Mixing and matching methylotrophic enzymes to design a novel methanol utilization pathway in E. coli

Alessandro de Simone, Cláudia M Vicente, Camille Peiro, Lara Gales, Floriant Bellvert, Brice Enjalbert, Stephanie Heux

Mixing and matching methylotrophic enzymes to design a novel methanol utilization pathway in *E. coli*

De Simone A.[1]; Vicente C.M.[1]; Peiro C.[1]; Gales L.[1,2], Bellvert F.[1,2]; Enjalbert B.[1]; Heux S.[1*]

[1]TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France

[2]MetaboHUB-MetaToul, National Infrastructure of Metabolomics and Fluxomics, Toulouse, 31077, France

* Corresponding author. Mailing address TBI, INSA Toulouse, 135 Avenue de Rangueil, 31077 Toulouse cedex 4, France. Fax: +33 561 55 96 89. E-mail: stephanie.heux@insa-toulouse.fr

ABSTRACT

One-carbon (C1) compounds, such as methanol, have recently gained attention as alternative low-cost and non-food feedstocks for microbial bioprocesses. Considerable research efforts are thus currently focused on the generation of synthetic methylotrophs by transferring methanol assimilation pathways into established bacterial production hosts. In this study, we used an iterative combination of dry and wet approaches to design, implement and optimize this metabolic trait in the most common chassis, *E. coli*. Through *in silico* modeling, we designed a new route that "mixed and matched" two methylotrophic enzymes: a bacterial methanol dehydrogenase (Mdh) and a dihydroxyacetone synthase (Das) from yeast. To identify the best combination of enzymes to introduce into *E. coli*, we built a library of 266 pathway variants containing different combinations of Mdh and Das homologues and screened it using high-throughput ^{13}C-labeling experiments. The highest level of incorporation of methanol into central metabolism intermediates (e.g. 22% into the PEP), was obtained using a variant composed of a Mdh from *A. gerneri* and a codon-optimized version of *P. angusta* Das. Finally, the activity of this new synthetic pathway was further improved by engineering strategic metabolic targets identified using omics and modelling approaches. The final synthetic strain had 1.5 to 5.9 times higher methanol assimilation in intracellular metabolites and proteinogenic amino acids than the starting strain did. Broadening the repertoire of methanol assimilation pathways is one step further toward synthetic methylotrophy in *E. coli*.

KEYWORDS: One-carbon metabolism, Methanol, *Escherichia coli*, Synthetic methylotrophy.

1. INTRODUCTION

In the quest to replace fossil fuel-based processes with more sustainable bio-based ones, low-cost and easy to use fermentation substrates are of great interest. Commonly used feedstocks such as hydrolyzed starch and molasses have the disadvantage of competing with food supply, and lignocellulosic biomass requires costly pre-treatment. A promising alternative feedstock is methanol, an abundant and pure raw material that can be utilized directly in bacterial fermentation processes. Furthermore, methanol's higher degree of reduction means that it is more electron rich than carbohydrates and these extra electrons can be expected to enhance product yields during fermentation (Whitaker et al., 2015). Methanol is currently one of the top five commodity chemicals with a global production capacity of about 110 million metric tons per year and a price similar to that of glucose (http://www.methanol.org/). Although methanol is mainly produced from fossil resources, a notable advantage is that it can be produced by polygeneration, as a product of any renewable resource that can be converted into an intermediated synthesis gas (syngas). This includes biomass, agricultural and timber waste, municipal solid waste, landfill gas, industrial waste and a number of other feedstocks (http://enerkem.com/fr/; http://www.methanol.org/). Bio-methanol can also be produced from the thermo-, electro- or photo- catalytic reduction of the notorious greenhouse gas CO_2. These approaches, which are still under development, can provide a way to recycle emitted CO_2 creating a carbon neutral cycle and, at the same time, store renewable or (excess) energy (Simakov, 2017). All these factors make methanol an attractive feedstock for biorefineries and the concept of a methanol economy has received considerable attention (Olah, 2013; Schrader et al., 2009).

Methylotrophy is the capacity of certain prokaryote and eukaryote microorganisms to use reduced one-carbon (C1) compounds such as methanol as their sole source of carbon and energy. This metabolism includes: (i) the oxidation of methanol to formaldehyde; (ii) the oxidation of formaldehyde to CO_2, and (iii) the assimilation of one carbon compounds, either formaldehyde or CO_2 or a combination thereof (Heux S. et al., 2018). The industrial-scale use of natural methylotrophs has already been attempted. In the 1970s, a process was developed to produce single-cell protein (SCP) from methanol (Matelbs and Tannenbaum, 1968; Windass et al., 1980), but the technology fell out of favor in the following decades because of the low prices of alternative sources such as soybean protein. Currently, the use of natural methylotrophs in bioprocesses is only seen in the production by methylotrophic yeasts of recombinant proteins such as enzymes, antibodies, cytokines, plasma proteins, and hormones (Ahmad et al., 2014). The production of small molecules and metabolites (e.g. PHAs (polyhydroxyalkanoates) and amino acids) is still at the proof-of-concept stage (Schrader et al., 2009). The main limitations to the use of natural methylotrophs in biotechnologies are our currently weak understanding of their cellular metabolism and physiology, and the general lack of genetic tools to modify them (Chung et al., 2010; Schrader et al., 2009). In contrast, *Escherichia coli* is a robust biotechnological chassis with a wide range of products and an extensive genetic toolbox (Becker and Wittmann, 2015). Engineering a methanol assimilation pathway in this microorganism has thus become a popular research topic.

Methylotrophy is quite challenging to engineer because all biomass production and energy requirements must be satisfied by a reduced C1 precursor. In addition, cells must be able to tolerate formaldehyde, a central but toxic compound in methanol metabolism, whose accumulation due to an imbalance between oxidation and assimilation in the pathway can be fatal for cells. Because formaldehyde oxidation is efficient, the main bottleneck is C1 assimilation, which is achieved through a cyclic

process involving a C1-acceptor to enable the formation of C-C bonds. Several attempts have been made to engineer synthetic methylotrophy in *E. coli* using naturally occurring cyclic pathways (Wang et al., 2020). Most of these involve the expression of three heterologous enzymes: a NAD^+-dependent methanol dehydrogenase (Mdh) for the oxidation of methanol to formaldehyde together with hexulose phosphate synthase (Hps) and phosphohexuloisomerase (Phi) from the ribulose monophosphate (RuMP) cycle for formaldehyde fixation. The *in vivo* operation of this pathway in *E. coli* has been confirmed by isotope-labeling experiments, which showed that methanol carbons were incorporated into cellular material (Muller et al., 2015). Similar results have also been reported in other model organisms such as *Corynebacterium glutamicum*, *Pseudomonas putida* and *Saccharomyces cerevisiae* (as reviewed recently by (Heux S. et al., 2018)). Improvements in methanol assimilation have been achieved using different strategies such as (i) optimizing the cultivation medium (Gonzalez et al., 2018), (ii) lowering the thermodynamic and kinetic constraints associated with NAD-dependent methanol oxidation (Roth et al., 2019; Wu et al., 2016), (iii) improving formaldehyde assimilation (Price et al., 2016; Woolston et al., 2018), (iv) increasing carbon fluxes through the autocatalytic cycle (Bennett et al., 2018), and (v) coupling the activity of the RuMP cycle to the growth of the host microorganism and then using adaptive laboratory evolution (Chen et al., 2018; He et al., 2018; Meyer et al., 2018). However, none of these synthetic strains are able to grow on methanol alone. The reasons for this and the obstacles to overcome include regenerating the C1-acceptor, protecting the cells against formaldehyde toxicity, channeling the substrate so that it can be integrated directly into the central metabolism, and lowering energetic constraints.

The approach outlined here to tackle the exciting challenge of synthetic methylotrophy is to develop a hybrid of naturally occurring cyclic methanol assimilation pathways. Using a “mix and match” approach, we created a new

synthetic pathway combining Mdh, a methylotrophic enzyme of bacterial origin, with dihydroxyacetone synthase (Das), a methylotrophic enzyme from yeast. The engineered strain was then optimized in an iterative process using omics (transcriptomics, metabolomics and fluxomics) and modelling approaches to identify bottlenecks. Overall, this approach allows non-natural pathways to be explored and tested while offering new perspectives on synthetic methylotrophy in *E. coli*.

2. RESULTS

2.1. Selecting the best design for a methanol assimilation pathway

Natural methylotrophs have developed multiple pathways that allow them to grow on methanol as the sole source of carbon and energy (Chistoserdova, 2011). From this metabolic diversity, we can estimate that there are more than 500 unique methanol assimilation pathways from methanol to biomass (Heux S. et al., 2018). To identify the best pathway for *E. coli* to consume methanol, we used FindPath, a tool that freely recombines a repertoire of existing reactions to create metabolic pathways (Vieira et al., 2014). FindPath uses a substrate-associated reaction database and flux balance analysis (FBA) based on a genome scale model (GSM) of the host to (i) find all the possible pathways, and (ii) rank them according to their length and the predicted growth rate on the substrate of interest. The tool identified two equally efficient synthetic routes: the already well-studied RuMP-based pathway involving the bacterial enzymes Mdh, Hps and Phi, and a hybrid metabolic pathway, involving methylotrophic bacterial Mdh and methylotrophic yeast derived Das (Figure 1). The latter is a transketolase that catalyzes the fixation of formaldehyde on xylulose 5-phosphate (Xu5P) to form glyceraldehyde 3-phosphate (GAP) and dihydroxyacetone (DHA) in the xylulose monophosphate (XuMP) cycle in methylotrophic yeasts. The GSM-predicted growth rate of *E. coli* on methanol with this pathway is 0.34 h^{-1}. The

153 predicted fluxes show that this optimal growth rate is achieved when 16% of the
154 methanol is incorporated into the biomass with the rest being used to recycle the C1
155 acceptor, Xu5P. No flux through the methanol oxidation pathway (i.e. through FrmA &
156 B) was predicted. In addition, the GSM predicted that Xu5P would be recycled by
157 fructose-6-phosphate aldolase and transaldolase (FSA/TAL pathway variant) rather
158 than by fructose 1,6-bisphosphate aldolase with transaldolase (FBA/TAL variant), or
159 by sedoheptulose biphosphatase (FBA/SBP variant) (Supplementary Figure S1). No
160 ATP is required for Xu5P regeneration in the FSA/TAL variant, while the other two
161 metabolic variants require two ATP molecules (Supplementary Figure S1). In
162 comparison, in the synthetic RuMP based pathway, the C1 acceptor, ribulose-5-
163 phosphate, is recycled using one or two ATP molecules. However, in term of
164 assimilation no ATP is required in the synthetic RuMP based pathway while one
165 equivalent of ATP is needed to phosphorylate DHA into DHAP (Supplementary
166 Figure S1). Considering both recycling and assimilation, the synthetic RuMP pathway
167 requires the same amount of ATP with the FBA/TA variant and one more ATP with
168 the FBA/SBP variant compared to the FSA/TA variant which can occur only with the
169 proposed pathway (Woolston et al., 2018). Finally, the transketolase activity of Das
170 may contribute to the regeneration of the C1 acceptor, making our pathway partially
171 independent of the pentose phosphate pathway (PPP). This has been demonstrated
172 recently in methylotrophic yeasts (Russmayer et al., 2015).

