



HAL
open science

Engineering and compartmentalisation of a synthetic methanol utilisation pathway in *E. coli*.

Camille Peiro

► **To cite this version:**

Camille Peiro. Engineering and compartmentalisation of a synthetic methanol utilisation pathway in *E. coli*. Biochemistry [q-bio.BM]. INSA de Toulouse, 2021. English. NNT: 2021ISAT0043 . tel-04213514

HAL Id: tel-04213514

<https://theses.hal.science/tel-04213514>

Submitted on 21 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE

**En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE
Délivré par l'Institut National des Sciences Appliquées de
Toulouse**

**Présentée et soutenue par
Camille PEIRO**

Le 18 mars 2021

**Ingénierie et compartimentation d'une voie synthétique
d'utilisation du méthanol chez E. coli**

Ecole doctorale : **SEVAB - Sciences Ecologiques, Vétérinaires, Agronomiques et
Bioingenieries**

Spécialité : **Ingénieries microbienne et enzymatique**

Unité de recherche :

TBI - Toulouse Biotechnology Institute, Bio & Chemical Engineering

Thèse dirigée par

Jean-charles PORTAIS et Stéphanie HEUX

Jury

M. Pau FERRER, Rapporteur

Mme Carole CAMARASA, Rapporteuse

Mme Anne-Marie DELORT, Examinatrice (Présidente du jury)

M. Diethard MATTANOVICH, Examineur

M. Jean-Charles PORTAIS, Directeur de thèse

Mme Stéphanie HEUX, Co-directrice de thèse

I began to realize how important it was to be an enthusiast in life. If you are interested in something, no matter what it is, go at it fullspeed. Embrace it with both arms, hug it, love it and above all become passionate about it. Lukewarm is no good.

- Roald Dahl

Acknowledgments

Je tiens tout d'abord à remercier Pau Ferrer, Carole Camarasa, Anne-Marie Delort et Diethard Mattanovich d'avoir accepté d'être membre de mon jury de thèse et d'avoir évalué ce travail.

Merci à Stéphanie et Jean-Charles, mes co-directeurs de thèse pour la relecture de ce manuscrit. Je tiens particulièrement à vous remercier pour l'influence que vous avez eue sur moi au cours de ces dernières années. Jean-Charles, tout d'abord merci de m'avoir fait découvrir la métabolomique à l'université, tes cours ont été de toute évidence les premières pierres qui m'ont aidée à bâtir mon chemin professionnel. Merci encore pour tes précieux conseils lorsque j'étais étudiante, sans eux, je n'aurais sûrement jamais sauté le pas pour faire une thèse. Stéphanie, quasiment 4 ans de thèse se terminent certes, mais c'est 6 ans passés ensemble qui se concluent. Je suis heureuse que mes stages aient été les prémices de cette thèse. Merci de m'avoir laissée toute la liberté dont j'avais besoin pour m'épanouir pleinement dans ce projet. Lors de mes moments de doutes, tu as toujours su être là pour m'épauler. Merci pour ton écoute, ta confiance et surtout, ton optimisme sans faille qui nous auront permis de mener ce projet aussi loin (on n'aura tout de même pas réussi à l'emmener jusqu'à Hawaï..!). Cette thèse aura été, on peut le dire, une première pour toutes les deux !

Merci à MetaSys et MétaToul (ou parfois malencontreusement appelés MetaSous). Merci à Pietro, Cláudia, Pierre, Philippe, Denis, Fabien, Sandrine, Florence, Cécilia, Maud, Edern, Loïc, Noémie, Nina, Lara, Floriant, Lindsay, Hanna, Amandine mais aussi à Alessandro, Mickaël, Baudoin, Clément, Matthieu et Brice pour toutes ces discussions, scientifiques ou non. Vous avez tous contribué à votre manière à faire en sorte que cette thèse se passe si bien. Chaque moment passé avec vous au labo ou lors des (trop ?) nombreuses soirées au Dubli a toujours été un réel plaisir à partager avec vous! Que serait le labo sans les runs entre midi et deux avec Neil et Pietro, les Jeudi Bières, les Methylo meetings, les gâteaux pour les anniversaires et pour finalement toutes les circonstances car on sait bien que toutes les occasions sont bonnes pour manger dans cette équipe ! Edern, ton aide lors de mes manip RMN a été plus que précieuse, merci ! On retiendra tout de même qu'il n'est pas bon de laisser un paquet de schtroumpfs sur le bureau de la RMN toute une nuit..! Matthieu, quel honneur d'avoir pu être ton bras droit pendant quelques temps! Nos « grandes conversations » et nos nombreux fous rires ont évidemment marqué ma thèse. Maud et Cécilia, bravo à vous, vous êtes les premières à avoir réussi à m'avoir fait accepter « Camillou » comme surnom, bien évidemment ce petit privilège vous est exclusivement réservé ! Les 133, je ne suis pas arrivée dans le bureau pour la partie la plus fun de la thèse mais votre patience et bonne humeur ont rendu la rédaction un peu plus facile. Après plus de 3 ans à partager le labo et le bureau avec toi, Pietro, cela a toujours été synonyme de rire (mais aussi de science (un peu) avec les nombreux samplings). Je n'allais tout de même pas te mentionner sans parler de tes légendaires memes et montages en tout genre qui m'ont tellement fait rire. Tu resteras le maître inconditionné dans cette catégorie pour moi!

Neil et Gilles, quoi de mieux que de partager le quotidien de thésard qu'avec des thésards ?! On aura passé les bons et les mauvais moments de nos thèses ensemble, et je suis contente que ce soit avec vous que j'ai pu partager les miens. Les pauses café et les pauses déj qui finissent en parties de domino ont été des rendez-vous immanquables dans le

quotidien du labo. Merci pour toutes ces discussions, fou rires et soirées. Bref, vous savez, ce qui fait que nous sommes les Trois Mousquetaires de la thèse.

Merci à mes amis: Maryne, Mayou, Guigui, Cyril, Cha', Nico', Victor et Guth. Merci pour ces innombrables moments et soirées que nous avons passés ensemble un partout en France, et ailleurs ! Vous êtes de réels piliers qui avez fait que la vie hors du labo est si sympa ! Also, a special thanks to Rozanne, my time in Canterbury wouldn't have been the same without you. I am glad we met and shared a part of this PhD journey together, I can't wait to see you again (and share G&T's)! Bertrand, may l'ordre cosmique be with us.

Enfin, merci à ma famille. A mes parents, Manuel et Luce pour votre soutien indéfectible. Vous avez toujours su me soutenir et me conforter dans mes choix. Merci pour votre aide et pour votre patience pendant mes nombreux moments de remise en question « à la Camille-SAFE » comme dirait Wazou ! Aussi, une pensée particulière à ceux qui ne sont plus là et qui doivent être bien heureux de cet accomplissement.

Résumé

Pour surmonter le problème rencontré dans la biotechnologie blanche concernant l'utilisation de sources carbonées spécifiques pour la production microbienne de composés d'intérêt, de nouveaux substrats plus durables et plus rentables sont nécessaires. Les composés à un seul carbone (C1) tels que le méthanol ont récemment gagné en intérêt comme substituts aux substrats conventionnels qui sont en concurrence avec l'alimentation humaine. Le méthanol peut à présent être produit de manière renouvelable à partir d'électricité, d'eau et de CO₂. Les méthylotrophes naturels peuvent croître en assimilant le méthanol mais leur utilisation lors de procédés industriels est encore limitée par un manque d'outils génétiques. Par conséquent, des efforts ont été faits pour générer des méthylotrophes synthétiques en transférant ce trait métabolique chez des hôtes de production bactériens connus (tels que *E. coli* ou *C. glutamicum*). Ici, nous avons modifié, introduit et compartimenté une nouvelle voie synthétique d'utilisation du méthanol (MUT) chez *E. coli*. La voie hybride MUT est composée d'une méthanol déshydrogénase (Mdh) d'origine bactérienne dépendante du NAD et d'une dihydroxyacétone synthase (Das) provenant d'une levure. Le réseau métabolique d'*E. coli* a tout d'abord été modifié pour s'adapter aux propriétés méthylotrophiques acquises. Le dihydroxyacétone (DHA), un métabolite résultant de l'activité de la Das, est une source naturelle de carbone pour *E. coli* mais son métabolisme natif reste inconnu. En utilisant des outils de modélisation, nous avons constaté que le métabolisme du DHA n'était pas optimal chez *E. coli*. De plus, nous avons pu déterminer quelles enzymes étaient impliquées dans le métabolisme du DHA et avons identifié certains verrous métaboliques limitant l'assimilation du DHA grâce à une analyse à l'échelle du système. Nous avons aussi pu identifier des cibles génétiques permettant d'améliorer l'assimilation du DHA. Ensuite, la voie MUT a été introduite chez *E. coli*. Nous l'avons construite en élaborant une bibliothèque combinatoire de gènes provenant de divers organismes codant pour *mdh* et *das* et l'avons testée en suivant l'incorporation du ¹³C-méthanol au niveau du PEP (un proxy pour suivre l'assimilation du méthanol). Le plus haut niveau d'incorporation de méthanol dans les intermédiaires du métabolisme central (22% dans le PEP) a été obtenu en utilisant la combinaison composée de la Mdh d'*A. gernerii* et d'une version codon optimisée de la Das de *P. angusta*. Puis, l'assimilation du méthanol a été testée dans le châssis construit précédemment pour une assimilation améliorée du DHA. La souche synthétique finale a une assimilation du méthanol 1,5 à 5,9 fois plus élevée dans les métabolites intracellulaires et les acides aminés protéinogéniques que la souche de départ. Cependant, la souche finale ne pouvait toujours pas pousser sur du méthanol pur. Enfin, pour optimiser spatialement le fonctionnement de la voie MUT *in vivo*, nous avons construit un microcompartiment bactérien méthylotrophique (mBMC) contenant la meilleure combinaison de la Mdh et de la Das. Nous avons comparé le fonctionnement de la voie MUT seule ou encapsulée *in vitro*. Nous avons démontré que la conversion du substrat était améliorée lorsque la voie MUT était encapsulée dans le BMC.

Dans l'ensemble, ces travaux ont élargi le répertoire des voies synthétiques permettant l'assimilation du méthanol, tout en apportant de nouvelles connaissances sur l'établissement de la méthylothrophie synthétique chez *E. coli*. Ils représentent également une base de départ pour l'utilisation des BMCs comme micro-réacteurs modulables pour l'encapsulation de voies métaboliques synthétiques avec des applications directes pour l'assimilation du méthanol.

Abstract

To overcome the issue faced in white biotechnology of using specific carbon sources for the microbial production of value-added compounds, more sustainable and cost-efficient substrates are needed. One-carbon (C1) compounds such as methanol have recently gained attention as substitutes for substrates competing with conventional food supply. Methanol can now be produced renewably from electricity, water or CO₂. Natural methylotrophs can grow on methanol but their full implementation in industrial processes is still limited by a lack of genetic tools. Therefore, efforts have been made to generate synthetic methylotrophs by transferring this metabolic trait into established bacterial production hosts (e.g. *E. coli* or *C. glutamicum*). Here, we shaped, implemented and compartmentalised a novel synthetic methanol utilisation (MUT) pathway in *E. coli*. The hybrid MUT pathway is composed of a bacterial NAD-dependent methanol dehydrogenase (Mdh) and a yeast dihydroxyacetone synthase (Das). First, we shaped the metabolic network of *E. coli* to fit the acquired methylotrophic properties. Dihydroxyacetone (DHA), a metabolite resulting from Das activity, is a natural carbon source for *E. coli* but its native metabolism remains unclear. Using constraint-based modelling, we found that DHA metabolism was suboptimal in *E. coli*. In addition, we deciphered enzymes involved in DHA metabolism and identified some bottlenecks in DHA assimilation by a system-level analysis. We identified genetic targets enabling to improve DHA assimilation. Next, we implemented the MUT pathway in *E. coli*. We constructed a combinatorial library of *mdh* and *das* encoding genes coming from various organisms and tested it by following ¹³C-methanol incorporation at the PEP-level (i.e. a proxy for methanol assimilation). The highest level of methanol incorporation into central metabolism intermediates (e.g. 22% into the PEP) was obtained using a variant composed of Mdh from *A. gernerii* and a codon-optimized version of Das from *P. angusta*. Methanol assimilation was tested in the chassis built previously with improved DHA assimilation. 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. However, the resulting strain did not grow on pure methanol. Finally, to spatially optimize the *in vivo* operation of the MUT pathway, we constructed a methylotrophic bacterial microcompartment (mBMC) containing the selected combination of Mdh and Das. We compared the *in vitro* operation of the MUT pathway alone and encapsulated in the BMC. We demonstrated that methanol conversion was improved when the MUT pathway was encapsulated into the BMC.

Overall, these results broadened the repertoire of synthetic pathways for methanol assimilation while bringing new knowledge on the establishment of synthetic methylotrophy in *E. coli*. It also represents a groundwork for the utilization of BMCs as versatile micro-reactors for the encapsulation of synthetic metabolic pathways with direct applications for the assimilation of methanol.

