bar-Even, A., Milo, R., 2016. Sugar synthesis from CO<sub>2</sub> in *Escherichia coli*. *Cell* 166, 115–125. <https://doi.org/10.1016/j.cell.2016.05.064>.
- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., Mori, H., 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* 2, 2006–2008. <https://doi.org/10.1038/msb4100050>.
- Bar-Even, A., 2016. Formate assimilation: the metabolic architecture of natural and synthetic pathways. *Biochemistry* 55, 3851–3863. <https://doi.org/10.1021/acs.biochem.6b00495>.
- Bar-Even, A., Flamholz, A., Noor, E., Milo, R., 2012. Thermodynamic constraints shape the structure of carbon fixation pathways. *Biochim. Biophys. Acta* 1817, 1646–1659. <https://doi.org/10.1016/j.bbabi.2012.05.002>.
- Bennett, B.D., Kimball, E.H., Gao, M., Osterhout, R., Van Dien, S.J., Rabinowitz, J.D., 2009. Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat. Chem. Biol.* 5, 593–599. <https://doi.org/10.1038/nchembio.186>.
- Bizouarn, T., van Boxel, G.I., Bhakta, T., Jackson, J.B., 2005. Nucleotide binding affinities of the intact proton-translocating transhydrogenase from *Escherichia coli*. *Biochim. Biophys. Acta* 1708, 404–410. <https://doi.org/10.1016/j.bbabi.2005.04.004>.
- Bogorad, I.W., Chen, C.T., Theisen, M.K., Wu, T.Y., Schlenz, A.R., Lam, A.T., Liao, J.C., 2014. Building carbon-carbon bonds using a biocatalytic methanol condensation cycle. *Proc. Natl. Acad. Sci. U. S. A.* 111, 15928–15933. <https://doi.org/10.1073/pnas.1413470111>.
- Bouzon, M., Perret, A., Loreau, O., Delmas, V., Perchat, N., Weissenbach, J., Taran, F., Marlière, P., 2017. A synthetic alternative to canonical one-carbon metabolism. *ACS Synth. Biol.* 6, 1520–1533. <https://doi.org/10.1021/acssynbio.7b00029>.
- Braatsch, S., Helmark, S., Kranz, H., Koebmann, B., Jensen, P.R., 2008. *Escherichia coli* strains with promoter libraries constructed by Red/ET recombination pave the way for transcriptional fine-tuning. *Biotechniques* 45, 335–337. <https://doi.org/10.2144/000112907>.
- Cardineau, G.A., Curtis 3rd, R., 1987. Nucleotide sequence of the *asd* gene of *Streptococcus mutans*. Identification of the promoter region and evidence for attenuator-like sequences preceding the structural gene. *J. Biol. Chem.* 262, 3344–3353.
- Chen, C.T., Chen, F.Y., Bogorad, I.W., Wu, T.Y., Zhang, R., Lee, A.S., Liao, J.C., 2018. Synthetic methanol auxotrophy of *Escherichia coli* for methanol-dependent growth and production. *Metab. Eng.* 49, 257–266. <https://doi.org/10.1016/j.ymben.2018.08.010>.
- Chen, Z., Geng, F., Zeng, A.P., 2015. Protein design and engineering of a *de novo* pathway for microbial production of 1,3-propanediol from glucose. *Biotechnol. J.* 10, 284–289. <https://doi.org/10.1002/biot.201400235>.
- Cheong, S., Clomburg, J.M., Gonzalez, R., 2016. Energy- and carbon-efficient synthesis of functionalized small molecules in bacteria using non-decarboxylative Claisen condensation reactions. *Nat. Biotechnol.* 34, 556–561. <https://doi.org/10.1038/nbt.3505>.
- Chou, A., Clomburg, J.M., Qian, S., Gonzalez, R., 2019. 2-Hydroxyacyl-CoA lyase catalyzes acyloin condensation for one-carbon bioconversion. *Nat. Chem. Biol.* 15, 900–906. <https://doi.org/10.1038/s41589-019-0328-0>.
- Claassens, N.J., Cotton, C.A.R., Kopljar, D., Bar-Even, A., 2019. Making quantitative sense of electromicrobial production. *Nat Catal* 2, 437–447. <https://doi.org/10.1038/>

- s41929-019-0272-0.
- Contestabile, R., Paiardini, A., Pascarella, S., di Salvo, M.L., D'Aguanno, S., Bossa, F., 2001. L-Threonine aldolase, serine hydroxymethyltransferase and fungal alanine racemase. A subgroup of strictly related enzymes specialized for different functions. *Eur. J. Biochem.* 268, 6508–6525. <https://doi.org/10.1046/j.0014-2956.2001.02606.x>.