173

174 ***2.2. Screening for the best matching methylotrophic enzymes***

175 To identify the combination of enzymes that would optimize the *in vivo* activity of the
176 pathway in *E. coli*, we built a combinatorial library of *mdh* and *das* homologues
177 derived from native and non-native methylotrophs (bacteria and yeasts). Starting with
178 the query protein sequences of *Bacillus methanolicus PB1* Mdh2 and *Pichia angusta*

179 Das, a BLAST search with an identity cut-off at 50% was used with the software CD-
180 HIT to filter out and cluster homologous templates (see Materials and Methods
181 section). This threshold ensured that only functional homologues were identified
182 (Sangar et al., 2007). Some of the 12 prokaryotic Mdh sequences and 17 eukaryotic
183 Das sequences selected in this way belong to genera known to contain
184 methylotrophs, such as *Bacillus*, *Burkholderia* (Chistoserdova et al., 2009),
185 *Acinetobacter* (Del Rocío Bustillos-Cristales et al., 2017), *Pichia* and *Candida*
186 (Supplementary Figure S2). We added an Mdh from *Bacillus stearothermophilus* and
187 a *P. angusta* methanol oxidase (MOX), since both have been reported to have good
188 affinity for methanol (Kms of 20 mM and 0.4 mM, respectively) (Shleev et al., 2006;
189 Whitaker et al., 2017). Finally we added a Das from *Mycobacterium* with a high
190 affinity for formaldehyde (Km of 1.86 mM) (Ro et al., 1997) and an *E. coli* codon-
191 optimized version of *P. angusta* Das. The Mdh and Das genes were respectively
192 cloned into the low-copy plasmid pSEVA424 and the middle-copy plasmid pSEVA134
193 (Silva-Rocha et al., 2013). All the selected sequences were assembled in a library of
194 266 combinations of genes (14 Mdh sequences * 19 Das sequences) and
195 transformed using a robotic platform (Supplementary Figure S3). To prevent the
196 induction of the formaldehyde detoxification pathway (formaldehyde to CO_2) (Figure
197 1), we used an *E. coli* strain Δ*frmA*, deleted for the formaldehyde dehydrogenase i.e.
198 the first gene of this pathway. Enzyme expression were measured at two reasonable
199 *E. coli* growth temperatures i.e. 30°C and 37°C. The higher expression obtained at
200 30°C overnight (Supplementary Figure S4) led us to use this temperature for all
201 subsequent experiments.

202 To analyze the performance of the 266 different enzyme combinations, methanol
203 incorporation was measured for each pathway variant using dynamic ^{13}C-labeling
204 experiments as shown in Supplementary Figure S3. We used the ^{13}C-labeling
205 incorporation into the phosphoenolpyruvate (PEP) as a proxy for methanol

assimilation since PEP is one of the first multi-carbon products of methanol assimilation (Figure 1). The ^{13}C-enrichment of PEP measured for each combination of Mdh and Das is shown in Figure 2A and Supplementary Table S2. ^{13}C enrichments of between 1% to 5% were observed for the combinations involving the *das* genes of *Pichia pastoris*, *Verruconis gallopava*, *Scedosporium apiospermum, Rasamsonia emersonii*, *Fonsecaea erecta* and *Kuraishia capsulata.* In comparison, combinations involving the *das* from *P. angusta*, had ^{13}C enrichments two to twelve times higher, up to 22% for the codon-optimized version, *P. angusta* (opt), representing an average 2.4-fold increase in ^{13}C-isotopic enrichment in PEP compared with the wild type (Supplementary Table S2). These results are consistent with the more stable expression of the codon-optimized version of Das, compared with the wild type version (Supplementary Figure S5).

We then investigated whether methanol assimilation could be increased by optimizing the expression of the Das enzymes which led to low- or non-labeled PEP. The codon-optimized Das genes from *Candida boidinii*, *P. methanolica A*, *Aspergillus fumigatus Z5*, *D. hansenii*, *R. emersonii* and *K. capsulata* were synthesized and individually co-expressed with *A. gerneri* Mdh. No significant improvement in ^{13}C-enrichment was observed compared with the native sequences, except for codon-optimized *R. emersonii* Das, whose PEP labeling was twice as high (4% vs 2% ^{13}C-enrichment) (Supplementary Table S3).

Labeling was observed for all these combinations regardless of the nature of the Mdh, suggesting Das compensated for the generally poor kinetic properties of the NAD$^+$-dependent Mdh enzymes (Brautaset et al., 2013; Krog et al., 2013) by shifting the equilibrium toward methanol oxidation and subsequent formaldehyde assimilation. However, the level of ^{13}C-incorporation was not linked with the expression levels of either Mdh or Das (Figure 2A and Supplementary Figure S5). It

232 is also worth noting that although *B. stearothermophilus* Mdh and *Mycobacterium*
233 Das have favorable *in vitro* and *in vivo* activities (Ro et al., 1997; Whitaker et al.,
234 2017), and were well-expressed in *E. coli* (Supplementary Figure S5), they led to
235 very low ^{13}C-enrichement (< 3%) in most of the tested combinations (Figure 2A).

236 Finally, the highest methanol incorporation was achieved when either *A. gerneri* Mdh
237 or *Burkholderia sp. TSV86* Mdh was expressed in combination with *P. angusta* Das
238 (opt). With these combinations, the ^{13}C-enrichments of PEP were respectively 13 and
239 12 times higher than with the query pathway (*B. methanolicus PB1* Mdh2 / *P.*
240 *angusta* Das). In particular, the fractions of PEP with one ^{13}C atom were 31% and
241 32% and reached 17% and 13.7% for two ^{13}C atoms, respectively (Figure 2B). The
242 incorporation of more than one labeled carbon into PEP demonstrates that the
243 recycling of Xu5P is functional in both combinations.

244 The higher ^{13}C-enrichment obtained for the combination involving *A. gerneri* Mdh and
245 *P. angusta* Das (opt) led us to use these enzymes for subsequent experiments.
246 Although this is the best matching of enzymes, the fact that 100% PEP labeling was
247 not achieved indicates that methanol alone cannot supply all the carbon atoms
248 required for molecular assembly and that pure methylotrophic growth is not yet
249 possible with this pathway.

250 ***2.3. Characterizing the cellular behavior of the synthetic methylotroph***

251 In order to uncover the specific make-up of the new synthetic methylotrophic *E. coli*
252 strain with regard to methanol utilization, we performed a physiological and
253 transcriptomic analysis of the strain expressing *A. gerneri* Mdh and *P. angusta* Das
254 (opt) grown on xylose with and without additional methanol (Table 1, Figure 3). For
255 this analysis, both genes were cloned into the pSEVA424 vector as a single operon
256 to avoid the metabolic burdening of the cells with double-antibiotic selection (Silva et
257 al., 2012). Similar levels of ^{13}C-methanol incorporation were measured in the single

and double plasmid strains after 90 min culture with methanol, but labeling continued to increase in the double plasmid strain up to twice the level observed in the single plasmid strain (Supplementary Figure S6). This can be explained by a decrease in Das levels when the gene is expressed from the low-copy plasmid pSEVA424 (10–15 copies/cell) instead of the middle-copy pSEVA131 plasmid (20–30 copies/cell) (Silva-Rocha et al., 2013).

The physiological response to methanol of the synthetic strain expressing the assimilation pathway on one plasmid is given in Table 1. Both the growth rate (+54%) and the specific xylose consumption rate (+45%) were higher in the methanol-supplemented medium than when cells were grown on xylose alone. Formate was only observed in the presence of methanol (Table 1) and its production increased once xylose was depleted (Figure 3A). In contrast, methanol consumption could not be formally assessed since the decrease in concentration occurred at a similar rate as evaporation and fell within the error range of the NMR instrument (4% of the measured value). These results clearly indicate a positive effect of methanol on the rate of xylose uptake, and thus on growth, but also show that formaldehyde was oxidized into formate even though the first step of this pathway had been deleted (i.e. *frmA*).

To characterize the cellular response of the synthetic strain to methanol, a transcriptional analysis was performed during exponential growth on xylose, with or without methanol supplementation. Specifically, we looked at the expression of the genes involved in methanol metabolism (Figure 3B). In the presence of methanol the *frmR* and *frmB* genes were strongly up-regulated. Because FrmR is a formaldehyde sensing factor which regulates expression of the *frmRAB* operon, frmR and frmB up-regulation indicates the production of formaldehyde into the cells which was expected from the presence of methanol (Figure 3B). However, the production of formate in the

284 Δ*frmA* strain (Table 1) suggests the presence of a promiscuous alcohol
285 dehydrogenase that replaces FrmA in the reaction from formaldehyde to S-
286 formylglutathione, which is then converted to formate by FrmB. Up-regulation of *dhaL*
287 and *dhaM*, which encode the dihydroxyacetone kinase (DAK) pathway, was also
288 observed in presence of methanol, (Figure 3B). Because the expression of the
289 *dhaKLM* operon is induced by DHA (Bächler et al., 2005), this confirms the presence
290 of DHA in the cells and thus the co-assimilation of methanol with xylose. However,
291 the genes encoding alternative DHA assimilation routes (i.e. the glycerol (*gldA*, *glpK*
292 and *glpD)* and the FSA (*fsaA* and *fsaB*) pathways) were not transcriptionally activated
293 or even down-regulated (Figure 3B). These results are consistent with the conclusion
294 of Peiro et al. that DHA is mainly assimilated via the dihydroxyacetone kinase (DAK)
295 (Peiro et al., 2019). However, they appear to contradict those of the flux balance
296 analysis that predict that Fsa may be involved in the regeneration of the C1 acceptor.
297 Finally, the gene encoding the transketolase *tktA* was up-regulated on methanol
298 (Figure 3B). This enzyme catalyzes the formation of Xu5P, which plays a key role in
299 the cyclic operation of our synthetic pathway. However, Xu5P is also the entry point
300 of xylose in the metabolism and, interestingly, the presence of methanol improved the
301 expression of *xylE* involved in its transport through the cellular membrane (Figure
302 3B), particularly during the exponential growth phase. This result corroborates the
303 higher specific xylose uptake rate observed when the synthetic *E. coli* strain was
304 grown in media supplemented with methanol (Table 1).