Table of contents

Abstract	7
Table of contents	9
Abbreviations	11
Chapter 1 – General introduction	13
Chapter 2 – Chemical and metabolic controls on dihydroxyacetone metabolism lead to a suboptimal growth of <i>E. coli</i>	71
Chapter 3 – Mixing and matching methylotrophic enzymes to a design a novel methanol utilization pathway in <i>E. coli</i>	109
Chapter 4 – Encapsulation of a synthetic methyltrophic pathway in a bacterial microcompartment	153
Conclusion and outlooks	191

Abbreviations

ALE: Adaptive laboratory evolution	Mdh: Methanol dehydrogenase
AO: Alcohol oxidase	MUT: Methanol utilization pathway
BMC: Bacterial microcompartment	NAD(H): Nicotinamide adenine dinucleotide
eBMC: empty bacterial microcompartment	NADP(H): Nicotinamide adenine dinucleotide phosphate
mBMC: methylotrophic bacterial microcompartment	NMR: Nuclear magnetic resonance
C1: one carbon	OD: Optical density
CBB: Calvin Benson Bassahm	P18: N-terminal 18 amino acids of PduP
D18: N-terminal 18 amino acids of PduD	Pdu: 1,2-Pronapendiol utilization
DAK: Dihydroxyacetone kinase pathway	PEP: Phosphoenol-pyruvate
Das: Dihydroxyacetone synthase	Phi: 6-phospho-3-hexuloisomerase
DHA: Dihydroxyacetone	PPP: Pentose phosphate pathway
EMC: Ethylmalonyl-CoA	PQQ: Pyrroloquinoline quinone
Eut: Ethanolamine utilization	Rpe : ribulose-phosphate 3-epimerase
F6P: Fructose-6-phosphate	Rpi : ribose-5-phosphate isomerase
FSA: Fructose-6-phosphate aldolase pathway	RuMP: Ribulose monophosphate
Hps: 3-hexulose-6-phosphate synthase	SDS-PAGE: Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
Ga3P: Glyceraldehyde-3-phosphate	TCA cycle: Tricarboxylic acid cycle
GLD: Glycerol pathway	TEM: Transmission electron microscopy

XuMP: Xylulose monophosphate

Xu5P: Xylulose-5-phosphate

Chapter 1

General introduction

Since few years, we are witnessing our society aiming to become more independent from fossil fuel-based processes by replacing them by more sustainable bio-based ones. In this effort, thinking about microbial cell factories is a good alternative to produce key chemical, food or health care products. However, to grow and produce those value-added compounds, microorganisms are using sugars or food-derived raw substrates. In the long term, it might lead to an unwanted competition with the food supply. One challenge is to find low-cost and easy to use fermentation substrates. That is why the possibility to use reduced one-carbon (C1) substrates such as methanol to overcome this problem starts to rise. Methanol caught people's attention over the past years as a new economy can be generated by using it as a feedstock for biorefineries (Olah, 2013; Schrader et al., 2009). Consequently, in biotech, a new field of research is in development, based on the use of methanol or other C1-compounds as carbon source for natural methylotrophs and synthetic methylotrophs (i.e. production hosts in which methylotrophy has been integrated) to produce compounds of interest. Although, methanol is mainly produced from fossil resource, new processes are emerging enabling the production of bio-methanol from any renewable resource as long as this resource can be converted into synthesis gas (i.e. syngas). Nowadays, new techniques are developed to produce bio-methanol using thermo-, electro- or photo-catalytic reduction of CO₂ (Simakov, 2017). Other substrates as biomass, agricultural and timber waste are also used for the production of biomethanol (<http://enerkem.com/fr/>; <https://www.methanol.org/>). In addition to its wide availability, another advantage to use methanol as a carbon source for biotech purposes is that its degree of reduction is higher compared to glucose, which leads to higher product yields (Villadsen et al., 2011).

I. Methylotrophy

Methylotrophy is defined as the capacity of an organism to use reduced C1 compound, such as methanol, as its sole carbon and energy source. In nature, methylotrophs are a wide group of microorganisms, including among others, methylotrophic bacteria and yeasts. Among these kingdoms, microorganisms had to adapt to different lifestyles and ecological niches which led them to evolve and to set up a diverse range of pathways and enzymes to grow on methanol and to ensure their survival (Figure 1). All these pathways include processes

(i) to oxidize methanol to formaldehyde (ii) to oxidize formaldehyde to CO_2 and (iii) to assimilate a C1-intermediate (i.e. formaldehyde, $\text{CH}_2\text{H}_4\text{F}$ or CO_2) to biomass. Four different assimilation pathways are used to ensure growth: Calvin-Benson-Bassham (CBB) cycle, Serine Cycle, Ribulose MonoPhosphate (RuMP) and the Xylulose MonoPhosphate (XuMP) pathways.

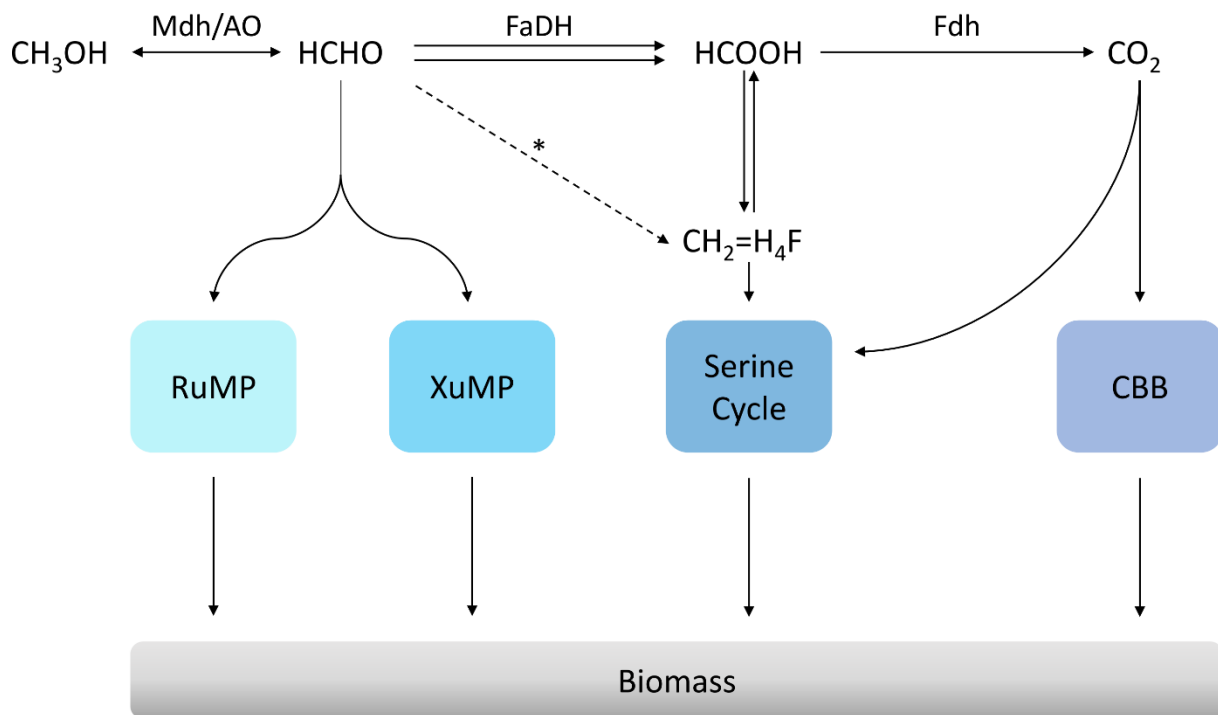


Figure 1: Natural pathways to use methanol. Adapted from (Wang et al., 2020).

*CBB, Calvin-Benson-Bassham cycle; RuMP, Ribulose MonoPhosphate pathway; XuMP, Xylulose MonoPhosphate pathway. AO, alcohol oxidase; FaDH, formaldehyde dehydrogenase; Fdh, formate dehydrogenase; MDH, methanol dehydrogenase. * indicates spontaneous reaction*

A. Methanol oxidation to formaldehyde

Methanol oxidation to formaldehyde is the first step in methanol assimilation or dissimilation. This reaction is ensured by oxidoreductases named methanol dehydrogenase (Mdh) in bacteria and alcohol oxidase (AO) in yeasts. Two types of Mdhs are described: Pyrroloquinoline quinone (PQQ)-dependent Mdhs found in Gram-negative bacteria and Nicotinamide Adenine Dinucleotide (NAD)-dependent Mdhs commonly found in Gram-positive bacteria, while AOs use oxygen as an electron acceptor.

1. PQQ-dependent MDHs

Gram-negative methylotrophs possess multimeric PQQ-Mdhs located in the periplasm. These proteins are composed by two subunits: one large and one small, respectively encoded by the genes *mxoF* and *mxoI* (Goodwin and Anthony, 1998). However, some bacteria such as *Burkholderia* spp. do not possess these genes but have instead a *mdh2* gene that encodes for a monomeric PQQ-Mdh (Kalyuzhnaya et al., 2008). Recently, a third type of PQQ-Mdh encoded by *xoxF* gene was described. It turned out that XoxF are widely present in the PQQ-Mdhs phylogenetic tree and unlike MxaFI, XoxF are not using Ca^{2+} as a cofactor but specifically bind a rare element : Lanthanide (Ln^{3+}) (Keltjens et al., 2014). So far, MxaFI are the best characterized PQQ-Mdhs. PQQ is a prosthetic group capturing electrons from methanol oxidation to transfer them specifically to cytochrom c (Goodwin and Anthony, 1998). PQQ connects redox reactions to the respiratory chains (Matsushita et al., 1997). PQQ-Mdhs have a higher affinity for methanol than for higher alcohols (Kalyuzhnaya et al., 2008).

2. NAD-dependent MDHs

Thermophilic Gram-positive of the *Bacillus* spp use NAD-dependent Mdh. Unlike PQQ-Mdhs, NAD-Mdhs are located in the cytoplasm (Arfman et al., 1989). Among all, NAD-Mdhs from *B. methanolicus* and *B. stearothermophilus* are the best described (Krog et al., 2013; Sheehan et al., 1988). Interestingly in *B. methanolicus*, there are three Mdhs encoded by three different genes: two are chromosomal and the third one is carried on a natural plasmid (Heggeset et al., 2012). pBM19 and pBM20 are the plasmids respectively found in *B. methanolicus* MGA3 and PB1. Other genes involved in methanol assimilation are present in these plasmids. NAD-Mdhs are decameric proteins. In each of the subunit, one molecule of NAD(H), one Zn^{2+} and two Mg^{2+} cofactors are found. NAD-Mdhs interact with an activator protein (ACT) which facilitates the oxidation of NADH (Arfman et al., 1997). *In vitro*, the activity of NAD-Mdhs can be enhanced by ACT (Krog et al., 2013) however it does not seem to be effective *in vivo* (Müller et al., 2015a). The *act* gene encoding ACT is found on *B. methanolicus* chromosome. It was shown, *in vitro* that the three Mdhs from *B. methanolicus* have more affinity for formaldehyde ($K_m \leq 7$ mM) than for methanol ($K_m \geq 170$ mM) (Krog et al., 2013). At the end of the 80's, a new alcohol dehydrogenase has been described in *B. stearothermophilus* (Sheehan et al., 1988). This enzyme was shown to have a better affinity for methanol ($K_m=20$ mM) than the Mdh from *B. methanolicus*. This was confirmed during *in vivo* experiment where

the alcohol dehydrogenase from *B. stearothermophilus* showed a better performance than the tested Mdh from *B. methanolicus* (Whitaker et al., 2017).

3. Oxygen-dependent AO

In eukaryotic methylotrophs, AO is specifically located in the peroxisome. AO is an homo-octameric protein: each inactive AO monomer is synthesised in the cytoplasm and then are assembled together to form the active octomeric AO protein in the peroxysome. Each subunit contains one molecule of flavin adenosine dinucleotide that is not bound covalently (Sahm and Wagner, 1973). During growth on methanol, AO can account for up to 30% of the total cellular protein (Ozimek et al., 2003). During methanol oxidation, a toxic compound other than formaldehyde is produced: hydrogen peroxide (H_2O_2). H_2O_2 is detoxified into oxygen and water by a catalase (CAT), also located in the peroxisome.

B. Formaldehyde dissimilation to CO_2

Formaldehyde presence can not be tolerated in cells as it is a highly toxic compound that nonspecifically interacts with proteins and nucleic acid damaging the cell integrity. Formaldehyde can be produced from methanol oxidation but also from demethylation reactions, hence facing formaldehyde toxicity is a common challenge faced by cells. If they can cope with low formaldehyde concentrations, any accumulation is necessarily resulting in the activation of detoxification processes. Cofactors, such as glutathione (GSH) (or mycothiol (MySH)), tetrahydrofolate (H_4F) and tetrahydromethanopterin (H_4MPT), trap formaldehyde to efficiently oxidize it. Depending on the cofactor used, different pathways exist to ensure formaldehyde dissimilation (Figure 2). Sometimes, various dissimilation pathways are existing in the same organism (Vorholt, 2002). All of them convert formaldehyde to formate which can be ultimately converted to CO_2 .