- Cotton, C.A., Claassens, N.J., Benito-Vaquerizo, S., Bar-Even, A., 2020. Renewable methanol and formate as microbial feedstocks. *Curr. Opin. Biotechnol.* 62, 168–180. <https://doi.org/10.1016/j.copbio.2019.10.002>.
- Crowther, G.J., Kosaly, G., Lidstrom, M.E., 2008. Formate as the main branch point for methylotrophic metabolism in *Methylobacterium extorquens* AM1. *J. Bacteriol.* 190, 5057–5062. <https://doi.org/10.1128/JB.00228-08>.
- Dai, Z., Gu, H., Zhang, S., Xin, F., Zhang, W., Dong, W., Ma, J., Jia, H., Jiang, M., 2017. Metabolic construction strategies for direct methanol utilization in *Saccharomyces cerevisiae*. *Bioresour. Technol.* 245, 1407–1412. <https://doi.org/10.1016/j.biortech.2017.05.100>.
- Datsenko, K.A., Wanner, B.L., 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* 97, 6640–6645. <https://doi.org/10.1073/pnas.120163297>.
- Ebrahim, A., Lerman, J.A., Palsson, B.O., Hyduke, D.R., 2013. COBRApy: Constraints-based reconstruction and analysis for Python. *BMC Syst. Biol.* 7, 74. <https://doi.org/10.1186/1752-0509-7-74>.
- Erb, T.J., Jones, P.R., Bar-Even, A., 2017. Synthetic metabolism: metabolic engineering meets enzyme design. *Curr. Opin. Chem. Biol.* 37, 56–62. <https://doi.org/10.1016/j.cbpa.2016.12.023>.
- Fang, J., Hait, D., Head-Gordon, M., Chang, M.C.Y., 2019. Chemoenzymatic platform for synthesis of chiral organofluorines based on Type II aldolases. *Angew. Chem. Int. Ed. Engl.* 58, 11841–11845. <https://doi.org/10.1002/anie.201906805>.
- Fesko, K., 2016. Threonine aldolases: perspectives in engineering and screening the enzymes with enhanced substrate and stereo specificities. *Appl. Microbiol. Biotechnol.* 100, 2579–2590. <https://doi.org/10.1007/s00253-015-7218-5>.
- Fitzherbert, E.B., Struebig, M.J., Morel, A., Danielsén, F., Brühl, C.A., Donald, P.F., Phalan, B., 2008. How will oil palm expansion affect biodiversity? *Trends Ecol. Evol.* 23, 538–545. <https://doi.org/10.1016/j.tree.2008.06.012>.
- Giavalisco, P., Li, Y., Matthes, A., Eckhardt, A., Hubberten, H.M., Hesse, H., Segu, S., Hummel, J., Köhl, K., Willmitzer, L., 2011. Elemental formula annotation of polar and lipophilic metabolites using  $^{13}\text{C}$ ,  $^{15}\text{N}$  and  $^{34}\text{S}$  isotope labelling, in combination with high-resolution mass spectrometry. *Plant J.* 68, 364–376. <https://doi.org/10.1111/j.1365-3113.2011.04682.x>.
- Gleizer, S., Ben-Nissan, R., Bar-On, Y.M., Antonovsky, N., Noor, E., Zohar, Y., Jona, G., Krieger, E., Shamshoum, M., Bar-Even, A., Milo, R., 2019. Conversion of *Escherichia coli* to generate all biomass carbon from  $\text{CO}_2$ . *Cell* 179, 1255–1263. <https://doi.org/10.1016/j.cell.2019.11.009>. e12.
- Gresham, D., Hong, J., 2015. The functional basis of adaptive evolution in chemostats. *FEMS Microbiol. Rev.* 39, 2–16. <https://doi.org/10.1111/1574-6976.12082>.
- He, H., Edlich-Muth, C., Lindner, S.N., Bar-Even, A., 2018. Ribulose monophosphate shunt provides nearly all biomass and energy required for growth of *E. coli*. *ACS Synth. Biol.* 7, 1601–1611. <https://doi.org/10.1021/acssynbio.8b00093>.
- Hernandez, K., Bujons, J., Joglar, J., Charnock, S.J., de Maria, P.D., Fessner, W.D., Clapes, P., 2017. Combining aldolases and transaminases for the synthesis of 2-Amino-4-hydroxybutanoic acid. *ACS Catal.* 7, 1707–1711. <https://doi.org/10.1021/acscatal.6b03181>.