305 Overall, these data demonstrate that methanol can be assimilated by the new
306 synthetic *E. coli* strain and identified genetic engineering targets to limit its
307 dissimilation and improve the cyclic operation of the pathway.

308

309 ***2.4. Optimizing the methylotrophic chassis***

The choice of the genes (Figure 4B) to be targeted in the current strain (Strain 1) to engineer a superior methanol assimilation phenotype was based on: (i) computational prediction highlighting the key role of Fsa for the regeneration of the C1 acceptor, XU5P, making the DHA a key junction between assimilation and recycling in our synthetic strain; (ii) a previous study which demonstrated that increasing expression of fsaB and gldA leads to an improved DHA assimilation (Peiro et al., 2019); (iii) the above transcriptomic analysis identifying the genes encoding enzymes for formaldehyde dissimilation, xylulose-5-phosphate recycling and alternative DHA assimilation pathways as potential targets to boost the assimilation of methanol in our synthetic strain. Strain 2 was built by knocking out the entire *frmRAB* operon to avoid drainage of formaldehyde to the detoxification pathway. Strain 3 was built by knocking out the *frmRAB* operon in a Δ*ptsA::kan* mutant. In this strain the *ptsA* gene is replaced by a kanamycin cassette leaving the downstream genes within the same operon, *gldA* and *fsaB,* under the control of the kanamycin promoter. As previously observed in a Δ*ptsA*::kan strain (Peiro et al., 2019), a full activation of both the GLD and FSA pathways is expected in the strain 3. In strain 4 finally, *tktA*, a gene encoding a key enzyme in the regeneration of Xu5P, was overexpressed to promote this process.

To study the impact of these genetic modifications on methanol assimilation, the genealogy of the new rationally designed strains was characterized by following the incorporation of ^{13}C-methanol atoms into intracellular and extracellular metabolites (Figure 4 and controls in supplementary Figure S7). Knocking out the *frmRAB* operon (strain 2) resulted in a small increase in ^{13}C-methanol incorporation in all the measured intracellular metabolites compared with the starting Δ*frmA* strain 1 (Figure 4A), in line with measurements of the extracellular production of formate (Figure 4C). Upon xylose depletion in the medium, ^{13}C-formate production was detected in strain 1 and increased constantly during the stationary phase. In contrast, strain 2 did not

produce ^{13}C-formate, even after several hours in the stationary phase. In this strain, we observed a small but significant increase of the ^{13}C-enrichement of the pool of 2 and 3 phosphoglycerate (23PG, + 2%) and phosphoenolpyruvate (PEP, + 1.9%) and, more specifically, of the fractions with one ^{13}C atom (M1) compared with the starting strain (strain 1) (Figure 4A). The most significant improvement in methanol assimilation was observed in strain 3, in which *fsaB* and *gldA* were overexpressed. In line with the activation of the glycerol pathways in strain 3, a large fraction of glycerol-3-phosphate (GLYC3P) with two ^{13}C atoms (M2) was measured resulting in an increase of 11.5 % of the ^{13}C-enrichement compared with strain 1. GLYC3P is an important precursor of membrane constituents and therefore of biomass. In strain 3, all the measured central metabolites had more than one ^{13}C atom (Figure 4A), which can only have resulted from recycling of the C1 acceptor, XU5P. This is in accordance with the computational prediction that Fsa plays a key role in the cyclic operation of the synthetic pathway (Figure 1). Strain 4, in which *tktA* was overexpressed, showed higher ^{13}C-methanol incorporation into F6P. The fraction carrying two ^{13}C atoms was twice as high in F6P compared with strain 3 and traces of F6P containing three ^{13}C atoms (M3) were also detected. In addition, traces of labeling were measured in the pentose phosphate pool (P5P) containing XU5P.However, ^{13}C-formate was once again detected in this strain (Figure 4C). This might derive from another reaction catalysed by the 3,4-dihydroxy-2-butanone-4-phosphate synthase, RibB, which converts ribulose-5-phosphate (Ru5P) into formate (Richter et al., 1992). By overexpressing tktA, the pool of Ru5P is indeed expected to be higher.

To confirm that carbon molecules originating from methanol were used in biosynthetic pathways, we also analyzed ^{13}C incorporation into proteinogenic amino acids after 48h of cultivation on ^{13}C-methanol (Supplementary Figure S8). Low but significant levels of ^{13}C were found. In agreement with the labeling observed in the

glycolytic and TCA intermediates, labeling was also observed in their derived amino acids i.e. serine (SER, derived from glyceraldehyde-3-phosphate), alanine (ALA, derived from pyruvate), aspartate and glutamate/glutamine (ASP and GLX, derived respectively from oxaloacetate and α-ketoglutarate). As expected from the small amounts of labeled carbon in the P5P pool, no labeling was found in histidine (HIS), which is derived from ribose-5-P. However, some labeling was detected in phenylalanine (PHE), which is derived from erythrose-4-phosphate. The fraction of labeled carbons increased systematically from strain 1 to strain 4 and, more importantly, a fraction of proteinogenic amino acids were found to carry more than one ^{13}C atom (Supplementary Figure S8).

3. DISCUSSION

Methanol is an attractive feedstock for the production of fuels and chemicals but engineering a C1 fixation pathway into an industrially relevant microorganism, such as *E. coli*, remains challenging. To tackle this problem, this article describes a new computationally designed pathway as an alternative to the well-studied RuMP based pathway (Bennett et al., 2018; Chen et al., 2018; Gonzalez et al., 2018; He et al., 2018; Meyer et al., 2018; Muller et al., 2015; Price et al., 2016; Whitaker et al., 2015; Woolston et al., 2018). This new pathway is a hybrid of naturally occurring cyclic methanol assimilation pathways and consists of a Mdh from *A. gerneri* in combination with a codon-optimized version of *P. angusta* Das. Although the new pathway does not allow the cell to grow on methanol alone, 22% incorporation of methanol carbon was observed in the multi-carbon compound PEP. This is similar to the values measured previously in a synthetic methylotrophic *E. coli* strain expressing cyclic RuMP based-pathways and cultivated under comparable conditions i.e. without yeast extract (Supplementary Table S4). Importantly, this article reports the discovery of

two novel NAD-dependent alcohol dehydrogenases from Gram-negative, mesophilic, non-methylotrophic organisms (*A. gerneri* and *Burkholderia* sp.) with significant *in vivo* affinity for methanol. Representatives of the Burkholderia order have recently been recognized as true facultative methylotrophs (Chistoserdova et al., 2009) and one NAD-dependent Mdh from this order has the highest *in vitro* affinity for methanol reported to date (Woolston et al., 2018; Wu et al., 2016; Yu and Liao, 2018). In our setting, the two novel Mdhs performed better *in vivo* than the Mdhs from *B. methanolicus* and *B. stearothermophilus* which were used previously to implement methylotrophy in *E. coli* (Bennett et al., 2018; Chen et al., 2018; Gonzalez et al., 2018; Kim et al., 2020; Meyer et al., 2018; Muller et al., 2015; Whitaker et al., 2017). However, the efficiency of the pathway was mostly improved by using a codon-optimized version of Das, indicating that this enzyme is very likely rate-limiting for methanol assimilation. Since Das was not overexpressed as much as Mdh was (Supplementary Figure S4), further increasing its expression should also increase methanol assimilation.

Our iterative process of strain analysis and engineering combining omics and modelling approaches was decisive in the selection of strategic genetic targets to maximize methanol assimilation. The final optimized strain incorporated 1.5 to 5.9 times more methanol — as measured by ^{13}C-enrichment and depending on the metabolite — than did the starting strain. A maximum ^{13}C-enrichment of 37% was achieved in GLYC3P. In addition, the increase in the number of labeled carbons per molecule for most metabolites shows that cyclic operation of the synthetic pathway was improved in the final strain. Finally, the presence of labeling in biomass constituents showed that carbon molecules originating from methanol were not only assimilated into the central metabolism but also used in biosynthetic pathways. This is evidence of true methanol metabolism and confirms the establishment of methylotrophy in this *E. coli* strain. In the optimized strain, the most significant

improvement was achieved by activating alternative DHA assimilatory pathways. This is consistent with a previous study demonstrating that the specific DHA uptake rate in a similar engineered strain was increased by 60% (Peiro et al., 2019). We further improved methanol assimilation in the synthetic strain by overexpressing a transketolase and, therefore, improving the recycling of the C1 acceptor. This is in agreement with the conclusion of a previous study that expressing the non-oxidative pentose phosphate pathway (PPP) from *B. methanolicus* improves methanol assimilation in a synthetic *E. coli* methylotroph (Bennett et al., 2018) .

Finally, we also observed that methanol improved the growth of our synthetic strain on xylose by up-regulating the genes involved in xylose transport through the cellular membrane. Up-regulation of genes encoding transmembrane transporters in the presence of methanol has also been observed in *S. cerevisiae* (Espinosa et al., 2019). The chemical properties of methanol are known to modify the physical properties of cell membranes, such as their fluidity (Joo et al., 2012). These changes can be perceived by the cells and trigger the expression of genes that are involved in the acclimation of cells to new conditions (Los and Murata, 2004).

In this work, we successfully created an *E. coli* strain able to efficiently assimilate methanol through a brand new synthetic metabolic pathway. However, there is still room for optimization and our results suggest that the overall metabolic capacity for methanol can be improved in several ways. For example, one could (i) improve the expression of Das, (ii) block all the dissimilatory pathways, (iii) improve the recycling of the C1 acceptor, and (iv) coordinate the catabolic pathway with the overall cellular infrastructure by engineering methanol-sensitive elements to improve the global response to the substrate (Rohlhill et al., 2017) or by directed evolution (Chen et al., 2018; He et al., 2018; Meyer et al., 2018). However, a recent study demonstrating the slow growth (doubling time of 54 h) on a mixture of methanol and CO_2 of an *E. coli*

443 strain expressing a linear methanol assimilation pathway (Kim et al., 2020) raises
444 questions about the relevance of establishing methylotrophy in *E. coli* using cyclic
445 pathways. Arguments in favor of pursuing the quest for growth on pure methanol
446 using cyclic pathways are (i) the independence of such pathways from other carbon
447 sources, and (ii) a recent study reporting an *E. coli* strain expressing an autotrophic
448 cycle capable of producing all its biomass carbon from CO_2 (Gleizer et al., 2019).