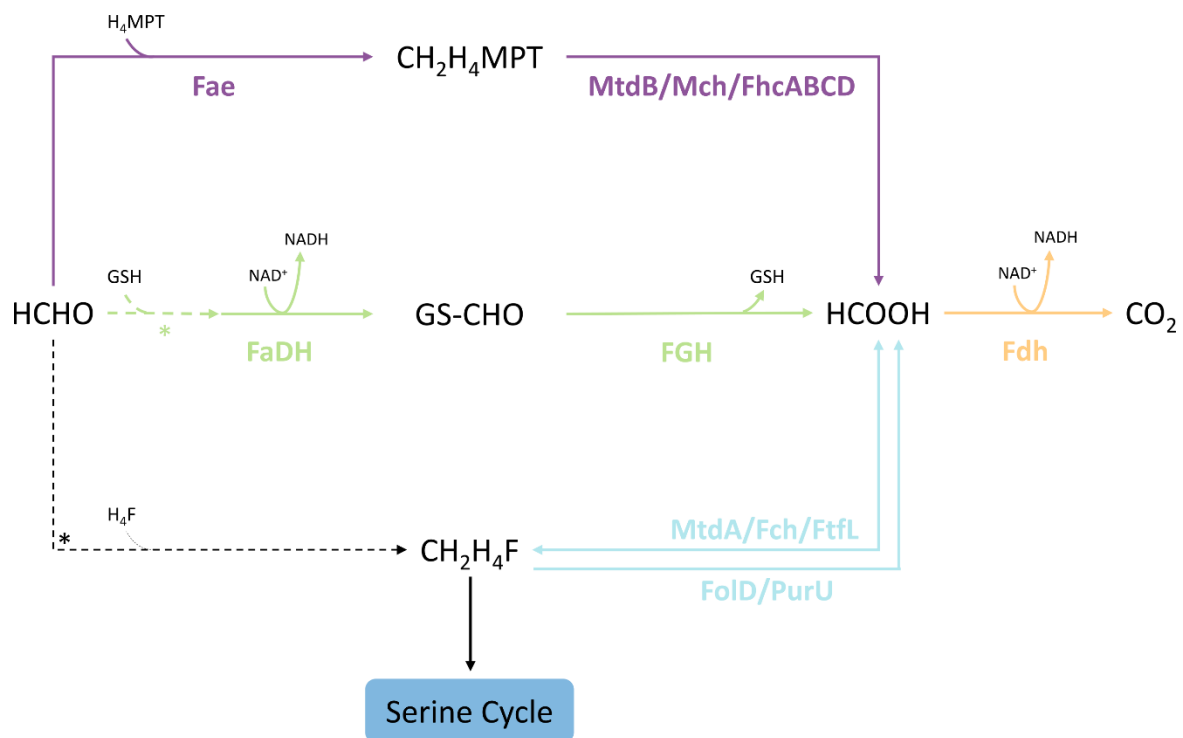


Figure 2: Pathways for the dissimilation of formaldehyde into CO_2 found in methylotrophs. Formaldehyde oxidation by H_4MPT - (in purple), H_4F - (in blue) or GSH - (in green) pathways. Formate oxidation to CO_2 by Fdh (in orange). Adapted from (Chistoserdova, 2011). Enzymes are written in bold.

FaDH, formaldehyde dehydrogenase; *Fae*, formaldehyde-activating enzyme; *Fch*, methenyl- H_4F -cyclohydrolase; *FhcABCD*, methenyl- H_4F cyclohydrolase; *Fdh*, formate dehydrogenase; *FGH*, *S*-formylglutathione hydrolase; *FOLD*, methylene- H_4F dehydrogenase/ methenyl- H_4F cyclohydrolase; *FtFL*, formate- H_4F ligase; *Mch*, methenyl- H_4MPT cyclohydrolase; *MtdA*, methylene- H_4F dehydrogenase/ methylene- H_4MPT dehydrogenase; *MtdB*, methylene- H_4MPT dehydrogenase; *PurU*, formyl- H_4F hydrolase.

1. H_4MPT -dependent pathway

H_4MPT pathway is one of the most elaborated pathways to oxidize formaldehyde to formate. It is a widespread pathway, commonly found in obligate anaerobic archae, methanogens and acetogens (Chistoserdova, 1998). However, this pathway was first described in the aerobic mesophilic bacterium *Methylobacterium extorquens* AM1 (Chistoserdova, 1998). Formaldehyde condensation with H_4MPT to methylene- H_4MPT ($\text{CH}_2\text{H}_4\text{MPT}$) can occur spontaneously but is catalyzed by a formaldehyde-activating enzyme (*Fae*). Then, $\text{CH}_2\text{H}_4\text{MPT}$ is converted to formate *via* a series of enzymatic reactions. *MtdB* and *MtdA* can both oxidize $\text{CH}_2\text{H}_4\text{MPT}$ to methenyl- H_4MPT . However, it was shown that *MtdB* is the main enzyme involved in this pathway (Chistoserdova et al., 2009). *MtdB* is specific to H_4MPT but can either use NAD or NADP as cofactors (Hagemeier et al., 2000).

2. H₄F-dependent pathway

H₄F reacts non enzymatically with formaldehyde to generate methylene-H₄F (CH₂H₄F) which can either be used in the serine cycle for biosynthesis or enter the oxidation pathway to be converted to formate (Vorholt, 2002). The bifunctional methylene-H₄F dehydrogenase/methenyl-H₄F cyclohydrolase (FOLD) is associated with formyl-H₄F hydrolase (PurU) to ensure CH₂H₄F conversion to formate. However, some methylotrophs lack FOLD and are using instead two enzymes: a methylene-H₄F dehydrogenase (MtdA), and a methenyl cyclohydrolase (Fch). MtdA and Fch reactions are reversible and were described in *M. extorquens* AM1 to work preferably in the reductive direction so as to ensure C1-assimilation via the serine cycle (Crowther et al., 2008). MtdA can be involved also in the H₄MPT pathway as it can oxidize both CH₂H₄F and CH₂H₄MPT but it uses specifically NADP as a cofactor (Vorholt et al., 1998). The organisms using MtdA/Fch are always associating with the formyl-H₄F ligase (FtfL) whereas those using FOLD can be associated with either PurU or FtfL or both (Chistoserdova, 2011).

Except for PurU, all the reactions of this pathway are reversible (Nagy et al., 1995). A hypothesis suggests that the CH₂H₄F intermediate could be generated via the direct condensation of formaldehyde with H₄F, via H₄MPT and via the H₄F pathway; in this sense H₄MPT would then play a key role in both the dissimilation and assimilation processes.

3. GSH-dependent pathway

GSH-dependent pathway is found in methylotrophic bacteria and yeast. It is the simplest pathway to convert formaldehyde to formate. It is also the most common reaction to oxidize formaldehyde as it is found in many non-methylotrophic organisms from all kingdoms (i.e. plants, mammals, bacteria). Formaldehyde reacts first spontaneously with GSH. Then, two enzymes are involved: a NAD⁺ and GSH-dependent formaldehyde dehydrogenase (FaDH) and formyl-GSH hydrolase (FGH) generating respectively formyl-GSH and formate. Instead of glutathione, some Gram-positive bacteria employ a “sugar thiol”, mycothiol and a MySH-dependent formaldehyde dehydrogenase (Vorholt, 2002). Recently, bacillithiol was found as a novel cofactor involved in formaldehyde oxidation in *B. methanolicus* MGA3 (Müller et al., 2015b).

4. Formate oxidation to CO₂

No matter the pathway used, for some methylotrophs the produced formate can be assimilated or is most commonly oxidized to CO₂ thanks to a formate dehydrogenase (Fdh). As in many organisms, methylotrophs have at least one Fdh. For example, in *M. extorquens* sp four different Fdhs are found (Chistoserdova et al., 2004, 2007; Laukel et al., 2003).

C. C1-intermediate assimilation

Four assimilatory pathways are known in methylotrophy: the CBB cycle that assimilates carbon exclusively using CO₂, the serine cycle that assimilates carbon at the CH₂H₄F and CO₂ levels, the RuMP and XuMP pathways that enable carbon assimilation at the formaldehyde level. Interestingly, all these pathways are cyclic processes and are sharing one common feature, the condensation of the C1-intermediate with a C1-acceptor molecule. Obviously, an efficient regeneration of this C1-acceptor is crucial to ensure the long-term operation of these processes.

1. Calvin-Benson-Bassham (CBB) cycle

The CBB cycle, also known as the Ribulose BiPhosphate (RuBP) pathway or Calvin cycle, occurs in few known methylotrophs - mostly methanotrophs - and some autotrophic bacteria (e.g. *Paracoccus denitrifians*) (Baker et al., 1998; Ettwig et al., 2010; Khadem et al., 2011). In some cases, the CBB cycle is the only C1-assimilation pathway but sometimes another C1-assimilation is found as well. In the case of methanotrophs, methane is first oxidized to CO₂ to be assimilated via the CBB cycle. The CBB cycle can be divided in three steps (Figure 3). During carbon fixation, CO₂ is condensed with RuBP by the ribulose biophosphate carboxylase (RuBisCo) to produce two molecules of 3-phosphoglycerate (3PG). 3PG is either used to complete biomass production or subsequently phosphorylated into 1,3-BiPG and reduced to form glyceraldehyde 3-phosphate (Ga3P). Ga3P is then converted to ribulose 5-phosphate (Ru5P) that will be ultimately be phosphorylated to regenerate the RuBP pool.

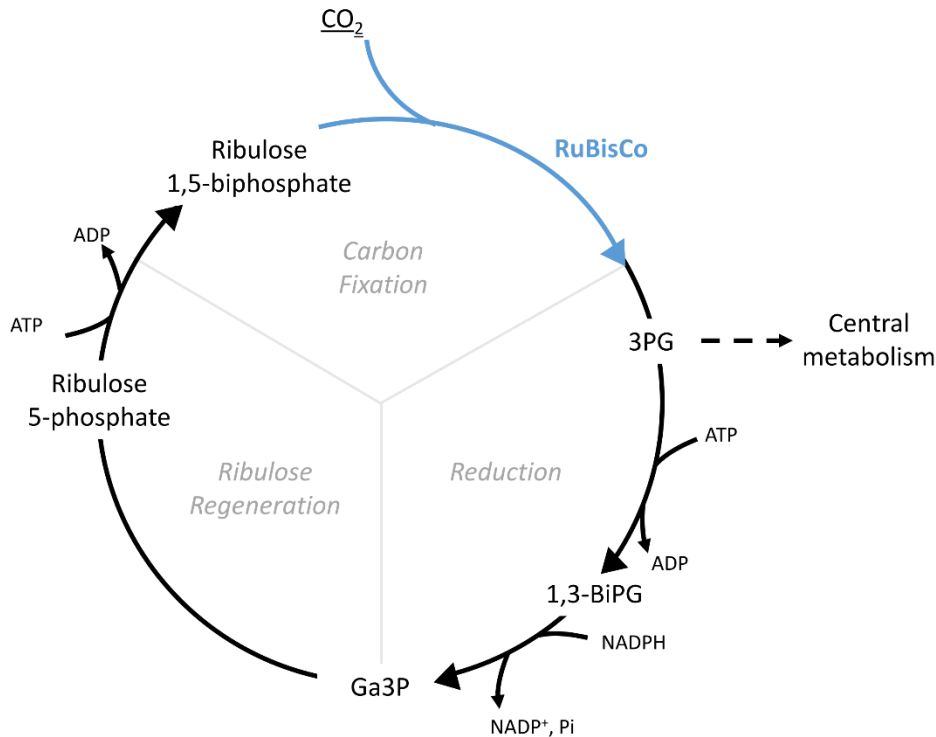


Figure 3: CO₂ assimilation in the Calvin-Benson-Bassahm cycle. Enzyme is in bold. C1-intermediate is underlined.

RuBisCo, ribulose biphosphate carboxylase. *Ga3P*, glyceraldehyde-3-phosphate; *3PG*, 3-phosphoglycerate; *1,3-BiPG*, 1,3-biphosphoglycerate.

2. Serine Cycle

The serine cycle occurs in *Alphaproteobacteria* (e.g. *M. extorquens*). In this cycle, the C1-intermediate is assimilated via the condensation of CH₂H₄F with glycine to form serine. After several reaction steps resulting in phosphoenol-pyruvate (PEP) production, PEP is condensed with CO₂ to form oxaloacetate (OAA). Then, OAA is transformed into malate which is further converted to malyl-CoA. Malyl-CoA is subsequently cleaved into two molecules: glyoxylate and acetyl-CoA (Ac-CoA) (Figure 4). Glyoxylate is transformed back into glycine to complete the serine cycle while Ac-CoA is used either as a precursor of biosynthetic pathways or to regenerate glyoxylate. In many methylotrophs, glyoxylate regeneration is done via the ethylmalonyl-CoA (EMC) pathway that has been described in *M. extroquens* (Erb et al., 2007; Peyraud et al., 2009). In the cyclic EMC pathway, two molecules of CO₂ are needed to ensure glyoxylate regeneration. However, some methylotrophs lack the EMC pathway and regenerate glyoxylate from Ac-CoA by using the traditional glyoxylate shunt (Chen et al.,

2010). The EMC pathway is not specific of C1-assimilation as it has been described to be functional during *M. extorquens* growth on acetate (Schneider et al., 2012).

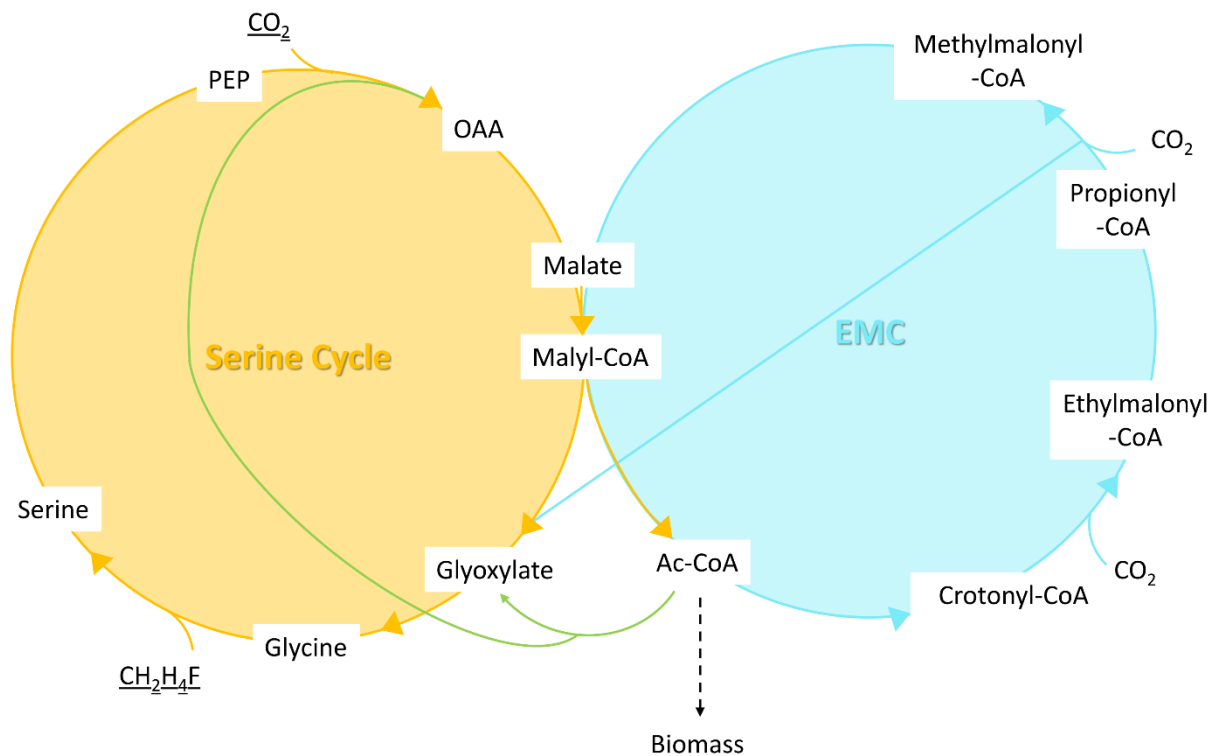


Figure 4: CH₂H₄F assimilation via the serine cycle.