- Jensen, S.I., Lennen, R.M., Herrgard, M.J., Nielsen, A.T., 2015. Seven gene deletions in seven days: fast generation of *Escherichia coli* strains tolerant to acetate and osmotic stress. *Sci. Rep.* 5, 17874. <https://doi.org/10.1038/srep17874>.
- Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. <https://doi.org/10.1093/molbev/mst010>.
- Keseler, I.M., Collado-Vides, J., Gama-Castro, S., Ingraham, J., Paley, S., Paulsen, I.T., Peralta-Gil, M., Karp, P.D., 2005. EcoCyc: a comprehensive database resource for *Escherichia coli*. *Nucleic Acids Res.* 33, D334–D337. <https://doi.org/10.1093/nar/gki108>.
- Kim, J., Copley, S.D., 2012. Inhibitory cross-talk upon introduction of a new metabolic pathway into an existing metabolic network. *Proc. Natl. Acad. Sci. U. S. A.* 109, E2856–E2864. <https://doi.org/10.1073/pnas.1208509109>.
- Kumar, S., Stecher, G., Li, M., Nkay, C., Tamura, K., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. <https://doi.org/10.1093/molbev/msy096>.
- Lessmeier, L., Pfeifenschneider, J., Carnicer, M., Heux, S., Portais, J.C., Wendisch, V.F., 2015. Production of carbon-13-labeled cadaverine by engineered *Corynebacterium glutamicum* using carbon-13-labeled methanol as co-substrate. *Appl. Microbiol. Biotechnol.* 99, 10163–10176. <https://doi.org/10.1007/s00253-015-6906-5>.
- Li, G.W., Burkhardt, D., Gross, C., Weissman, J.S., 2014. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell* 157, 624–635. <https://doi.org/10.1016/j.cell.2014.02.033>.
- Lu, X., Liu, Y., Yang, Y., Wang, S., Wang, Q., Wang, X., Yan, Z., Cheng, J., Liu, C., Yang, X., Luo, H., Yang, S., Gou, J., Ye, L., Lu, L., Zhang, Z., Guo, Y., Nie, Y., Lin, J., Li, S., Tian, C., Cai, T., Zhuo, B., Ma, H., Wang, W., Ma, Y., Liu, Y., Li, Y., Jiang, H., 2019. Constructing a synthetic pathway for acetyl-coenzyme A from one-carbon through enzyme design. *Nat. Commun.* 10, 1378. <https://doi.org/10.1038/s41467-019-09095-z>.
- Marx, C.J., Lidstrom, M.E., 2004. Development of an insertional expression vector system for *Methylobacterium extorquens* AM1 and generation of null mutants lacking *mtaA* and/or *feh*. *Microbiology* 150, 9–19. <https://doi.org/10.1099/mic.0.26587-0>.
- Meyer, F., Keller, P., Hartl, J., Groninger, O.G., Kiefer, P., Vorholt, J.A., 2018. Methanol-essential growth of *Escherichia coli*. *Nat. Commun.* 9, 1508. <https://doi.org/10.1038/s41467-018-03937-y>.
- Monk, J.M., Lloyd, C.J., Brunk, E., Mih, N., Sastry, A., King, Z., Takeuchi, R., Nomura, W., Zhang, Z., Mori, H., Feist, A.M., Palsson, B.O., 2017. iML1515, a knowledgebase that computes *Escherichia coli* traits. *Nat. Biotechnol.* 35, 904–908. <https://doi.org/10.1038/nbt.3956>.
- Neidhardt, F.C., Ingraham, J.L., Schaechter, M., 1990. Building blocks needed to produce 1g of *E. coli* protoplasm. *Physiology of the Bacterial Cell: A Molecular Approach* 134–143.
- Noor, E., Bar-Even, A., Flamholz, A., Reznik, E., Liebermeister, W., Milo, R., 2014. Pathway thermodynamics highlights kinetic obstacles in central metabolism. *PLoS Comput. Biol.* 10, e1003483. <https://doi.org/10.1371/journal.pcbi.1003483>.
- Noor, E., Haraldsdottir, H.S., Milo, R., Fleming, R.M., 2013. Consistent estimation of Gibbs energy using component contributions. *PLoS Comput. Biol.* 9, e1003098. <https://doi.org/10.1371/journal.pcbi.1003098>.