449

450 **4. Materials and Methods**

451 ***4.1 Reagents***

452 All chemicals were purchased from Sigma-Aldrich (St. Louis, MO, USA) unless noted
453 otherwise. Unlabeled methanol (≥ 99.9%, LC-MS grade) was purchased from
454 Honeywell (Muskegon, MI, USA). Isotopically labeled ^{13}C-methanol (99% ^{13}C) was
455 purchased from Eurisotop (Saint-Aubin, France). Phusion® DNA polymerase and
456 restriction enzymes were purchased from New England Biolabs Inc. (Beverly, MA,
457 USA).

458 ***4.2 Bacterial strains and culture media***

459 All the strains, plasmids, primers and synthetic gene constructs used in this study are
460 listed in Supplementary Table S5. *E. coli* DH5α was used for plasmid construction
461 and propagation whereas *E. coli* BW25113 was used for methanol assimilation. *E.*
462 *coli* BW25113Δfrma::*neo* was obtained from the Keio collection and the Flp
463 recognition target (FRT)-flanked kanamycin cassette was removed using Flp
464 recombinase from pCP20 plasmid (Cherepanov and Wackernagel, 1995). After
465 recombination, loss of pCP20 was confirmed by re-streaking on ampicillin, and
466 removal of the resistance cassette was confirmed by polymerase chain reaction
467 (PCR). For operon construction, *A. gerneri* Mdh and *P. angusta* Das genes,

468 containing RBS and a 6xHis tag, were amplified from the pSEVA plasmids using
469 primers P91&P92 and P83&P93, respectively. The two fragments, designed to
470 overlap by 35 bp, were joined by overlapping PCR. The complete Mdh-Das operon
471 was subsequently cloned into the pSEVA424 vector using primers P102&P103 and
472 the In-Fusion® HD kit (Takara Bio, Otsu, Japan). The λ red recombination method
473 (Datsenko and Wanner, 2000) was used to generate knockout strains ΔptsA (primers
474 P129&P130) and ΔfrmRAB (primers P145&P146). The introduced antibiotic
475 resistance cassettes were removed using the FRT/FLP recombination system
476 (Cherepanov and Wackernagel, 1995). All constructs were subsequently verified by
477 colony PCR and sequencing (GATC, Konstanz, Germany).

478 All *E. coli* strains harboring plasmids were propagated in Luria-Bertani (LB) medium
479 or M9 minimal medium containing the appropriate antibiotics. The composition of the
480 M9 minimal medium was as follows (in $g \cdot L^{-1}$): 18 Na_2HPO_4, 3.13 KH_2PO_4, 0.53 NaCl,
481 2.11 NH_4Cl, 0.49 $MgSO_4 \cdot 7H_2O$, 0.00438 $CaCl_2 \cdot 2H_2O$, 0.1 thiamine hydrochloride,
482 trace elements (mg L–1) 15 $Na_2EDTA \cdot 2H_2O$, 4.5 $ZnSO_4 \cdot 7H_2O$, 0.3 $CoCl_2 \cdot 6H_2O$, 1
483 $MnCl_2$, 1 H_3BO_3, 0.4 $Na_2MoO_4 \cdot 2H_2O$, 3 $FeSO_4 \cdot 7H_2O$, 0.3 $CuSO_4 \cdot 5H_2O$. The
484 antibiotics were added when necessary in the following concentrations: ampicillin
485 (Amp, 100 µg/ml), kanamycin (Kan, 50 µg/ml), streptomycin (Strp, 50 µg/ml). The
486 optical density at 600 nm (OD600) was measured using a GENESYS 6™
487 spectrophotometer (Thermo Scientific).

488 ***4.3 In silico design of the synthetic pathway***

489 The synthetic pathway for methanol assimilation was designed using the software
490 FindPath (Vieira et al., 2014). The workflow starts with the creation of a substrate-
491 associated reaction database based on the literature and available metabolic
492 databases. This database consists of reactions involving the target molecule (in our
493 case, methanol). The database is then converted into a model that is subsequently

494 used to compute elementary flux modes (EFMs), i.e., all the possible flux distributions
495 in a metabolic network under steady state conditions. Among these EFMs, the best
496 pathways are selected and ranked according to their efficiency. Finally, the best
497 module combinations for efficient methanol conversion were identified. In our case,
498 the methanol database encompassed more than 100 reactions steps and 100
499 metabolic compounds involved in methanol metabolism. For each reaction, the
500 genes, reaction, EC number, KEGG name, localization, and reversibility were
501 reported. Finally, the model was built by bringing together all the reactions along with
502 the transporters and cofactor recycling, i.e. 47 reactions and 114 metabolites. Using
503 this model, 10000 EFMs were generated of which 85 allowed the conversion of
504 methanol into *E. coli* metabolites. From these, 20 efficient EFMs were selected, i.e.
505 those involving a small number of reactions and with low cofactor consumption (ATP,
506 NAD(P)H). The hypothesis was that their introduction into the host would require little
507 genetic effort (the number of genes being correlated with the number of reactions)
508 and would have little or no effect on the host's energy and redox machinery. The
509 reactions composing the 20 EFMs were implemented in a genome scale model *E.*
510 *coli* (iAF1620). Finally, the biomass yields on methanol of each of the 20 EFMs were
511 simulated using *in-silico* flux balance analysis (FBA),.

512 ***4.4 Library generation by combinatorial assembly***

513 A BLAST search against UniRef50 (Suzek et al., 2014) using *B. methanolicus* PB1
514 Mdh2 (UniProt ID: I3DVX6) and *P. angusta* Das (UniProt ID: P06834) as query
515 sequences returned two clusters with 177 and 230 members, respectively. The
516 sequence clustering tool H-CD-HIT (Huang et al., 2010) was used to hierarchically
517 merge similar sequences at varying levels of sequence identity. Proteins were first
518 clustered at a high identity (90%) before the non-redundant sequences were further
519 clustered at a low identity (80% and eventually 70%). Among the representatives of

520 the different clusters, we selected 12 putative Mdh variants and 17 putative Das
521 variants from aerobic and mesophilic microorganisms. The corresponding Mdh
522 genes, as well as the *Bacillus stearothermophilus* Mdh (Dowds et al., 1988) and
523 *Pichia angusta* Mox genes (Shleev et al., 2006), were cloned in the expression vector
524 pSEVA424 (Silva-Rocha et al., 2013) between restriction sites AvrII and NotI. The
525 selected Das genes, plus the *Mycobacterium* Das gene (Ro et al., 1997) and an *E.*
526 *coli* codon-optimized version of P06834, were cloned in the expression vector
527 pSEVA134 (Silva-Rocha et al., 2013) between restriction sites AvrII and SpeI. All the
528 constructs were synthesized and cloned by BaseClear (Leiden, The Netherlands).
529 The same ribosome binding site (RBS) (AGGAGGAAAAACAT) and 6xHis tag was
530 used for all the genes. The two gene libraries were co-transformed in the
531 BW25113ΔfrmA::frt strain using the rubidium chloride method (Green and Rogers,
532 2013) and plated on LB-Amp-Strp plates (Supplementary Figure S3).

533 ***4.5. Dynamic ^{13}C-labeling incorporation***

534 To study the incorporation of ^{13}C-methanol into intracellular metabolites and
535 proteinogenic amino acids, cells were first cultured in M9 minimal medium in the
536 presence of 15 mM xylose, antibiotics and 0.1 mM IPTG, in 96-deep-well plates, at
537 30°C and 220 rpm until exponential phase (OD600 = 0.5-1). The cells were then
538 centrifuged at 4400g for 3 min and resuspended in M9 minimal medium with reduced
539 (five times less) phosphate and sulfate, IPTG, antibiotics, and ^{13}C-methanol (655
540 mM). The methanol concentration was chosen to be sufficiently above the Km of
541 Mdh.

542 After 90 and 180 min incubation at 30°C and 220 rpm, intracellular metabolites were
543 sampled as follows: 120 µL of culture was taken and mixed with 1 mL of cold (–20°C)
544 acetonitrile:methanol:water:formic acid (40:40:20:0.1) extraction solution. The
545 samples were vacuum-dried overnight. The next morning, dried metabolites were

546 resuspended in 120 µL water, centrifuged at 16,000 × g for 2 min, and injected into
547 the LC-MS. Central metabolites were separated on a Dionex™ IonPac AS11-HC
548 anion-exchange column (250 × 2 mm) equipped with an AG11 guard column (50 x 2
549 mm) with KOH as the mobile phase using a Dionex™ ICS-5000+ Reagent-Free™
550 HPIC™ system (Thermo Fisher Scientific™, Sunnyvale, CA, USA). Separation of
551 PEP shown in Figure 2 was carried out with a flow rate set at 0.38 ml/min and the
552 following elution gradient: 0 min, 0.5 mM; 1 min, 0.5 mM; 9.5 min, 4.1 mM; 14.6 min,
553 4.1 mM; 24 min, 9.65 mM; 31.1 min, 100 mM and 43 min, 100 mM. For separation of
554 central metabolites shown in Figure 4, the elution gradient was as follows: 0 min, 7
555 mM; 1 min, 7 mM; 9.5 min, 15 mM; 20 min, 15 mM; 30 min, 45 mM; 33 min, 70 mM;
556 33.1 min, 100 mM; 42 min, 100mM; 42.5 min, 7 mM and 50 min, 7 mM. Metabolites
557 were detected using a Thermo Scientific™ LTQ Orbitrap Velos™ mass spectrometer
558 in negative electrospray ionization mode. The spray voltage was 2.7 kV, the capillary
559 and desolvatation temperatures were 350°C, and the maximum injection time was 50
560 msec. The spectrometer was operated in full-scan mode at a resolution of 60,000
561 (400 m/z).

562 After 48 h of incubation at 30°C and 220 rpm, proteinogenic amino acids were
563 sampled as follows: the plates were centrifuged at 4400g for 3 min and the
564 supernatant was removed. To release protein-bound amino acids from cellular
565 proteins, the cell pellets collected were hydrolyzed for 15 h with 6N HCl at 100°C.
566 HCl was evaporated at low pressure (20 mbar, room temperature). Biomass
567 hydrolysates were washed twice in water using the same evaporation method. The
568 dried hydrolysates were resuspended in 200 µL water and centrifuged. A 10-fold
569 dilution was prepared, and samples were analyzed by LC-HRMS. Proteinogenic
570 amino acids were separated on a Supelco™ HS F5 DISCOVERY column (150 × 2.1
571 mm; 5 µm) equipped with a SUPELGUARD KIT HS F5 guard column (20 x 2.1 mm;
572 5 µm) with 0.1% formic acid (solvent A) and 0.1% acetonitrile/formic acid (solvent B)

as the mobile phase using a UHPLC Vanquish system (Thermo Fisher Scientific™, Sunnyvale, CA, USA). The flow rate was set to 0.25 ml/min and the elution gradient was (% B): 0 min at 2%, 2 min at 2%, 10 min at 5%, 15 min at 35%, 20 min at 100%, 24 min at 100%, 24,1 min at 2% and 30 min at 100%. Metabolites were detected using a Thermo Scientific™ Orbitrap Q-Exactive+™ mass spectrometer in positive electrospray ionization mode, with a spray voltage of 5 kV, and capillary and desolvatation temperatures of 250°C. The spectrometer was operated in full-scan mode at a resolution of 60,000 (400 m/z).