C1-intermediates are underlined. Glyoxylate is regenerated either via the ethylmalonylCoA (EMC) pathway (in blue) or via the glyoxylate shunt (in green), the serine cycle is in orange.

Ac-CoA, acetyl-CoA; CH₂H₄F, methylene-H₄F; OAA, oxaloacetate; PEP, phosphoenol pyruvate.

3. Ribulose Monophosphate (RuMP) Pathway

The RuMP pathway occurs in *Betaproteobacteria* and *Gammaproteobacteria* as well as in Gram-positive bacteria (e.g. *B. methanolicus*). The RuMP pathway can be divided in three parts (Figure 5). In the first part - C1 fixation-, formaldehyde is condensed with Ru5P to form hexulose-6-phosphate (H6P) by a 3-hexulose-6-phosphate synthase (Hps). H6P is then isomerized to fructose-6-phosphate (F6P) by a 6-phospho-3-hexuloisomerase (Phi). Then, come the cleavage and regeneration steps. There are two known variants for the cleavage phase and for the regeneration phase. F6P can be further cleaved to pyruvate and glyceralde-3-phosphate (Ga3P) via the 2-keto-3-deoxy-6-phosphogluconate aldolase (Kdpga), an enzyme

from the Entner-Doudoroff pathway. Alternatively, F6P is converted to Ga3P and dihydroxyacetone phosphate (DHAP) by fructose-1,6-biphosphate aldolase (Fba), an enzyme of glycolysis. Finally, Ru5P regeneration is achieved via two variants sharing a combination of transketolase and some enzymes from the pentose phosphate pathway (PPP). Ru5P can be regenerated via the sedoheptulose-1,7-biphosphatase (SBPase variant) or via the transaldolase (TA variant). Either way, ribulose-phosphate 3-epimerase (Rpe) or ribose-5-phosphate isomerase (Rpi) ultimately produce Ru5P respectively from xylulose-5-phosphate and from ribose-5-phosphate.

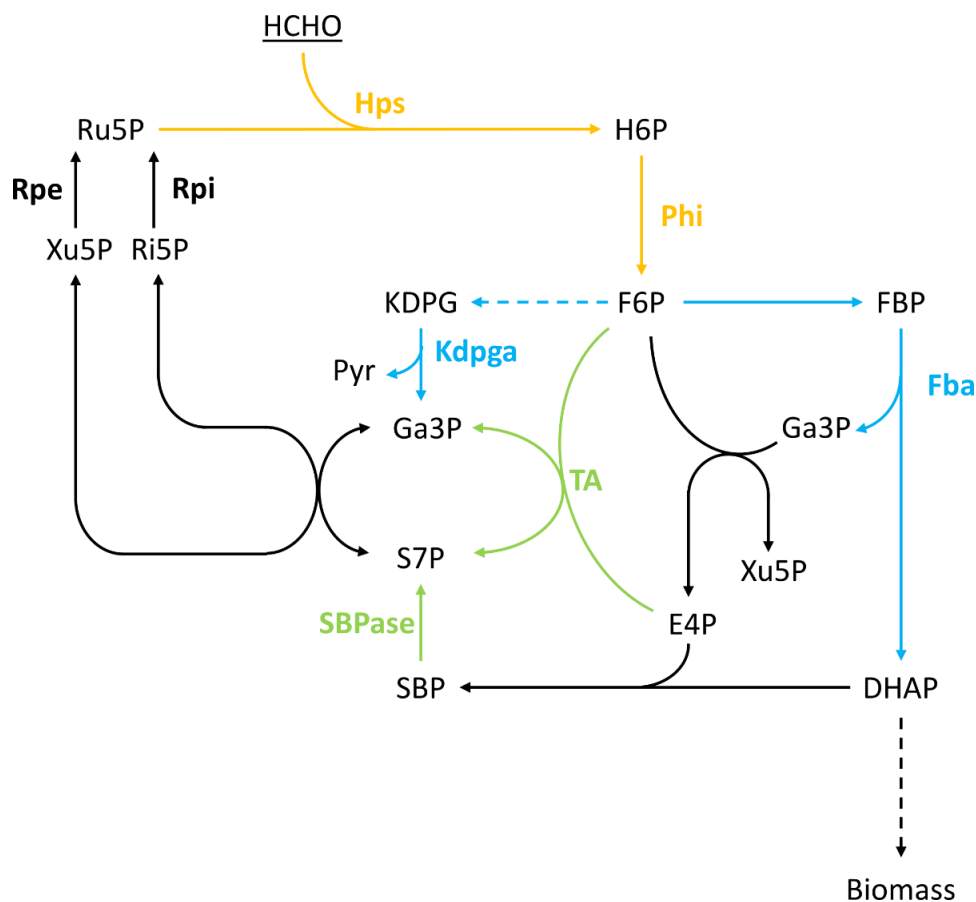


Figure 5: Formaldehyde assimilation in the ribulose monophosphate pathway.

The C1-assimilation pathway is in orange, the two variants for fructose-6-phosphate cleavage are in blue and the two variants for ribulose-5-phosphate regeneration are in green.

Enzymes are written in bold. C1-intermediate is underlined.

Fba, fructose-1,6-biphosphate aldolase; *Hps*, 3-hexulose-6-phosphate synthase; *Kdpga*, deoxy-6-phosphogluconate aldolase; *Phi*, 6-phospho-3-hexuloisomerase; *Rpe*, ribulose-phosphate 3-epimerase; *Rpi*, ribose-5-phosphate isomerase; *SBPase*, sedoheptulose-1,7-biphosphatase; *TA*, transaldolase. *DHAP*, dihydroxyacetone phosphate; *E4P*, erythrose-4-phosphate; *F6P*, fructose-6-phosphate; *FBP*, fructose-1,6-biphosphate; *Ga3P*, glyceraldehyde-3-phosphate; *H6P*, hexulose-6-phosphate; *KDPG*, 2-keto-3-deoxy-6-phosphogluconate; *Pyr*, pyruvate; *Ri5P*, ribose-5-phosphate; *Ru5P*, ribulose-5-phosphate; *S7P*, sedoheptulose-7-phosphate; *SBP*, sedoheptulose-1,7-biphosphate; *Xu5P*, xylulose-5-phosphate.

4. Xylulose Monophosphate (XuMP) Pathway

The XuMP pathway, also named dihydroxyacetone (DHA) pathway, is the only methylotrophic pathway specific of yeasts (e.g. *Pichia angusta* or *Candida boidinii*) (Figure 6). The XuMP pathway is similar to the CBB cycle or the RuMP pathway as the C1-unit (here formaldehyde) is condensed with a phosphorylated pentose. More precisely, formaldehyde is condensed with xylulose-5-phosphate (Xu5P) by a dihydroxyacetone synthase (Das) to form two C3-molecules: DHA and Ga3P. A catalase (Cat) detoxifies the H₂O₂ produced during AO activity. AO, Das and Cat are specifically located in the peroxisome. Indeed, the C-terminus sequences of these three proteins contain a peroxisomal targeting signal (PTS) PTS1. PTS1 sequence enables proteins to cross the peroxisomal membrane (van der Klei et al., 2006). A second PTS exist in yeast to address proteins to the peroxisome, PTS2, which is found on the N-terminus of proteins but is rarer than PTS1. In a *pex* mutant (lacking of peroxisome), despite AO, Das and Cat presence in the cytosol, *Pichia angusta* was not able to grow on methanol meaning that this spatial organisation is crucial for the survival of methylotrophic yeasts (Tan et al., 1995). Once produced, DHA is phosphorylated to DHAP by the dihydroxyacetone kinase (Dak). During methylotrophic growth, this enzyme is essential as its gene deletion prevented growth on methanol in *Pichia pastoris*, highlighting the key role of DHA in methanol assimilation (Lüers et al., 1998). Moreover, it was shown in *P. pastoris* upon growth on methanol, that isoforms of PPP enzymes were upregulated (Russmayer et al., 2015). These isoforms contained a PTS1 sequence. It is most likely that Xu5P recycling is completed in the peroxisome (Russmayer et al., 2015).

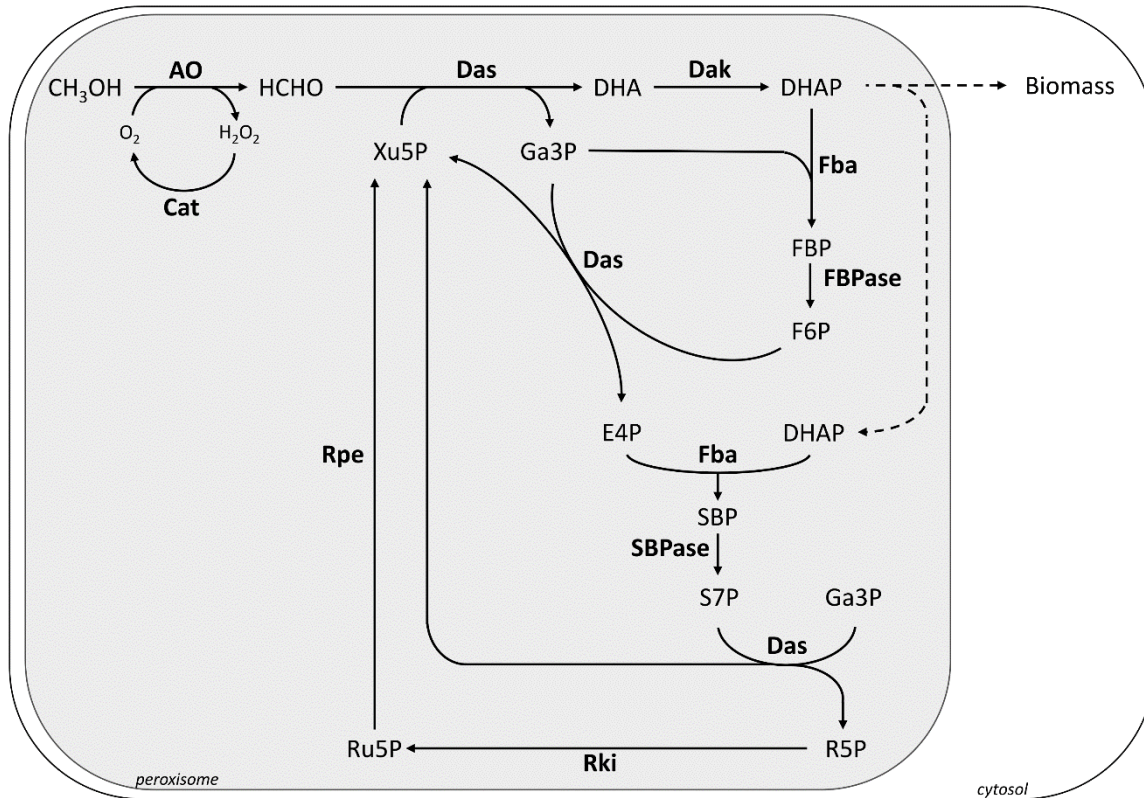


Figure 6: Methanol assimilation in the xylulose monophosphate pathway in yeast. Adapted from (Russmayer et al., 2015). Enzymes are written in bold.

AO, alcohol oxidase; Cat, catalase; Dak, dihydroxyacetone kinase; Das, dihydroxyacetone synthase; Fba, fructose-1,6-biphosphate aldolase; FBPase, fructose-1,6-biphosphatase; Rki, ribose-5-phosphate ketol-isomerase; Rpe, ribulose-5-phosphate 3-epimerase; SBPase, sedoheptulose-1,7-biphosphatase. DHA, dihydroxyacetone; DHAP, dihydroxyacetone phosphate; E4P, erythrose-4-phosphate; F6P, fructose-6-phosphate; FBP, fructose-1,6-biphosphate; Ga3P, glyceraldehyde-3-phosphate; R5P, ribose-5-phosphate; Ru5P, ribulose-5-phosphate; S7P, sedoheptulose-7-phosphate; SBP, sedoheptulose-1,7-biphosphate; Xu5P, xylulose-5-phosphate.

II. Natural methylotrophy in the biotechnology industry

The use of methylotrophs in the biotechnology industry is not that recent as it has been more than forty years that they are used for large-scale single cell protein production (Windass et al., 1980). Since then, the biotechnological use of natural methylotrophs has expanded. By using various methylotrophic organisms and their various pathways, the production of a wide-range of value-added compounds could be achieved: organic acids, polyhydroxyalkanoates (PHAs), amino acids, chemicals and proteins (Bourque et al.; Brautaset et al., 2010; Ettwig et al., 2010; Irla et al., 2017; Liu et al., 2018; Sonntag et al., 2014). Here, we decided to focus on the three most considered organisms i.e. *B. methanolicus*, *M. extorquens* and *P. Pastoris*.