- Pfeifenschneider, J., Brautaset, T., Wendisch, V.F., 2017. Methanol as carbon substrate in the bio-economy: metabolic engineering of aerobic methylotrophic bacteria for production of value-added chemicals. *Biofuel Bioprod Biorefin* 11, 719–731. <https://doi.org/10.1002/bbb.1773>.
- Sanderson, K., 2011. Lignocellulose: a chewy problem. *Nature* 474, S12–S14. <https://doi.org/10.1038/474S012a>.
- Schada von Borzyskowski, L., Remus-Emsermann, M., Weishaupt, R., Vorholt, J.A., Erb, T.J., 2015. A set of versatile brick vectors and promoters for the assembly, expression, and integration of synthetic operons in *Methylobacterium extorquens* AM1 and other alphaproteobacteria. *ACS Synth. Biol.* 4, 430–443. <https://doi.org/10.1021/sb500221v>.
- Schrader, J., Schilling, M., Holtmann, D., Sell, D., Filho, M.V., Marx, A., Vorholt, J.A., 2009. Methanol-based industrial biotechnology: current status and future perspectives of methylotrophic bacteria. *Trends Biotechnol.* 27, 107–115. <https://doi.org/10.1016/j.tibtech.2008.10.009>.
- Szima, S., Cormos, C.-C., 2018. Improving methanol synthesis from carbon-free  $\text{H}_2$  and captured  $\text{CO}_2$ : a techno-economic and environmental evaluation. *J. CO<sub>2</sub> Util.* 24, 555–563. <https://doi.org/10.1016/j.jcou.2018.02.007>.
- Takors, R., Kopf, M., Mampel, J., Bluemke, W., Blombach, B., Eikmanns, B., Bengelsdorf, F.R., Weuster-Botz, D., Dürre, P., 2018. Using gas mixtures of  $\text{CO}$ ,  $\text{CO}_2$  and  $\text{H}_2$  as microbial substrates: the do's and don'ts of successful technology transfer from laboratory to production scale. *Microb. Biotechnol.* 11, 606–625. <https://doi.org/10.1111/1751-7915.13270>.
- Thomason, L.C., Costantino, N., Court, D.L., 2007. *E. coli* genome manipulation by P1 transduction. *Curr. Protoc. Mol. Biol.* 79 (1), 1–17. <https://doi.org/10.1002/0471142727.mb0117s79>.
- Trudeau, D.L., Edlich-Muth, C., Zarzycki, J., Scheffen, M., Goldsmith, M., Khersonsky, O., Avizemer, Z., Fleishman, S.J., Cotton, C.A.R., Erb, T.J., Tawfik, D.S., Bar-Even, A., 2018. Design and *in vitro* realization of carbon-conserving photorespiration. *Proc. Natl. Acad. Sci. U. S. A.* 115, E11455–E11464. <https://doi.org/10.1073/pnas.1812605115>.
- Walker, D.A., 2009. Biofuels, facts, fantasy, and feasibility. *J. Appl. Phycol.* 21, 509–517. <https://doi.org/10.1007/s10811-009-9446-5>.
- Walther, T., Calvayrac, F., Malbert, Y., Alkim, C., Dressaire, C., Cordier, H., Francois, J.M., 2018. Construction of a synthetic metabolic pathway for the production of 2,4-dihydroxybutyric acid from homoserine. *Metab. Eng.* 45, 237–245. <https://doi.org/10.1016/j.jmben.2017.12.005>.
- Wang, C., Ren, J., Zhou, L., Li, Z., Chen, L., Zeng, A.P., 2019. An aldolase-catalyzed new metabolic pathway for the assimilation of formaldehyde and methanol to synthesize 2-Keto-4-hydroxybutyrate and 1,3-propanediol in *Escherichia coli*. *ACS Synth. Biol.* 8, 2483–2493. <https://doi.org/10.1021/acssynbio.9b00102>.
- Wang, X.L., Wang, Y., Liu, J., Li, Q.G., Zhang, Z.D., Zheng, P., Lu, F.P., Sun, J.B., 2017. Biological conversion of methanol by evolved *Escherichia coli* carrying a linear methanol assimilation pathway. *Bioresour. Bioprocess* 4, 41–46. <https://doi.org/10.1186/s40643-017-0172-6>.
- Wang, Y., Fan, L., Tuyishime, P., Zheng, P., Sun, J., 2020. Synthetic methylotrophy: a practical solution for methanol-based biomanufacturing. *Trends Biotechnol.* <https://doi.org/10.1016/j.tibtech.2019.12.013>.