^{13}C-carbon isotopologue distributions were identified by matching masses from the mass spectra (mass tolerance of 5 ppm) and retention times using the software TraceFinder (v. 4.1). The peaks of different isotopologues were integrated and corrected for the natural abundance and isotopic purity of the tracer using the software IsoCor (Millard et al., 2019). Levels of ^{13}C-isotopic enrichment were then determined as follows: ^{13}C-enrichement (%) = sum(Mi*i)/n, where n is the number of carbon atoms for the measured fragment and Mi is the corrected abundance of the mass isotopologue.

4.6 Supernatant analysis

Metabolite utilization and the production of the synthetic methylotroph were analyzed by quantitative 1D ^{1}H-NMR at 280 K using a zgpr30 sequence with water pre-saturation prior to acquisition on an Avance III 500 MHz spectrometer (Bruker, Rheinstetten, Germany) equipped with a 5 mm QPCI cryogenic probe head. The parameters were as follows: 286°K, 128K points, 8 s relaxation time, 2 dummy scans, 32 scans. Free induction decays (FIDs) were converted into frequency domain spectra by Fourier transform. All spectra were processed using the software TopSpin (v. 3.5). Phases were adjusted manually, baselines were adjusted automatically, and the spectra were aligned and quantified using 3-trimethylsilylpropionic-2,2,3,3-d4 acid

599 sodium salt (TSP-d4, 1 mM) as a chemical shift and concentration standard. The
600 concentrations of the different metabolites (xylose, methanol, formate, and acetate)
601 were calculated with the following equation: concentration = integrated peak
602 area*TSP concentration*dilution of the sample/number of protons in the molecule.
603 For xylose, only the peaks corresponding to the anomeric protons were integrated.

604 ***4.7 Transcriptomic analysis***

605 Cells were grown in flasks of M9 minimal media containing 15 mM xylose with or
606 without 150 mM MeOH. At T1 (OD600 = 1, exponential phase) and T2 (OD600 = 2,
607 stationary phase), 4 mL of each culture was centrifuged for 90 s at 14000 rpm before
608 discarding the supernatant and immediately freezing the pellets in liquid nitrogen.
609 Total RNA was extracted according to the Qiagen RNAeasy MiniKit procedure and
610 quantified using a Nanodrop® spectrophotometer. Double-stranded complementary
611 DNA (cDNA) synthesis and array processing were performed using the Agilent
612 Technologies One-Color Microarray-Based Gene Expression Analysis protocol. The
613 images were analyzed with the software DEVA (v. 1.2.1). All array procedures were
614 performed using the GeT-Biopuces platform (http://get.genotoul.fr/). For each data
615 set, corresponding to time point T1 or T2, the log2 intensities obtained in the
616 presence of methanol were divided by the log2 intensities obtained without methanol.
617 These ratios were then normalized by the log median intensity. Genes whose
618 expression level differed by a factor of 2 or more between the two conditions were
619 selected for further analysis. Gene ontology analyses were performed using Ecocyc
620 (https://ecocyc.org/). Gene expression data have been deposited in the ArrayExpress
621 database at EMBL-EBI (www.ebi.ac.uk/arrayexpress) under accession number E-
622 MTAB-8909.

623

624 ***4.8 In silico analysis of methanol metabolism***

We used the functions “flux balance analysis” (FBA) and “flux variability analysis” of the R environment (R Development Core Team, 2009; Team, 2015) Sybil Package (Gelius-Dietrich et al., 2013) and the genome scale model of *E. coli* ij01366 (Orth et al., 2011) amended with the heterologous reactions catalyzed by Mdh, Das and Glpx , and their associated metabolites to simulate the growth and fate of methanol. The objective function was the growth rate whereas the model was constrained using the methanol uptake rate measured experimentally for the wild-type methylotroph *Methylobacter extorquens* (15 mmol/gW/h).

4.9 Growth and methanol consumption calculations

Specific growth rates, uptake rates and production rates were determined using PhysioFit, provided open source at https://github.com/MetaSys-LISBP/PhysioFit. A conversion factor of 0.37 g dry weight/OD600 was used.

5. Acknowledgment

The authors thank MetaToul-MetaboHUB (Metabolomics & Fluxomics Facitilies, Toulouse, France, www.metatoul.fr, ANR-11-INBS-0010, www.metabohubmetabohub.fr) and its staff members for technical support and access to MS and NMR facilities. This work was funded by the ANR (ANR-16-CE20-0018-01 & ANR-17-COBI-0003-05) and TWB (https://www.toulouse-white-biotechnology.com/). The authors also thank H. Cordier and O. Galy from TWB for technical support and access to robotic platform, and the following students, L. Balcells, M. Beaudor, C. Brodeau, M. Cayet, T. Chaillet, C. Ledoux and E. Sylvander for their help producing transcriptomic data and initiating the data analysis.

6. Author contribution

A. De Simone, C.M. Vicente and C. Peiro built the strains and performed the physiological experiments. L. Gales and F. Bellvert performed the MS analysis. B.

Enjalbert performed the transcriptomic analysis. S. Heux designed the study and wrote the paper with the help of all the co-authors. The authors declare that they have no conflicts of interest.

7. References

Ahmad, M., Hirz, M., Pichler, H., Schwab, H., 2014. Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. Appl Microbiol Biotechnol. 98**,** 5301-17.

Bächler, C., Schneider, P., Bähler, P., Lustig, A., Erni, B., 2005. *Escherichia coli* dihydroxyacetone kinase controls gene expression by binding to transcription factor *DhaR*. The EMBO Journal. 24**,** 283-293.

Becker, J., Wittmann, C., 2015. Advanced biotechnology: metabolically engineered cells for the bio-based production of chemicals and fuels, materials, and health-care products. Angewandte Chemie. 54**,** 3328-50.

Bennett, R. K., Gonzalez, J. E., Whitaker, W. B., Antoniewicz, M. R., Papoutsakis, E. T., 2018. Expression of heterologous non-oxidative pentose phosphate pathway from *Bacillus methanolicus* and phosphoglucose isomerase deletion improves methanol assimilation and metabolite production by a synthetic *Escherichia coli* methylotroph. Metabolic engineering. 45**,** 75-85.

Brautaset, T., Heggeset, M., Marita, B., Heux, S., Kiefer, P., Krog, A., Lessmeier, L., Muller, J. E., Portais, J. C., Potthoff, E., Quax, W. J., Sibbald, M., Vorholt, J. A., Wendisch, V. F., Novel Methanol Dehydrogenase Enzymes from Bacillus. Vol. WO 2013110797., 2013.

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. Metabolic engineering. 49**,** 257-266.

Cherepanov, P. P., Wackernagel, W., 1995. Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. Gene. 158**,** 9-14.

Chistoserdova, L., 2011. Modularity of methylotrophy, revisited. Environmental microbiology. 13**,** 2603-22.

Chistoserdova, L., Kalyuzhnaya, M. G., Lidstrom, M. E., 2009. The expanding world of methylotrophic metabolism. Annu Rev Microbiol. 63**,** 477-99.

Chung, B. K., Selvarasu, S., Andrea, C., Ryu, J., Lee, H., Ahn, J., Lee, H., Lee, D. Y., 2010. Genome-scale metabolic reconstruction and in silico analysis of methylotrophic yeast *Pichia pastoris* for strain improvement. Microbial cell factories. 9**,** 50.

Datsenko, K. A., Wanner, B. L., 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proceedings of the National Academy of Sciences of the United States of America. 97**,** 6640-5.

Del Rocío Bustillos-Cristales, M., Corona-Gutierrez, I., Castañeda-Lucio, M., Águila-Zempoaltécatl, C., Seynos-García, E., Hernández-Lucas, I., Muñoz-Rojas, J., Medina-Aparicio, L., Fuentes-Ramírez, L. E., 2017. Culturable Facultative Methylotrophic Bacteria from the Cactus Neobuxbaumia macrocephala Possess the Locus xoxF and Consume Methanol in the Presence of Ce(3+) and Ca(2). Microbes Environ. 32**,** 244-251.

Dowds, B. C., Sheehan, M. C., Bailey, C. J., McConnell, D. J., 1988. Cloning and characterization of the gene for a methanol-utilising alcohol dehydrogenase from Bacillus stearothermophilus. Gene. 68**,** 11-22.

Espinosa, M. I., Williams, T. C., Pretorius, I. S., Paulsen, I. T., 2019. Benchmarking two Saccharomyces cerevisiae laboratory strains for growth and transcriptional response to methanol. Synthetic and Systems Biotechnology. 4**,** 180-188.

Gelius-Dietrich, G., Desouki, A. A., Fritzemeier, C. J., Lercher, M. J., 2013. Sybil--efficient constraint-based modelling in R. BMC systems biology. 7**,** 125-125.

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 CO2. Cell. 179**,** 1255-1263 e12.

Gonzalez, J. E., Bennett, R. K., Papoutsakis, E. T., Antoniewicz, M. R., 2018. Methanol assimilation in *Escherichia coli* is improved by co-utilization of threonine and deletion of leucine-responsive regulatory protein. Metabolic engineering. 45**,** 67-74.

Green, R., Rogers, E. J., 2013. Transformation of chemically competent E. coli. Methods Enzymol. 529**,** 329-36.

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.

Heux S., Brautaset T., Vorholt J.A., Wendisch V.F., J.C., P., 2018. Synthetic Methylotrophy: Past, Present, and Future. In: Kalyuzhnaya M., XH., X., Eds.), Methane Biocatalysis: Paving the Way to Sustainability. Springer, Cham.

http://enerkem.com/fr/.

http://www.methanol.org/.

Huang, Y., Niu, B., Gao, Y., Fu, L., Li, W., 2010. CD-HIT Suite: a web server for clustering and comparing biological sequences. Bioinformatics. 26**,** 680-682.

Joo, H.-J., Ahn, S.-H., Lee, H.-R., Jung, S.-W., Choi, C.-W., Kim, M.-S., Bae, M.-K., Chung, I.-K., Bae, S.-K., Jang, H.-O., Yun, I., 2012. The Effect of Methanol on the Structural Parameters of Neuronal Membrane Lipid Bilayers. Korean J Physiol Pharmacol. 16**,** 255-264.