A. *Bacillus methanolicus*

B. methanolicus is a Gram-positive facultative methylotrophic bacterium able to naturally use methanol as well as mannitol, glucose and arabitol as carbon sources (Delépine et al., 2020). One particularity of this bacterium is that its optimal growth temperature is around 50-55°C. Its growth rate on methanol is of 0.65 h⁻¹ (Bozdag et al., 2015). Methanol is assimilated via a NAD-dependent Mdh associated with the RuMP cycle. This microorganism is the first example of plasmid-dependent methylotrophy (Brautaset et al., 2004). This plasmid is a natural plasmid of 19 kbp encoding for *mdh* and five genes involved in the RuMP pathway while *hps* and *phi* are found in the same operon on the chromosome. The RuMP pathway relies on both the SBPase and the TA variants (Müller et al., 2015c; Pfeifenschneider et al., 2020). In the last two decades, with the objective to develop methanol-based cell factories, significant progress has been made in the development of synthetic and system biology tools as well as in the understanding of the metabolic background enabling methanol utilization in *B. methanolicus*. In particular, *B. methanolicus* has been intensively investigated, including genome sequencing, transcriptome, proteome and metabolome analysis, biochemical characterization and metabolic modelling (Delépine et al., 2020). Very recently, CRISPRi was developed in *B. methanolicus* MGA3 (Schultenkämper et al., 2019). The different compounds that *B. methanolicus* can produce are listed in Table 1.

Product	Concentration (g/L)	Reference
L-Glutamate	59	(Heggeset et al., 2012)
L-Lysine	65	(Brautaset et al., 2010)
GABA	9	(Irla et al., 2017)
Cadaverine	11.3	(Naerdal et al., 2015)
Acetoin	0.42	(Erythropel et al., 2018)

Table 1 : Chemicals produced using *B. methanolicus* chassis.

Among these, only L-glutamate and L-lysine can naturally be produced. So far, no L-lysine production could not be achieved without co-producing L-glutamate. These amino acids are both derived from intermediates of TCA cycle: L-glutamate from 2-oxoglutarate and L-lysine

from oxaloacetate. Both L-glutamate and L-lysine are industrially important compounds as L-glutamate is a flavor enhancer and L-lysine is part of the essential amino acids. Since few years, the progress in the genetic engineering of *B. methanolicus* has enabled to expand the production to compounds that are not synthesized by the native metabolism: cadaverine (Naerdal et al., 2015), γ -aminobutyric acid (GABA) (Irla et al., 2017) and acetoin more recently (Erythropel et al., 2018). Cadaverine is a product of lysine decarboxylation and is building block used for the production of polyamids, GABA is a non-protein amino acid used in the pharmacology and acetoin is used in the food industry as a flavor enhancer.

B. *Methylobacterium extorquens*

M. extorquens (recently renamed *Methylorubrum extorquens*) is a Gram-negative facultative methylotrophic bacterium able to use methanol as well as succinate, oxalate and acetate as carbon sources and can co-consume some of them (Peyraud et al., 2012). Methanol is assimilated via a PQQ-dependent Mdh associated with the serine cycle. *M. extorquens* AM1 strain is a model organism for enzyme and pathway discovery that enabled the elucidation of the H₄MPT dissimilation pathway and of the EMC pathway (see section I.B.1 & I.C.2). *M. extorquens* AM1 growth rate on methanol is of $\mu=0.168 \text{ h}^{-1}$ (Peyraud et al., 2011). A suite of omics (i.e. genome scale model, ¹³C-metabolomics) & genetic engineering tools have been developed and are now used to turn *M. extorquens* AM1 into a platform host to produce value-added products from methanol (Kiefer et al., 2011; Peyraud et al., 2011; Reaser et al., 2016).

Product	Concentration (g/L)	Reference
PHB	40-52.9	(Bourque et al.,1995)
Mesaconic	0.07	(Sonntag et al., 2014)
Methylsuccinic acid	0.06	(Sonntag et al., 2014)
α -Humulene	1.65	(Sonntag et al., 2015)
Mevalonate	2.67	(Liang et al., 2017)
GFP	4	(Bélanger et al., 2004)

Table 2: Chemicals produced using *M. extorquens* chassis.

The different compounds that *M. extorquens* can produce from methanol are listed in Table 2. Historically, *M. extorquens* was first used to produce PHAs, which are of great interest to replace oil-based plastics while being sustainable, biocompatible and biodegradable (Shah et al., 2008). *M. extorquens* is a natural producer of poly-3-hydroxybutyrate (PHB). PHB is synthesized from 3-hydroxybutyryl-CoA, a metabolite from EMC pathway. PHB production optimization has focused on media formulation rather than on the engineering of the production host (Mokhtari-Hosseini et al., 2009; Orita et al., 2014). In contrast, the production of novel compounds from CoA-intermediates of the EMC pathway was achieved by the introduction of heterologous genes: mesaconic and methylsuccinic acid after introduction of thioesterases (Sonntag et al., 2014), α -humulene after introduction of the mevalonate pathway, α -humulene synthase and farnesyl pyrophosphate synthase (Sonntag et al., 2015), and mevalonate after the introduction of the mevalonate pathway and by the engineering of a mevalonate biosensor (Liang et al., 2017). Mesaconic, methylsuccinic acid and mevalonate are key building blocks for the chemical industry and α -humulene gets attention in the pharmaceutical field for its anti-inflammatory and anti-cancer properties. As a first proof of concept for recombinant protein production, green fluorescent protein was produced in *M. extorquens* ATCC 551366 (Bélanger et al., 2004), since then few other proteins were produced but the final concentration reached was very low (Choi et al., 2008; Gutiérrez et al., 2005). One particularity of this bacterium is its pink colour coming from its natural capacity to produce carotenoids. Genetic engineering of *M. extorquens* by using CRISPRi to overproduce carotenoids was achieved but still further developments are needed to biosynthesize new products of interest (Mo et al., 2020). Few years ago, butanol production from ethylamine has been demonstrated (Hu and Lidstrom, 2014).

C. Pichia pastoris

The methylotrophic yeast *P. Pastoris* (syn. *Komagataella* spp.) has been the first yeast to be used at industrial scale, more than 40 years ago, for the commercial production of single cell proteins (SCP) using methanol (Wegner, 1983). *P. pastoris* is a chassis of great interest for the protein production thanks to its high secretory capacity and its strong methanol inducible promoters, like AOX1 promoter, that are used to express heterologous proteins (Puxbaum et

al., 2015; Weinhandl et al., 2014). During protein production, to improve the amounts of protein secreted, mixing methanol with another carbon source as glycerol (Jordà et al., 2014), sorbitol (Jungo et al., 2007) or glucose (Jordà et al., 2012) has been a strategy commonly used. So far, most efforts have been focused on mixed cultures because methanol bioconversion still remains difficult. Indeed, methylotrophic yeasts such as *P. pastoris* typically shows a relatively slow specific growth rate on pure methanol ($\mu = 0.10 \text{ h}^{-1}$ to 0.15 h^{-1} (Looser et al., 2015; Tomàs-Gamisans et al., 2016)). Moreover, during the first step of methanol assimilation (methanol oxidation to formaldehyde), the energy is released in the form of heat and is not conserved as reducing power (NADH). In addition, ^{13}C -MFA studies revealed that at least 50% of methanol is directly dissimilated to CO_2 , hence the C1 assimilation ratio is low (Jordà et al., 2012, 2014). With the development of synthetic and systems biology tools (e.g. CRISPR/Cas9, genome-scale metabolic models, omics analysis), the knowledge of methanol metabolism and its regulation improved (Baumann et al., 2010; Russmayer et al., 2015; Tomàs-Gamisans et al., 2016; Peña et al., 2018; Fischer and Glieder, 2019). The use of *P. pastoris* is now expanding towards the production of value-added products from methanol as illustrated by the recent examples listed in Table 3. A co-culture of *P. pastoris*-*P. pastoris* system was used, to reduce metabolic burden, one strain was expressing the hydromonacolin-L (DML) biosynthesis pathway and the other one was expressing genes encoding for proteins enabling the conversion of DML into monacolin J and lovastatin. However, this process was limited by DML diffusion through the cell membrane. Monacolin J and lovastatin are molecules used in the treatment of cardiovascular diseases.

Product	Concentration (g/L)	Reference
Monacolin J	0.5939	(Liu et al., 2018)
Lovastatin	0.2508	(Liu et al., 2018)

Table 3: Chemicals produced using *P. pastoris* chassis.

D. Limitations

Although native methylotrophs are attractive platforms for the methanol industry, several limitations are still complicating their full exploitation at the industrial scale. The first limitation is the lack of genetic tools to engineer these organisms, in particular in methylotrophic bacteria where genetic deletions or integrations are not yet possible. The second limitation is that our current understanding of the cellular metabolism, physiology and regulation is still incomplete despite the extensive systems-level studies. Synthetic methylotrophy offers an alternative strategy by integrating methylotrophy into non-native methylotrophic production hosts. This allows using methanol as feedstock for well established, industrially-relevant organisms for which a large genetic toolbox is available. In addition, the reconstruction of methylotrophic pathways in such organisms will help to better understand the biological complexity of methylotrophy.

III. Engineering synthetic methylotrophic assimilation pathways

Synthetic methylotrophy is the implementation of natural or artificial methanol assimilation pathways in platform microorganisms to produce value-added products using methanol as a carbon source. *E. coli*, *Corynebacterium glutamicum* and *Saccharomyces cerevisiae* are workhorses in biotechnology field, able to produce a wide range of products including amino acids and biofuels to synthons used in the pharmaceutical and chemistry industry (Becker and Wittmann, 2015). Many tools are available to modify them and to evaluate the potential of the engineered strains for methanol-based bioproduction. They are thus ideal candidates for establishing synthetic methylotrophy.

At first, from a technical point of view, the implementation of methylotrophy in non-native methylotrophs can be seen as quite straightforward, as “only” the genes encoding for methanol catabolism need to be added to the production host (Figure 7). However, engineering methylotrophy is a true challenge since all energy and biosynthetic requirements have to be fulfilled from a unique reduced C1 precursor. In addition, cells have to cope with a toxic compound (i.e. formaldehyde) which is an obligate intermediate in methanol metabolism. Hence, any imbalance between formaldehyde oxidation and assimilation would result in its accumulation with fatal consequences on the cell integrity and survival. Because

formaldehyde oxidation to CO_2 is an efficient process, C1-assimilation is the main bottleneck for synthetic methylotrophy. As presented in section I.C, natural C1-assimilation pathways are cyclic processes requiring a C1-acceptor (i.e. Ru5P in the RuMP pathway and Xu5P in XuMP pathway) to enable the C-C bond formation. Finding the fine tune between the production and recycling of the C1-acceptor is thus determinant to efficiently assimilate methanol and enable methylotrophic growth.

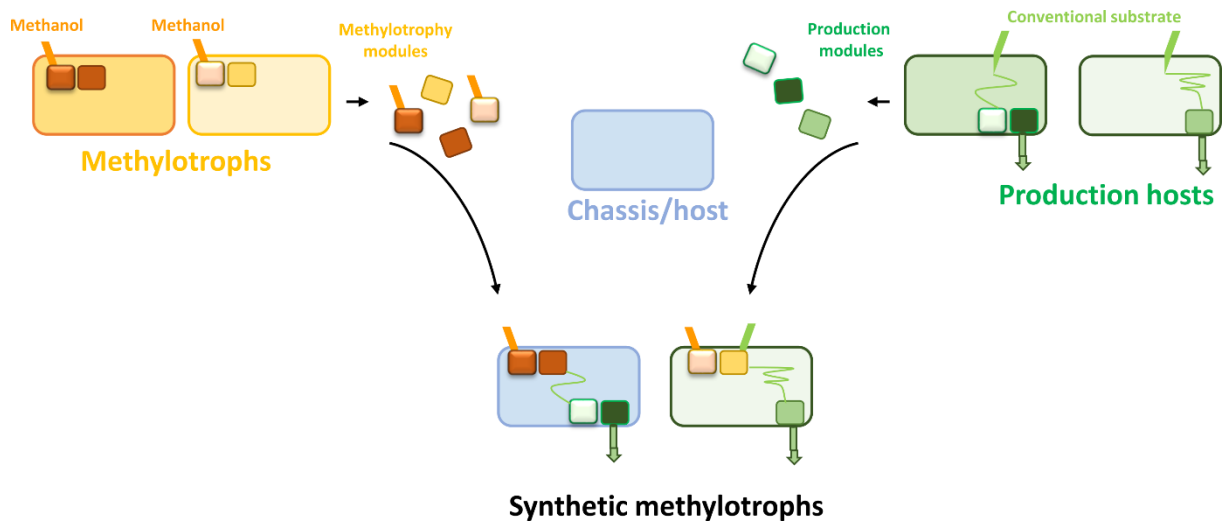


Figure 7: Implementation of methylotrophic modules in chassis or production hosts to engineer synthetic methylotrophs.

A. NAD-dependent Mdh and RuMP pathway and its optimization

During the past years, most attempts to introduce methylotrophy into non-methylotrophic hosts have relied on the RuMP pathway. Indeed, this pathway is the simplest to implement in hosts: only three enzymes (Mdh, Hps and Phi) are needed to connect the methylotrophic modules to the host central metabolism. Among the different Mdhs found in nature, the choice of a NAD-dependent Mdh seems the best as NAD synthesis does not require oxygen unlike PQQ synthesis. Moreover, hosts like *E. coli* cannot natively synthesize PQQ while NAD is a ubiquitous cofactor found in cells during both aerobic and anaerobic growth. RuMP pathway has the great advantage to only need three enzymes to be connected with the native central metabolism which ease its introduction in hosts while not requiring any cofactor

(Yurimoto et al., 2009). Moreover, the native metabolic network of the host can support both assimilation and regeneration of the C1-acceptor.