- Wenk, S., Yishai, O., Lindner, S.N., Bar-Even, A., 2018. An engineering approach for re-wiring microbial metabolism. *Methods Enzymol.* 608, 329–367. <https://doi.org/10.1016/bs.mie.2018.04.026>.
- Whitaker, W.B., Sandoval, N.R., Bennett, R.K., Fast, A.G., Papoutsakis, E.T., 2015. Synthetic methylotrophy: engineering the production of biofuels and chemicals based on the biology of aerobic methanol utilization. *Curr. Opin. Biotechnol.* 33, 165–175. <https://doi.org/10.1016/j.copbio.2015.01.007>.
- Withhoff, S., Schmitz, K., Niedner, S., Noh, K., Noack, S., Bott, M., Marienhagen, J., 2015. Metabolic engineering of *Corynebacterium glutamicum* for methanol metabolism. *Appl. Environ. Microbiol.* 81, 2215–2225. <https://doi.org/10.1128/AEM.03110-14>.
- Yang, X., Yuan, Q., Luo, H., Li, F., Mao, Y., Zhao, X., Du, J., Li, P., Ju, X., Zheng, Y., Chen, Y., Liu, Y., Jiang, H., Yao, Y., Ma, H., Ma, Y., 2019. Systematic design and *in vitro* validation of novel one-carbon assimilation pathways. *Metab. Eng.* 56, 142–153. <https://doi.org/10.1016/j.jmben.2019.09.001>.
- Yishai, O., Goldbach, L., Tenenboim, H., Lindner, S.N., Bar-Even, A., 2017. Engineered assimilation of exogenous and endogenous formate in *Escherichia coli*. *ACS Synth. Biol.* 6 (9), 1722–1731. <https://doi.org/10.1021/acssynbio.7b00086>.
- Yishai, O., Lindner, S.N., Gonzalez de la Cruz, J., Tenenboim, H., Bar-Even, A., 2016. The formate bio-economy. *Curr. Opin. Chem. Biol.* 35, 1–9. <https://doi.org/10.1016/j.cbpa.2016.07.005>.
- You, L., Page, L., Feng, X., Berla, B., Pakrasi, H.B., Tang, Y.J., 2012. Metabolic pathway confirmation and discovery through  $^{13}\text{C}$ -labeling of proteinogenic amino acids. *J. Vis. Exp.* 59. <https://doi.org/10.3791/3583>.

- Yu, H., Liao, J.C., 2018. A modified serine cycle in *Escherichia coli* converts methanol and CO<sub>2</sub> to two-carbon compounds. Nat. Commun. 9, 3992. <https://doi.org/10.1038/s41467-018-06496-4>.
- Yu, J.L., Xia, X.X., Zhong, J.J., Qian, Z.G., 2014. Direct biosynthesis of adipic acid from a synthetic pathway in recombinant *Escherichia coli*. Biotechnol. Bioeng. 111, 2580–2586. <https://doi.org/10.1002/bit.25293>.
- Zakaria, Z., Kamarudin, S.K., 2016. Direct conversion technologies of methane to methanol: an overview. Renew. Sustain. Energy Rev. 65, 250–261. <https://doi.org/10.1016/j.rser.2016.05.082>.
- Zelcbuch, L., Antonovsky, N., Bar-Even, A., Levin-Karp, A., Barenholz, U., Dayagi, M., Liebermeister, W., Flamholz, A., Noor, E., Amram, S., Brandis, A., Bareia, T., Yofe, I., Jubran, H., Milo, R., 2013. Spanning high-dimensional expression space using ribosome-binding site combinatorics. Nucleic Acids Res. 41, e98. <https://doi.org/10.1093/nar/gkt151>.
- Zhang, W.M., Zhang, T., Wu, S.H., Wu, M.K., Xin, F.X., Dong, W.L., Ma, J.F., Zhang, M., Jiang, M., 2017. Guidance for engineering of synthetic methylotrophy based on methanol metabolism in methylotrophy. RSC Adv. 7, 4083–4091. <https://doi.org/10.1039/c6ra27038g>.
- Zhao, M., Lu, X., Zong, H., Li, J., Zhuge, B., 2018. Itaconic acid production in microorganisms. Biotechnol. Lett. 40, 455–464. <https://doi.org/10.1007/s10529-017-2500-5>.
- Zhong, W., Zhang, Y., Wu, W., Liu, D., Chen, Z., 2019. Metabolic engineering of a homoserine-derived non-natural pathway for the *de novo* production of 1,3-propanediol from glucose. ACS Synth. Biol. 8, 587–595. <https://doi.org/10.1021/acssynbio.9b00003>.