Kim, S., Lindner, S. N., Aslan, S., Yishai, O., Wenk, S., Schann, K., Bar-Even, A., 2020. Growth of E. coli on formate and methanol via the reductive glycine pathway. Nature chemical biology.

Krog, A., Heggeset, T. M., Muller, J. E., Kupper, C. E., Schneider, O., Vorholt, J. A., Ellingsen, T. E., Brautaset, T., 2013. Methylotrophic Bacillus methanolicus encodes two chromosomal and one plasmid born NAD+ dependent methanol dehydrogenase paralogs with different catalytic and biochemical properties. PloS one. 8**,** e59188.

Los, D. A., Murata, N., 2004. Membrane fluidity and its roles in the perception of environmental signals. Biochimica et Biophysica Acta (BBA) - Biomembranes. 1666**,** 142-157.

Matelbs, R. I., Tannenbaum, S. E., 1968. Single-Cell protein. Economic Botany. 22**,** 42-50.

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.

Millard, P., Delepine, B., Guionnet, M., Heuillet, M., Bellvert, F., Letisse, F., 2019. IsoCor: isotope correction for high-resolution MS labeling experiments. Bioinformatics. 35**,** 4484-4487.

Muller, J. E., Meyer, F., Litsanov, B., Kiefer, P., Potthoff, E., Heux, S., Quax, W. J., Wendisch, V. F., Brautaset, T., Portais, J. C., Vorholt, J. A., 2015. Engineering *Escherichia coli* for methanol conversion. Metabolic engineering. 28**,** 190-201.

Olah, G. A., 2013. Towards oil independence through renewable methanol chemistry. Angewandte Chemie International Edition. 52**,** 104-107.

Orth, J. D., Conrad, T. M., Na, J., Lerman, J. A., Nam, H., Feist, A. M., Palsson, B. Ø., 2011. A comprehensive genome-scale reconstruction of Escherichia coli metabolism--2011. Molecular systems biology. 7**,** 535-535.

Peiro, C., Millard, P., de Simone, A., Cahoreau, E., Peyriga, L., Enjalbert, B., Heux, S., 2019. Chemical and Metabolic Controls on Dihydroxyacetone Metabolism Lead to Suboptimal Growth of Escherichia coli. Applied and environmental microbiology. 85.

Price, J. V., Chen, L., Whitaker, W. B., Papoutsakis, E., Chen, W., 2016. Scaffoldless engineered enzyme assembly for enhanced methanol utilization. Proceedings of the National Academy of Sciences of the United States of America. 113**,** 12691-12696.

R Development Core Team, R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, 2009.

Richter, G., Volk, R., Krieger, C., Lahm, H. W., Rothlisberger, U., Bacher, A., 1992. Biosynthesis of riboflavin: cloning, sequencing, and expression of the gene coding for 3,4-dihydroxy-2-butanone 4-phosphate synthase of Escherichia coli. J Bacteriol. 174**,** 4050-6.

760 Ro, Y. T., Eom, C. Y., Song, T., Cho, J. W., Kim, Y. M., 1997. Dihydroxyacetone synthase
761 from a methanol-utilizing carboxydobacterium, Acinetobacter sp. strain JC1 DSM
762 3803. J Bacteriol. 179**,** 6041-7.
763 Rohlhill, J., Sandoval, N. R., Papoutsakis, E. T., 2017. Sort-Seq Approach to Engineering a
764 Formaldehyde-Inducible Promoter for Dynamically Regulated Escherichia coli Growth
765 on Methanol. Acs Synth Biol. 6**,** 1584-1595.
766 Roth, T. B., Woolston, B. M., Stephanopoulos, G., Liu, D. R., 2019. Phage-Assisted
767 Evolution of Bacillus methanolicus Methanol Dehydrogenase 2. Acs Synth Biol. 8**,**
768 796-806.
769 Russmayer, H., Buchetics, M., Gruber, C., Valli, M., Grillitsch, K., Modarres, G., Guerrasio,
770 R., Klavins, K., Neubauer, S., Drexler, H., Steiger, M., Troyer, C., Al Chalabi, A.,
771 Krebiehl, G., Sonntag, D., Zellnig, G., Daum, G., Graf, A. B., Altmann, F.,
772 Koellensperger, G., Hann, S., Sauer, M., Mattanovich, D., Gasser, B., 2015. Systems-
773 level organization of yeast methylotrophic lifestyle. BMC biology. 13**,** 80.
774 Sangar, V., Blankenberg, D. J., Altman, N., Lesk, A. M., 2007. Quantitative sequence-
775 function relationships in proteins based on gene ontology. BMC Bioinformatics. 8**,**
776 294.
777 Schrader, J., Schilling, M., Holtmann, D., Sell, D., Filho, M. V., Marx, A., Vorholt, J. A., 2009.
778 Methanol-based industrial biotechnology: current status and future perspectives of
779 methylotrophic bacteria. Trends in biotechnology. 27**,** 107-15.
780 Shleev, S. V., Shumakovich, G. P., Nikitina, O. V., Morozova, O. V., Pavlishko, H. M., Gayda,
781 G. Z., Gonchar, M. V., 2006. Purification and characterization of alcohol oxidase from
782 a genetically constructed over-producing strain of the methylotrophic yeast
783 Hansenula polymorpha. Biochemistry (Mosc). 71**,** 245-50.
784 Silva-Rocha, R., Martinez-Garcia, E., Calles, B., Chavarria, M., Arce-Rodriguez, A., de Las
785 Heras, A., Paez-Espino, A. D., Durante-Rodriguez, G., Kim, J., Nikel, P. I., Platero,
786 R., de Lorenzo, V., 2013. The Standard European Vector Architecture (SEVA): a
787 coherent platform for the analysis and deployment of complex prokaryotic
788 phenotypes. Nucleic acids research. 41**,** D666-75.
789 Silva, F., Queiroz, J. A., Domingues, F. C., 2012. Evaluating metabolic stress and plasmid
790 stability in plasmid DNA production by Escherichia coli. Biotechnol Adv. 30**,** 691-708.
791 Simakov, D., 2017. Renewable Synthetic Fuels and Chemicals from Carbon Dioxide.
792 Springer International Publishing.
793 Suzek, B. E., Wang, Y., Huang, H., McGarvey, P. B., Wu, C. H., the UniProt, C., 2014.
794 UniRef clusters: a comprehensive and scalable alternative for improving sequence
795 similarity searches. Bioinformatics. 31**,** 926-932.
796 Team, R., RStudio: Integrated Development for R. RStudio, Inc., Boston, MA, 2015.
797 Vieira, G., Carnicer, M., Portais, J. C., Heux, S., 2014. FindPath: a Matlab solution for *in*
798 *silico* design of synthetic metabolic pathways. Bioinformatics. 30**,** 2986-8.
799 Wang, Y., Fan, L., Tuyishime, P., Zheng, P., Sun, J., 2020. Synthetic Methylotrophy: A
800 Practical Solution for Methanol-Based Biomanufacturing. Trends in biotechnology.
801 Whitaker, W. B., Jones, J. A., Bennett, R. K., Gonzalez, J. E., Vernacchio, V. R., Collins, S.
802 M., Palmer, M. A., Schmidt, S., Antoniewicz, M. R., Koffas, M. A., Papoutsakis, E. T.,
803 2017. Engineering the biological conversion of methanol to specialty chemicals in
804 Escherichia coli. Metabolic engineering. 39**,** 49-59.
805 Whitaker, W. B., Sandoval, N. R., Bennett, R. K., Fast, A. G., Papoutsakis, E. T., 2015.
806 Synthetic methylotrophy: engineering the production of biofuels and chemicals based
807 on the biology of aerobic methanol utilization. Current opinion in biotechnology. 33**,**
808 165-75.
809 Windass, J. D., Worsey, M. J., Pioli, E. M., Pioli, D., Barth, P. T., Atherton, K. T., Dart, E. C.,
810 Byrom, D., Powell, K., Senior, P. J., 1980. Improved conversion of methanol to single-
811 cell protein by *Methylophilus methylotrophus*. Nature. 287**,** 396-401.
812 Woolston, B. M., King, J. R., Reiter, M., Van Hove, B., Stephanopoulos, G., 2018. Improving
813 formaldehyde consumption drives methanol assimilation in engineered *E. coli*. Nat
814 Commun. 9**,** 2387.
815 Wu, T. Y., Chen, C. T., Liu, J. T., Bogorad, I. W., Damoiseaux, R., Liao, J. C., 2016.
816 Characterization and evolution of an activator-independent methanol dehydrogenase
817 from Cupriavidus necator N-1. Appl Microbiol Biotechnol. 100**,** 4969-83.

Yu, H., Liao, J. C., 2018. A modified serine cycle in Escherichia coli coverts methanol and CO2 to two-carbon compounds. Nat Commun. 9**,** 3992.

8. Supplementary Materials

As noted in the text, supplementary materials are available in the online version of this paper. Supplementary Figures contains Figures S1 to S8. Figure S1 shows and overview of the recycling of the C1 acceptor Xu5P and its operation. Figure S2 shows the unrooted phylogenetic trees of selected Mdh and Das homologues.Figure S3 shows the overall scheme of the combinatorial assembly and screening of the synthetic pathway. Figure S4 shows the expression analysis of *B. methanolicus* Mdh and *P. angusta* Das in different conditions. Figure S5 shows the western Blot analysis of expression of Mdh and Das homologues.Figure S6 shows the ^{13}C-Methanol assimilation in the methylotrophic *E. coli* Δ*frmA* expressing the synthetic pathway from one or two vectors. Figure S7 shows the ^{13}C-Methanol assimilation into central metabolism intermediates in the control strains of the genealogy of methylotrophic *E. coli*. Figure S8 shows ^{13}C-Methanol assimilation into proteinogenic amino acids in the genealogy of methylotrophic *E. coli*.

Supplementary Tables contains Tables S1 to S5. Table S1 is the list of selected Mdh and Das homologues and the associated optimal growth temperature range of the source organisms. Table S2 contains the mean isotopic enrichment of PEP in % using the combinatorial library. Table S3 contains the mean isotopic enrichment of PEP using different Das enzymes with codon-optimized sequences in combination with MDH from *A. gerneri*. Table S4 contains an overview of the ^{13}C-enrichement obtained in different synthetic methylotrophic strains from previous studies. Table S5 is the list of stains and plasmids used in this stu

Mixing and matching methylotrophic enzymes to design a novel "hybrid" metabolic pathway for methanol assimilation in *E. coli*

De Simone A.[1]; Vicente C.M.[1]; Peiro C.[1]; Gales L.[1,2], Bellvert F.[1,2]; Enjalbert B.[1]; Heux S.[1*]

[1]TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France

[2]MetaboHUB-MetaToul, National Infrastructure of Metabolomics and Fluxomics, Toulouse, 31077, France

* Corresponding author. Mailing address TBI, INSA Toulouse, 135 Avenue de Rangueil, 31077 Toulouse cedex 4, France. Fax: +33 561 55 96 89. E-mail: stephanie.heux@insa-toulouse.fr

Tables 1 and Figures 1 to 4.