Mdh, Hps and Phi have already been introduced in *E. coli* (Müller et al., 2015a), in *C. glutamicum* (Witthoff et al., 2015) but also in *Pseudomonas putida* (Koopman et al., 2009) and *S. cerevisiae* (Dai et al., 2017). Last year, Hps and Phi were implemented in *Yarrowia lipolytica*, methanol was oxidised in formaldehyde by a native NAD-dependent alcohol dehydrogenase. However, like for Dai et al, methanol was not shown to be used as a carbon source (Vartiainen et al., 2019). By using ^{13}C -methanol, Müller and al were the first to provide direct evidence of methanol assimilation. The same year, when medium was supplemented with ^{13}C -methanol and another carbon source, the production of value-added product deriving from methanol was demonstrated for the first time as ^{13}C -cadaverine could be produced in *C. glutamicum* (Leßmeier et al., 2015). Later, production of other value-added compounds deriving from methanol was also achieved in *E. coli*, including naringenin (Whitaker et al., 2017) and more recently acetone (Bennett et al., 2020).

Very recently, by combining rational design and laboratory evolution, a synthetic methylotroph growing solely on methanol has been engineered in *E. coli* (Chen et al., 2020). The final strain reached a growth rate of $\mu=0.08\text{ h}^{-1}$ on methanol. Among all the attempts, this achievement is the first one of the kind in the community. This shows that implementation of synthetic methylotrophic pathway is far from being easy but opens the way to successful achievements in this field.

1. Rational engineering

a) Engineering of the NAD-dependent Mdh

During the first attempts, Mdh, Hps and Phi expression relied on the genes from *B. methanolicus* (Leßmeier et al., 2015; Meyer et al., 2018; Müller et al., 2015a; Witthoff et al., 2015). Even though, the different variants of Mdh were tested, their efficiency was pretty low, therefore scientists aimed to find new NAD-dependent Mdh with better kinetic parameters by studying Mdhs coming from different organisms. An alcohol dehydrogenase from the non-methylotrophic *B. stearothermophilus* was found to be able to use methanol as substrate and to exhibited a low K_m towards methanol (Sheehan et al., 1988). Whitaker and colleagues used this Mdh combined with Hps and Phi from *B. methanolicus* and showed that this combination

resulted in a higher biomass yield and a higher flux from methanol to intracellular metabolites compared to the association using the Mdh from *B. methanolicus* (Whitaker et al., 2017). Another alcohol dehydrogenase from *Cuprividus necator* was found to exhibit similar activity towards methanol at 30°C than the Mdh from *B. methanolicus* at 45°C (its optimal temperature activity). By using directed molecular evolution, a variant of the Mdh from *C. necator* with a catalytic efficiency for methanol 6-fold higher than the WT was engineered (Wu et al., 2016). Mdh engineering by using phage assisted evolution was also done on the Mdh2 of *B. methanolicus* MGA3 and resulted in two times higher methanol incorporation than the strain expressing the native Mdh2 (Roth et al., 2019).

b) Engineering of the cell redox state

The redox state of the cell is directly affecting the activity of the Mdh, the higher the NADH/NAD⁺ ratio is, the less the methanol oxidation reaction will be favourable. Results obtained after directed evolution experiments (see section III.A.4.a for more details) aiming at improving the growth on methanol showed mutations in genes involved in *nadR* gene encoding for a NAD repressor in the evolved strain (Meyer et al., 2018). The activity of this repressor was found to be reduced, highlighting the importance of the NADH/NAD⁺ ratio during growth on methanol. To balance the redox state of the cell, authors also showed that when NAD-dependent malate dehydrogenase *maldh* was knockout, growth was improved on methanol and gluconate. This was also observed in *E. coli* during growth on methanol and yeast extract by Rohlhill and al even if methanol assimilation was not improved (Rohlhill et al., 2020). In the same line, methanol oxidation rate was improved when Mdh was coupled with a 'NADH sink' by using the lactate dehydrogenase to recycle NADH into NAD⁺ (Price et al., 2016).

c) Engineering of the regulation of gene expression

In natural methylotrophs, methanol and formaldehyde are used to regulate dynamically the expression of genes (Jakobsen et al., 2006). The use of promoters induced by methanol have been widely used in native methylotrophs to produce recombinant proteins (see section II.C). In contrast, synthetic methylotrophs do not recognize methanol as a substrate, so no gene regulation linked to methanol presence is set. Therefore, attempts to make synthetic methylotrophs sensing methanol have been done. In *E. coli*, a MxaYZ, a chimeric two-

components system, was created by fusing the periplasmic methanol sensing domain MxaY of the Mdh from *Paracoccus denitrificans* with the cytoplasmic catalytic transmitter domain of EnvZ from *E. coli* (Ganesh et al., 2017). Results showed that MxaYZ sensed extracellular methanol and activated the expression of GFP. Formaldehyde is a common metabolite known by cells, native formaldehyde-inducible promoter (P_{frm}) exist that activates the dissimilatory pathway to avoid formaldehyde accumulation in cells. Several examples of engineered P_{frm} in synthetic methylotrophs have been described (Denby et al., 2016; Rohlhill et al., 2017). In *E. coli*, P_{frm} is upstream to *frmRAB* operon and is repressed by FrmR, a transcriptional repressor. When formaldehyde is present in cells, it interacts with FrmR modifying its conformation, leading to the disruption of FrmR with the DNA-binding site, activating the transcription of the operon (Denby et al., 2016). When placed upstream the *mdh*, *hps* and *phi* operon, the native and engineered P_{frm} enabled both to improved biomass production in the engineered strain (Rohlhill et al., 2017). In another study, FmrR and P_{frm} , were used to build a formaldehyde biosensor in *E. coli* and used to modulate the expression of *mdh*, *hps* and *phi* (Woolston et al., 2017).

d) Engineering of the dissimilatory and recycling pathways

In order to drive the metabolism towards methanol assimilation, a common strategy is to delete the gene encoding formaldehyde dehydrogenase (i.e. *frmA*), the first enzyme involved in the formaldehyde detoxification pathway, to avoid a carbon loss as CO₂ (Whitaker et al., 2017; Bennett et al., 2018; Meyer et al., 2018; Rohlhill et al., 2020).

Another strategy set to stimulate methanol assimilation was to favour the recycling of the C1-acceptor (i.e. Ru5P). In synthetic methylotrophs, Ru5P regeneration is ensured by the non-oxidative part of the PPP. In the natural methylotroph *B. methanolicus*, it was shown that *pfk*, *rpe*, *tkt*, *glpX* and *fba* (i.e. genes involved in the SBPase variant of the RuMP pathway) are key genes involved for Ru5P regeneration. Their deletion resulted in a loss capacity to grow on methanol of the bacteria (Brautaset et al., 2004). Therefore, enhancing the host capacity to regenerate Ru5P by overexpressing heterologous enzymes appears critical to achieve efficient synthetic methylotrophy. To boost Ru5P regeneration, Bennett et al overexpressed PPP enzymes by integrating *pfk*, *rpe*, *tkt*, *glpX* and *fba* genes from *B. methanolicus* on *E. coli* chromosome (Bennett et al., 2018). Even if culture media had to be supplemented with yeast extract, the resulting strain achieved a 20% improvement in biomass production during

growth on methanol compared to the parental strain and 59% of mean enrichment was reached at the PEP-level. PEP is an interesting metabolite to follow methanol assimilation and Ru5P recycling. Indeed, the more carbons PEP are labelled, the more methanol had been assimilated and so the more Ru5P had been recycled. Woolston et al used iodoacetate to block the glyceraldehyde-3-phosphate dehydrogenase (Ga3PDH) and succeeded to activate the native SBPase variant to recycle Ru5P leading to a higher methanol incorporation into central metabolites in *E. coli*. The mean enrichment in F6P reached 27.5% when the iodoacetate was present. Moreover, the intracellular concentrations of F6P, S7P and Ru5P were higher (Woolston et al., 2018). Rohlhill et al. went one step further, by adding another layer of complexity by modulating the expression of *rpe* and *tkt* using P_{frm} , when combined with the disruption of the malate dehydrogenase improved methanol carbon incorporation into intracellular metabolites, but yeast extract was needed during the experiment (Rohlhill et al., 2020).

To build their methylotrophic strain, Chen et al. first integrated two operons in *E. coli*'s genome. The first one consisted of *mdh*, *hps* and *phi* genes while the second one consisted of the same *mdh* and *phi* genes than in the first operon but with a *hps* gene coming from a different organism and added *tkt* and *tal* genes in the operon to enhance Ru5P recycling (Chen et al., 2020). Moreover, after the first adaptive laboratory evolution (ALE) experiment (see section III.A.2.c), using Ensemble Modelling for Robustness Analysis (EMRA), authors identified that the high activity of phosphofructokinase and Ga3PDH was channelling the flux away from the RuMP cycle, which tended to unbalance the metabolic system. To reduce the activity of these two enzymes, *pfkA* was knockout and *gapA* gene was replaced by another *gapC* from *E. coli* BL21 encoding for a less efficient Ga3PDH than the native one.

2. Evolutionary engineering

Despite efforts, rational engineering strategies were not sufficient to ensure full synthetic C1-assimilation in microorganisms. The rational engineering had to be associated with evolutionary engineering to improve the microorganism performance. To do so, one strategy adopted by several groups was to first build a strain in which co-consumption of methanol with another carbon source (e.g. glucose, xylose, ribose or gluconate) was required to grow (Meyer et al., 2018; Chen et al., 2018; Bennett et al., 2020; Chen et al., 2020) and then to subject the strain to ALE.

a) Chassis optimization to engineer methanol-dependent strain for growth

To ensure methanol is co-utilized with the other carbon source, cells are engineered in order to gain a benefit in utilizing methanol. To do so, the common strategy was to design a strain dependent on methanol utilization for growth while the other carbon source is used only to ensure Ru5P production. This is achieved by blocking either the conversion of glucose to F6P or the Ru5P catabolic pathways depending of the carbon source used (Figure 8). That way, cells have no other choice than producing F6P from the co-assimilation of methanol and the other carbon source while ensuring a high pool of the key Ru5P. This strategy was applied by Meyer et al. to engineer an *E. coli* methanol dependent-strain ($\Delta rpiAB\Delta edd\Delta maldh$) co-utilizing gluconate (Meyer et al., 2018). Similarly, Chen et al., engineered an *E. coli* strain ($\Delta rpiAB$) co-utilizing xylose and another one (Δrpe) co-utilizing ribose (Chen et al., 2018). Later, to build their fully methylotrophic strain, Chen et al. decided to adopt the same strategy than in 2018 by deleting *rpiAB* but switched from BL21 genetic background to BW25113 (Chen et al., 2020). To enable the co-utilization of methanol and glucose, Bennet and al first deleted phosphoglucose isomerase (*pgi*) that converts glucose-6-phosphate (G6P) to F6P, so that G6P is only used by the oxidative part of the PPP to directly fill the pool of Ru5P. They also deleted *frmA* in order to push formaldehyde towards assimilation (Bennett et al., 2018). Later, they went one step further by knocking-out *edd* and *rpiAB* resulting in their $\Delta frmA\Delta pgi$ *E. coli* strain to completely abolish glucose assimilation but amino acids supplementation in the cultivation medium was then required (Bennett et al., 2020).

This strategy is not specific to *E. coli*, it was applied also to engineer a *C. glutamicum* methanol-dependent strains $\Delta aldH\Delta fadH\Delta rpi$ and $\Delta aldH\Delta fadH\Delta rpe$ respectively using xylose or gluconate and ribose or gluconate as co-substrates (Hennig et al., 2020). The genes *aldH* and *fadH* are respectively coding for acetaldehyde dehydrogenase and formaldehyde dehydrogenase, so their deletion enabled to prevent formaldehyde oxidation to CO₂.

By engineering methanol-dependent strains, methanol assimilation is coupled with cell growth, this opens the way to submit the engineered strains to ALE to improve methanol assimilation.