Conditions	Growth rate	Specific xylose uptake rate	Specific formate production rate
+ Methanol	0.17+/-0.00	3.00+/-0.15	0.04+/-0.02
- Methanol	0.11+/-0.01	2.09+/-0.03	0

Table 1: Physiological response of the new synthetic methylotroph to methanol. Growth rate (h^{-1}), specific consumption and production rates ($mmol/g_{DW}/h$) of the *E. coli* ΔfrmA_pSEVA424-Mdh-Das(opt) strain during growth in M9 minimal media containing 15 mM xylose without methanol (- Methanol) and supplemented with 150 mM methanol (+ Methanol). Mean and standard deviation of two replicates are given.

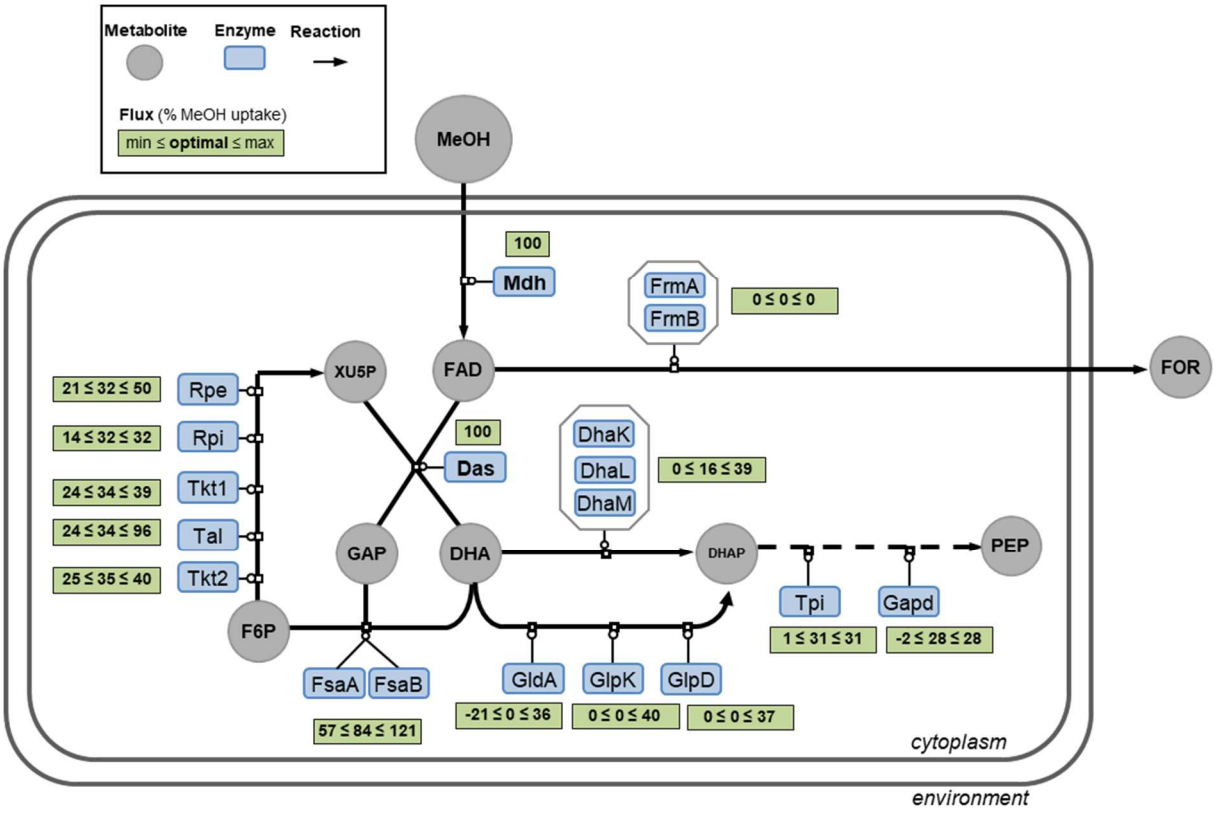

Figure 1. Overview of the synthetic methanol metabolism and its operation in *E. coli*. The new hybrid methanol assimilation pathway comprises a methanol dehydrogenase (Mdh) and a dihydroxyacetone synthase (Das). Green rectangles give the optimal and the ranges of simulated fluxes obtained using flux balance analysis and flux variability analysis, respectively, when growth rate is constrained to 90% of the optimal value. Flux values are given in % relative to a MeOH uptake rate of 15 mmol/gDW/h as defined in Peyraud et al., BMC Syst Biol. 2011.

Dihydroxyacetone kinase (DhaK, DhaL and DhaM); Glycerol dehydrogenase (GldA); Glycerol-3-phosphate dehydrogenase (GlpD); Glycerol kinase (GlpK); Fructose-6-phosphate aldolase (FsaA and FsaB); Triose phosphate isomerase (Tpi); Glyceraldehyde-3-phosphate dehydrogenase (Gapd); Ribulose 5-phosphate 3-epimerase (Rpe); Transketolase (Tkt1 & Tkt2); Transaldolase (Tal), Ribose-5-phosphate isomerase (Rpi); Formaldehyde dehydrogenase (FrmA); S-Formylglutathione hydrolase (FrmB); Methanol (MeOH); Formaldehyde (FAD); Xylulose-5-P (XU5P); Glyceraldehyde-3-phosphate (GAP); Dihydroxyacetone (DHA); Phosphoenolpyruvate (PEP); Dihydroxyacetone phosphate (DHAP); Fructose-6-phosphate (F6P), Formate (FOR).

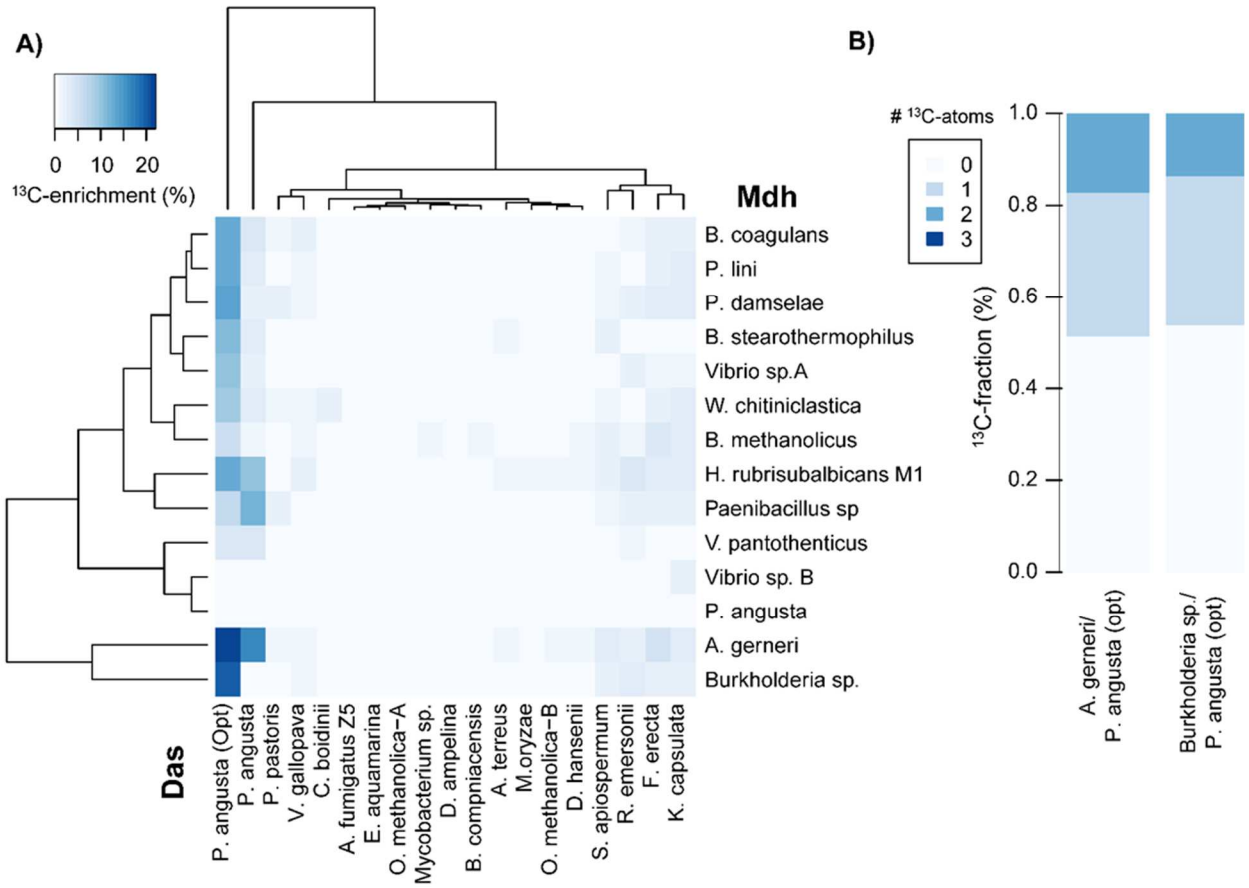

Figure 2. Screening of the different combinations of the hybrid methylotrophic pathway. (A) Heatmap showing the ^{13}C-enrichment of phosophoenlypyruvate (PEP) in *E. coli* Δ*frmA* expressing different combinations of Mdh and Das homologues. The ^{13}C-enrichment of PEP was measured at the steady state during exponential growth (90 min after cultivation in M9 medium containing 655 mM of ^{13}C-methanol). Rows and columns are ordered according to the cluster trees shown on the left and on the top. The Euclidean function was used as distance metric and complete linkage was used as clustering algorithm. (B) Labeling pattern of PEP at 90 min in *E. coli* Δ*frmA* expressing *P. angusta* Das (opt) with either *A. generi* Mdh or *Burkholderia* Mdh.