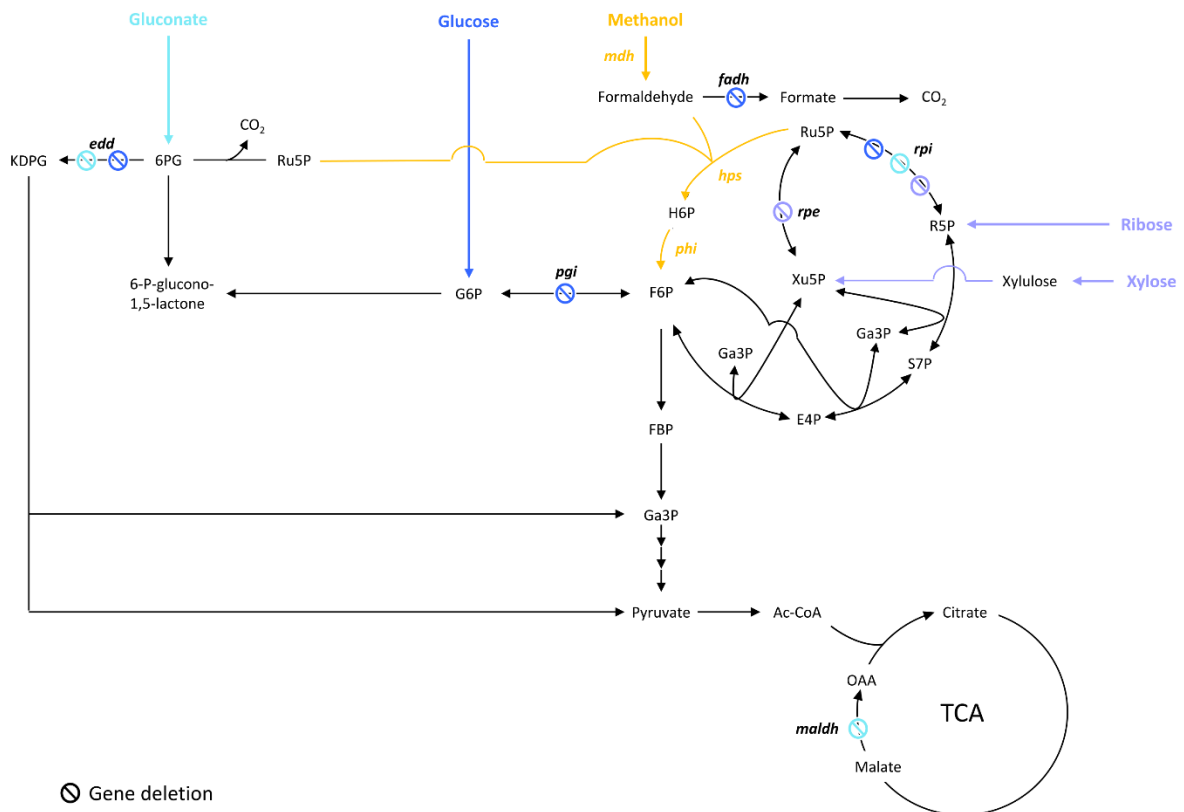


Figure 8: Targeted genes to construct methanol-dependent strains in *E. coli*. Adapted from (Wang et al., 2020). Engineered synthetic methanol-dependent strain co-assimilating gluconate (in light blue) by (Meyer et al., 2018); co-assimilating glucose (in blue) by (Bennett et al., 2020); co-assimilating ribose by (Chen et al., 2018) or xylose by (Chen et al., 2018, 2020) (in purple). Enzymes are written in bold.

edd, 6-phosphogluconate isomerase; *fadh*, formaldehyde dehydrogenase; *hps*, 3-hexulose-6-phosphate synthase; *maldh*, malate dehydrogenase; *mdh*, methanol dehydrogenase; *pgi*, glucose-6-phosphate isomerase; *phi*, 6-phospho-3-hexuloisomerase; *rpe*, ribulose-3-phosphate epimerase; *rpi*, ribose-5-phosphate isomerase. Ac-CoA, acetyl-CoA; E4P, erythrose-4-phosphate; F6P, fructose-6-phosphate; FBP, fructose-1,6-biphosphate; Ga3P, glyceraldehyde-3-phosphate; H6P, hexulose-6-phosphate; KDPG, 2-keto-3-deoxy-6-phosphogluconate; 6PG, 6-phosphogluconate; R5P, ribose-5-phosphate; Ru5P, ribulose-5-phosphate; S7P, sedoheptulose-7-phosphate; SBP, sedoheptulose-1,7-biphosphate; Xu5P, xylulose-5-phosphate.

b) Adaptive laboratory evolution (ALE)

Adaptive laboratory evolution (ALE) is used on microorganisms cultivated under defined conditions for weeks to years in order to select improved phenotypes on relevant carbon sources. Microorganisms can either be cultured in shake flasks with sequential serial passages or cells can be grown in chemostat with a controlled environment in which one component of the media is limiting. Shake flasks have the great advantage to enable to run many culture conditions in parallel but pH, oxygenation and cell density can vary during the experiment. By using the chemostat, oxygenation, pH, cell density (which can be higher than the one reached

in shake flasks) and the growth rate are controlled and kept constant. However, the operation costs of this device are higher than for shake flasks (Dragosits and Mattanovich, 2013).

Mutants evolving towards a faster growth can be selected, and methanol incorporation can be tested by measuring the incorporation of label in intracellular metabolites from ^{13}C -methanol (Table 4). Meyer et al decided to knockout *maldh* in $\Delta rpiAB\Delta edd$ background after *in silico* analysis (cf section III.A.1), and to apply ALE to this strain. A mutant co-utilizing methanol and gluconate at a growth rate of $\mu = 0.08 \text{ h}^{-1}$ and with a methanol uptake rate of $13 \text{ mmol.gCDW}^{-1}.\text{h}^{-1}$ - which is close to that reported for natural methylotrophs - was selected. 21% of PEP carbon atoms came from methanol (Meyer et al., 2018). When Chen et al used ALE on their $\Delta rpiAB$ strain, they succeeded to select a strain with a growth rate of $\mu = 0.17 \text{ h}^{-1}$. This strain was able to produce butanol from the co-consumption of methanol and xylose, and 22% of butanol carbons came from methanol (Chen et al., 2018). After ALE experiment, Bennet et al isolated an *E. coli* strain dependent on methanol + glucose without the need for amino acid supplementation. This strain was growing on glucose and methanol with a specific growth rate of $\mu = 0.15 \text{ h}^{-1}$. It was able to produce acetone, for which 22% of carbon atoms derived from methanol (Bennett et al., 2020).

Tuyshime et al performed ALE experiments on their *C. glutamicum* mutant and selected a strain growing on xylose and methanol at a growth rate of 0.03 h^{-1} with a methanol uptake rate of $0.86 \text{ mmol.gCDW}^{-1}.\text{h}^{-1}$ (Tuyshime et al., 2018). After ALE experiments, Henning et al selected, in *C. glutamicum*, an evolved $\Delta aldH\Delta fadH\Delta rpe$ strain co-utilizing methanol and gluconate. This strain was able to produce cadaverine, for which 43% of carbon atoms came from methanol. After a new round of ALE experiments started with the previous strain, another one was selected on ribose and methanol. This strain grew without yeast extract and exhibited a specific growth rate of $\mu = 0.10 \text{ h}^{-1}$ (Hennig et al., 2020).

Organism	Cultivations conditions	% of labelling incorporation	Reference
<i>E. coli</i> BW25113	500mM methanol + 5mM gluconate	21% at PEP-level	(Meyer et al., 2018)
<i>E. coli</i> BL21 DE3	250mM methanol + 50mM xylose	22% at butanol-level	(Chen et al., 2018)
<i>E. coli</i> BW25113	500mM methanol + 200mM glucose	22% at acetone-level	(Bennett et al., 2020)
<i>E. coli</i> BW25113	400mM methanol	100%	(Chen et al., 2020)
<i>C. glutamicum</i> ATCC 13032	125mM methanol + 27mM xylose	22 % at PEP-level	(Tuyishime et al., 2018)
<i>C. glutamicum</i> ATCC 13032	500mM methanol +20mM gluconate+ 0.5g/L yeast extract	30% at Cadaverine-level	(Hennig et al., 2020)

Table 4: Overview of ^{13}C -methanol incorporation level at the intracellular level or in final products in strains after ALE experiment.

c) Genes selection after ALE

As in nature, during ALE, genomic changes occur to select an improved phenotype enabling cells to cope with their environment. This is done by using chemical mutagens or transposon libraries or random mutagenesis. It is of great interest to get insight on which genes are essential (or non-essential) for methanol assimilation by investigating the mutations and their consequences. Mutations found in evolved strains can be classified in three categories, mutations affecting: (i) the dissimilatory pathway and the recycling pathways; (ii) the redox and energy balance; (iii) the substrate uptake and other enzymes connected to metabolism (Wang et al., 2020).

Mutations in *frmA* leading to the inactivation of the formaldehyde dissimilatory pathway were found in all the *E. coli* evolved strains (Chen et al., 2018; Meyer et al., 2018; Chen et al., 2020). This gene encodes for FaDH the first enzyme involved in formaldehyde dissimilatory pathway. It is most likely that to be efficient, methanol-dependent growth needs

formaldehyde to be redirected towards C1-assimilation. Mutations leading to *pykF* and *zwf* inactivation enabled to possibly increase F6P pool for Ru5P regeneration via the PPP (Chen et al., 2018). Similarly, *deoD* mutation was thought to increase Ru5P pool for formaldehyde assimilation (Chen et al., 2018).

Mutations were found in *nadR* (Meyer et al., 2018), *cyaA* (Chen et al., 2018) in *E. coli* and *mtrA* in *C. glutamicum* (Tuyishime et al., 2018), the enzymes encoded by these genes are involved in the redox state or in the energy supply of the cell. *nadR* and *mtrA* are genes encoding for enzymes involved in the balance of the redox state. Both mutations are thought to lead to the increase of NAD⁺ availability in cell, thereby most probably promoting methanol oxidation. *cyaA* encodes for an enzyme producing cyclic AMP (cAMP) from ATP. Transcription of TCA enzymes are activated by cAMP. Therefore, it is most likely that *cyaA* mutation would lower TCA cycle and consequently would lower NAD(P)H production at a beneficial level for methanol oxidation. Moreover, inactivation of *frmA* and *fodG*, for the synthetic methylotrophic strain, helps limiting the production of NADH (Chen et al., 2020).

Mutations in *gnd* encoding for 6-phosphogluconate dehydrogenase were found in the strain assimilating better methanol but still requiring xylose (Chen et al., 2020). GntR and AltR are transcriptional regulators, mutations in genes encoding these proteins were found respectively in the *E. coli* strain co-utilizing methanol and gluconate and *C. glutamicum* strain co-utilizing methanol and xylose (Meyer et al., 2018; Tuyishime et al., 2018). GntR represses two enzymes involved in gluconate uptake, therefore *gntR* mutations may have improved gluconate uptake. Similarly, AltR regulates among others the expression of *xylB* encoding for xylulose kinase and *adhA* encoding for alcohol dehydrogenase functioning as a methanol dehydrogenase in *C. glutamicum*. Therefore, *altR* mutation may have modified *xylB* and *adhA* expression leading to improved co-utilization of methanol and xylose by increasing substrate uptake.

Overall, ALE helped rerouting carbon flows towards methanol assimilation by limiting competing pathways and improving substrate uptakes in order to enhance efficient biomass production while improving the energy and redox state of the cell. The two ALE experiments were determinant for Chen et al. to select the first synthetic methylotrophic *E. coli* strain. The first one enabled to select a strain assimilating better methanol but still requiring xylose presence. After rational optimization and modification of the culture medium, the second ALE

experiment enabled to select a strain growing solely on methanol. From a general point of view, for future strain engineering, the genetic targets highlighted by ALE are very interesting as they unlocked key metabolic bottlenecks. However, one should keep in mind that these genes and their associated mutations are related to strains co-assimilating methanol with another carbon source in very specific culture conditions. To further engineer these evolved strains and switch them to condition on pure methanol remains a challenge as they were first modified by removing essential genes involved in Ru5P regeneration (Rpe or Rpi). It is most likely that these evolved strains will help in the evaluation, for instance, of methanol assimilating enzymes in the future.

3. Spatial engineering

Once produced formaldehyde has to be quickly condensed with Ru5P to avoid its toxic effect. To optimize formaldehyde assimilation, another strategy that has been proposed is to enhance the substrate channelling by locating enzymes in a close space. Orita et al were the first to address this question by fusing *hps* and *phi* from *Mycobacterium gastri* (Orita et al., 2007). They showed that *in vivo* when expressed in *E. coli* the protein fusion of Hps and Phi, *E. coli* growth rate was higher on formaldehyde than when the proteins were separated. Later, Price et al went one step further by engineering a supramolecular enzyme complex with Mdh from *B. methanolicus* and Hps and Phi from *M. gastri* (Price et al., 2016). They took advantage of the decameric structure of Mdh to design a supramolecular complex able to self-assemble by using SH3-ligand in order to “plug” the fusion protein of Hps and Phi previously described by Orita et al on it. This engineered complex enabled a faster conversion of methanol into F6P compared to unassembled proteins. In the same way, Fan et al used an alternative strategy to fuse Mdh from *B. stearothersophilus* with Hps and Phi from *M. gastri* by using flexible linkers (GGGS)_n (Fan et al., 2018). They also demonstrated an improvement in methanol conversion to F6P.

B. Other methylotrophic pathways tested

The implementation of synthetic methylotrophy in production hosts has been attempted extensively by using RuMP pathway. However, several groups have also evaluated the introduction of other pathways including native pathways and new synthetic ones.

1. Cyclic pathways

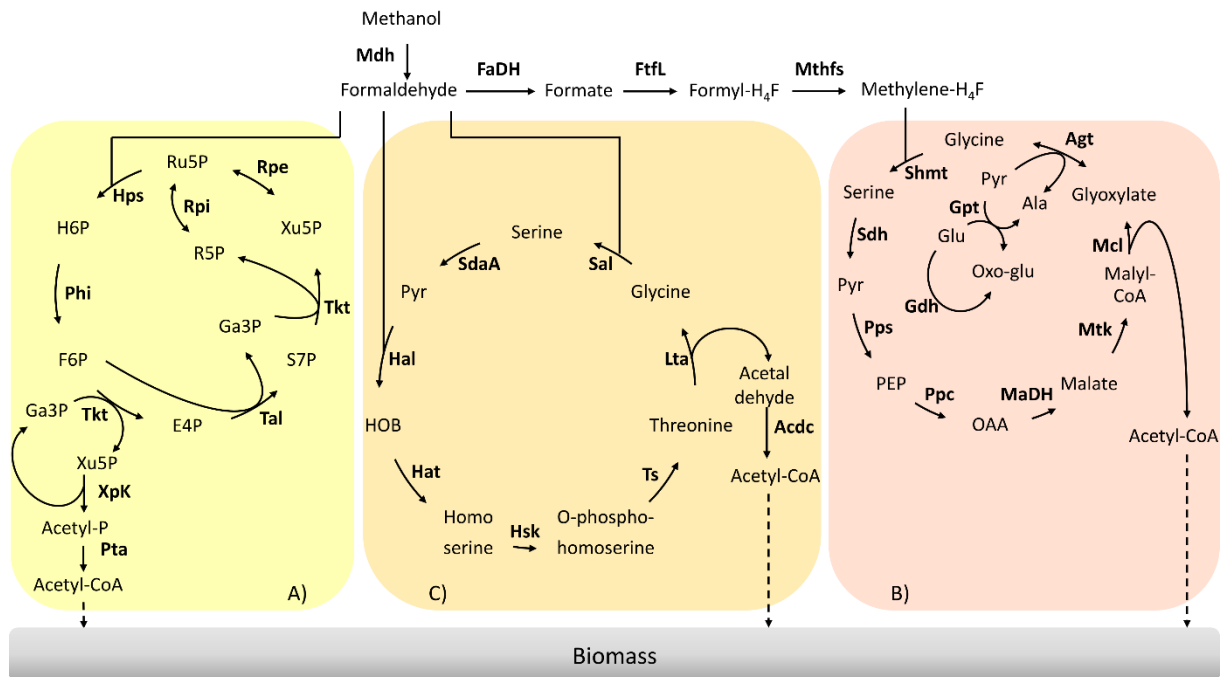


Figure 9: Synthetic cyclic pathways implemented for C1-assimilation. (A) Methanol condensation cycle pathway (B) Modified serine cycle pathway (C) Homoserine cycle pathway. Adapted from (Tuyishime and Sinumwayo, 2020). Enzymes are written in bold.