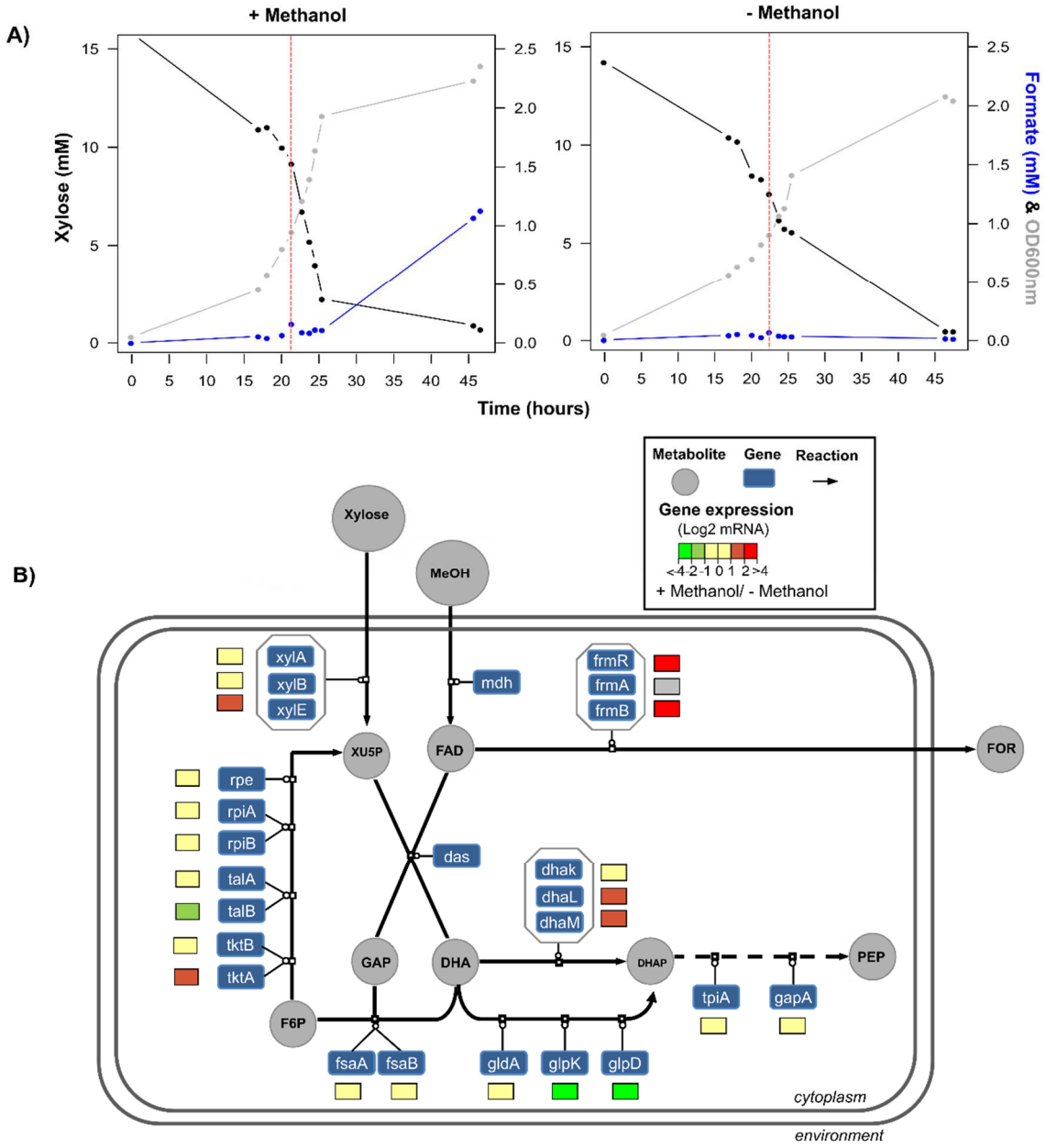

Figure 3. Response of the new synthetic methylotroph to methanol. A) Time course analysis of xylose (black), formate (blue) and biomass (gray). Red dotted lines indicates the sampling point for transcriptomic analysis. Cells were grown in a minimal synthetic medium containing 15mM xylose with or without 150 mM methanol at 30°C and 250 rpm. One representative experiment for each condition is shown (n=2). **B)** Gene expression profiling of the *E. coli* Δ*frmA*_pSEVA424-Mdh-Das (opt). The coloured squares represent the log2-ratios as measure of gene expression fold changes (+ Methanol / - Methanol) during exponential growth at OD_{600} = 1 (T1).

Methanol dehydrogenase (*mdh*); Dihydroxyacetone synthase (*das*); Glutathione-dependent formaldehyde detoxification operon (*frmRAB*); Dihydroxyacetone kinase operon (*dhaKLM*); Fructose-6-phosphate aldolase isoform A and B (*fsaA, fsaB*); Glycerol dehydrogenase (*gldA*); glycerol kinase (*glpK*); Glycerol-3-phosphate dehydrogenase (*glpD*); Transketolase isoforms A and B (*tktA, tktB*); Transaldolase isoforms A & B (*talA, talB*); Ribose phosphate isomerase isoforms A & B (*rpiA, rpiB*);

Ribulose phosphate epimerase (*rpe*); Triose phosphate isomerase (*tpiA*); Glyceraldehyde 3-phosphate dehydrogenase (*gapA*); Xylose isomerase (*xylA*); Xylulokinase (*xylB*); D-xylose/proton symporter (*xylE*); Methanol (MeOH); Formaldehyde (FAD); Xylulose-5-P (XU5P); Glyceraldehyde-3-phosphate (GAP); Dihydroxyacetone (DHA); Phosphoenolpyruvate (PEP); Dihydroxyacetone phosphate (DHAP); Fructose-6-phosphate (F6P); Formate (FOR).

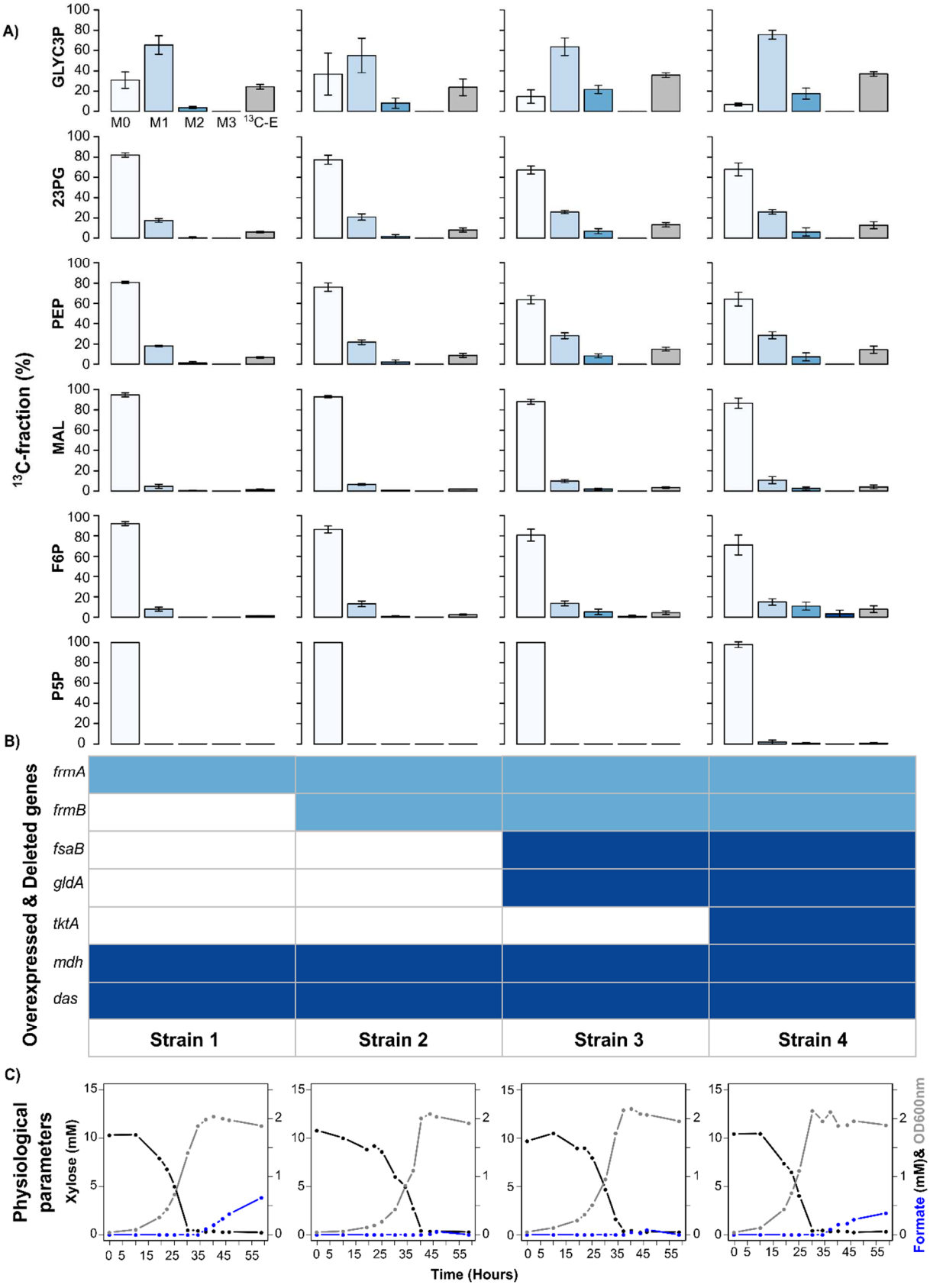
A)
GLYC3P
23PG
PEP
MAL
F6P
P5P
M0 M1 M2 M3 13C-E
13C-fraction (%)
B)
Overexpressed & Deleted genes
frmA
frmB
fsaB
gldA
tktA
mdh
das
Strain 1
Strain 2
Strain 3
Strain 4
C)
Physiological parameters
Xylose (mM)
Formate (mM)& OD600nm
Time (Hours)

Figure 4. Phenotypic characterization of rationally designed *E. coli* strains. A) Labeling patterns of the intracellular metabolites 2 & 3 phosphoglycerate (23PG), fructose-6-phospohate (F6P), glycerol-3-phosphate (GLYC3P), pool of pentoses-5-phosphate (P5P), phosphoenolpyruvate (PEP) and malate (Mal) within the different strains after 90 min of culture in M9 medium with 655 mM ^{13}C-methanol. Mean and standard deviation of of 3 replicates are shown. Fraction with zero (M0), one (M1), two (M2) , three (M3) ^{13}C-atom and 13C-enrichment (^{13}C-E) are show. **B)** Table of overexpressed (dark blue) and deleted (light blue) genes within the different strains. **C)** Time course analysis of xylose (black), formate (blue) and biomass (gray) within the different strains. Strains were grown in a minimal synthetic medium with 15 mM xylose and 150 mM ^{13}C-methanol at 30°C and 250 rpm. One representative experiment for each strain is shown (n=2).