Acdc, acetaldehyde dehydrogenase; *Agt*, Alanine-glyoxylate aminotransferase; *FaDH*, formaldehyde dehydrogenase; *Ftl*, formate- H_4F ligase; *Gdh*, glutamate dehydrogenase; *Gpt*, Glutamate-pyruvate transaminase; *Hps*, 3-hexulose-6-phosphate synthase; *Hal*, 4-hydroxy-2-oxobutanoate aldolase; *Hat*, 4-hydroxy-2-oxobutanoate aminotransferase; *Hsk*, homoserine kinase; *Lta*, threonine aldolase; *MaDH*, malate dehydrogenase; *Mcl*, malyl-CoA lyase; *Mdh*, methanol dehydrogenase; *Mthfs*, methylene- H_4F dehydrogenase; *Mtk*, malate thiokinase; *Phi*, 6-phospho-3-hexuloisomerase; *Ppc*, phosphoenolpyruvate carboxylase; *Pps*, phosphoenolpyruvate synthetase; *Pta*, phosphate acetyltransferase; *Rpe*, ribulose-3-phosphate epimerase; *Rpi*, ribose-5-phosphate isomerase; *Sal*, serine aldolase; *SdaA*, serine deaminase; *Sdh*, serine, dehydratase; *Shmt*, serine- H_4F hydroxymethyltransferase; *Tal*, transaldolase; *Tkt*, transketolase; *Ts*, threonine synthase; *XpK*, xilulose-5-phosphate phosphoketolase. *Ala*, alanine; *E4P*, erythrose-4-phosphate; *F6P*, fructose-6-phosphate; *Ga3P*, glyceraldehyde-3-phosphate; *Glu*, glutamate; *H6P*, hexulose-6-phosphate; *OAA*, oxaloacetate; *Oxo-Glu*, 2-oxoglutarate; *PEP*, phosphoenolpyruvate; *Pyr*, pyruvate; *R5P*, ribose-5-phosphate; *Ru5P*, ribulose-5-phosphate; *S7P*, sedoheptulose-7-phosphate; *SBP*, sedoheptulose-1,7-biphosphate; *Xu5P*, xylulose-5-phosphate.

a) Methanol condensation cycle

In 2014, Bogorad et al designed a non-natural pathway enabling methanol assimilation named methanol condensation cycle (MCC) (Figure 9A) (Bogorad et al., 2014). The MCC pathway is combining the nonoxidative part of the PPP with the RuMP pathway. The MCC pathway was first constructed *in silico* and then tested in a cell-free environment. This pathway ensures the conservation of phosphates in the catalytic cycle as well as the balance of the redox state. The MCC was found functional and acetyl-CoA production was achieved *in vitro*. Acetyl-CoA is a precursor for the production of biofuels (i.e. ethanol and butanol). However, there is no report of the *in vivo* functionality of the MCC pathway yet.

b) The serine and homoserine cycles

One interesting feature with the serine cycle is that it enables the synthesis of acetyl-CoA, a key C2 building block. Therefore, its introduction in production hosts would be of great interest. Yu and Liao implemented a modified serine cycle in *E. coli* (Figure 9B) (Yu and Liao, 2018). Methanol is assimilated via the H₄F-dependent pathway by being converted into formate. Rather than “copy-paste” the serine cycle, they decided to select a set of enzymes to replace key reactions in the natural pathway by other reactions to fit the activity of this pathway with *E. coli* metabolism. *E. coli* metabolic network was modified to become dependent of methanol to grow. Once the operation of the two modules was confirmed (i.e. formate conversion to CH₂H₄F and the modified serine cycle), the complete pathway was tested *in vivo* by using ¹³C-methanol and xylose. Label derived from methanol was found in intracellular metabolites. Later, He and al designed a complete homoserine cycle based on *E. coli* native enzymes and the promiscuous activity of the formaldehyde aldolase (Figure 9C) (He et al., 2020). Unlike the serine cycle, here the C1-intermediate for assimilation is formaldehyde. A codon-optimized Mdh from *C. glutamicum* was used to oxidise methanol to formaldehyde. Then, the *E. coli* native metabolic network was modified in order to design a strain in which methanol assimilation is required for serine biosynthesis. The functionality of the homoserine cycle was tested *in vivo* with labelling experiment by using *E. coli* auxotrophic strains in which the complete formaldehyde dissimilation pathway was blocked. However, optimization is still required to achieve *E. coli* growth using the homoserine cycle.

c) The XuMP pathway

Dai et al introduced the XuMP pathway in the non-methylotrophic yeast *S. cerevisiae* by integrating AO, Cat, Das and Dak from *P. pastoris* in its genome (Dai et al., 2017). They showed the *in vivo* operation of the XuMP pathway on pure methanol. By using the native peroxisome targeting peptide of *P. pastoris*, enzymes were addressed to *S. cerevisiae* peroxisome. When yeast extract was added to the medium, methanol consumption was further improved from 1.04 g/L to 2.35 g/L, suggesting that Xu5P recycling was enhanced by yeast extract addition.

2. Linear pathways

As mentioned before, the common feature of the native methylotrophic assimilation pathways is that these processes are cyclic to ensure a continuous recycling of the C1-acceptor. With the introduction of linear synthetic methylotrophic pathways, formaldehyde assimilation becomes independent from any C1-acceptor (Figure 10).

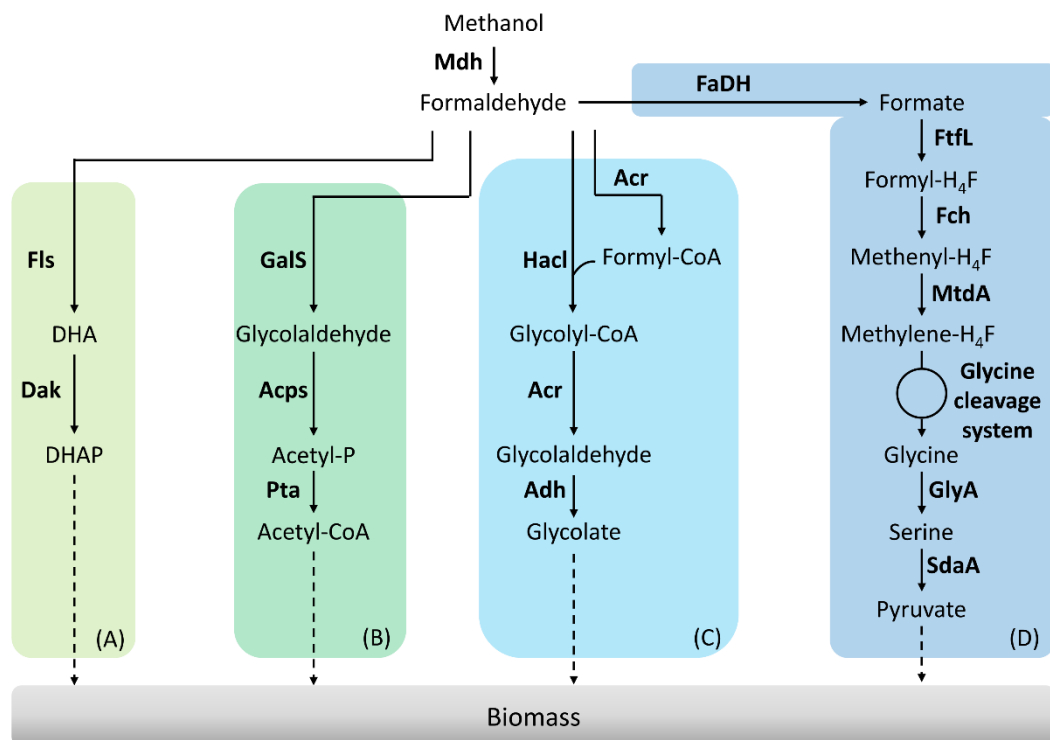


Figure 10: Linear pathways implemented for C1-assimilation. (A) Formolase pathway (B) Synthetic acetyl-CoA (SACA) pathway (C) 2-hydroxyacyl-CoA lyase (HACL) pathway (D) Reductive glycine pathway. Adapted from (Tuyishime and Sinumvayo, 2020). Enzymes are written in bold.

Acps, acetylphosphate synthase; *Acr*, acyl-CoA reductase; *Adh*, aldehyde dehydrogenase; *Dak*, dihydroxyacetone kinase; *FaDH*, formaldehyde dehydrogenase; *Fch*, methenyl H₄F-cyclohydrolase; *Fls*, formolase; *Ftl*, formate-H₄F ligase; *GalS*, glycolaldehyde synthase; *GlyA*, serine hydroxymethyltransferase; *Hacl*, 2-hydroxyacyl-CoA lyase; *Mdh*, methanol dehydrogenase; *MtdA*, methylene-H₄F dehydrogenase/methylene-H₄MPT dehydrogenase; *Pta*, phosphate acetyltransferase; *SdaA*, serine deaminase. *DHA*, dihydroxyacetone; *DHAP*, dihydroxyacetone phosphate.

a) Formolase pathway

Wang et al were the first in succeeding to implement a linear synthetic pathway (Wang et al., 2017). This pathway relies on two enzymes, an NAD-dependent Mdh and formolase (Fls) (Figure 10A). Fls is a synthetic enzyme, computationally designed by Siegel et al (Siegel et al., 2015). Fls condenses three molecules of formaldehyde to produce one molecule of DHA. The cooperation of the Mdh from *B. methanolicus* MGA3 or PB1 were first tested with Fls *in vitro*. From this experiment, they decided to introduce the Mdh from *B. methanolicus* PB1 combined with Fls in an *E. coli* strain deleted for *frmA*. To improve the performance of this strain, ALE was used. By using ^{13}C -methanol, the resulting evolved strain was shown to have higher level of labelling in the proteinogenic amino acids than the parental strain.

b) SACA pathway

The synthetic acetyl-CoA (SACA) is a linear pathway that was designed and constructed by Lu et al (Lu et al., 2019) (Figure 10B). This pathway was first designed to produce acetyl-CoA from formaldehyde. It relies on three enzymes. A glycolaldehyde synthase (GalS) condensates two molecules of formaldehyde to produce one molecule of glycolaldehyde. Then, glycolaldehyde is converted into acetyl-phosphate by the acetyl-phosphate synthase. Acetyl-phosphate is subsequently used to produce acetyl-CoA by the phosphate acetyltransferase. The authors first designed GalS and checked for glycolaldehyde production. Then, they optimized GalS by using directed evolution to improve the kinetic properties of the enzyme. The SACA pathway was functional *in vitro* and the authors then decided to implement the pathway in *E. coli*. The functionality of the pathway was tested *in vivo* in rich media. By combining the SACA pathway with the Mdh from *B. stearrowthermophilus*, they showed a slight improvement in the final OD reached. When tested on minimal media by using ^{13}C -methanol, label was also found in some proteinogenic amino acids after 26h.

c) HAcl pathway

The 2-hydroxyacyl-CoA lyase (Hacl) is a synthetic pathway designed by Chou et al. where Hacl condenses formaldehyde with formyl-CoA to produce glycolyl-CoA (Figure 10C) (Chou et al., 2019). Glycolyl-CoA is subsequently converted to glycolaldehyde by acyl-CoA (Acr). Finally, aldehyde dehydrogenase produces glycolate from glycolaldehyde. Hacl is a mammalian enzyme involved in α -oxidation. However, after Blast research limited to prokaryotes, authors

identified and tested one variant from *Rhodospiralleles bacterium* named RuHacI exhibiting the condensation activity of formaldehyde with formyl-CoA. They tested the pathway *in vivo* in an *E. coli* strain deleted for $\Delta frmA$, $\Delta fdhF\Delta fdnG\Delta fdoG$ and $\Delta glcD$ to avoid any competitive reactions using formaldehyde, formate or gluconate. The strain produced 0.5 g/L of glycolate corresponding to a yield of 67%. After enzyme engineering using direct evolution, a variant of RuHacI with improved enzyme kinetics was selected. By expressing this variant, *E. coli* produced 1.2 g/L of glycolate. However, the pathway has not been tested yet *in vitro* or *in vivo* on methanol with the addition of a Mdh.

d) Reductive glycine pathway

Later, Kim et al designed a fourth linear pathway: the reductive glycine pathway which, combined with Mdh, enabled methanol assimilation in *E. coli* (Kim et al., 2020) (Figure 10 D). In this pathway, formate is the key C1-intermediate that enables carbon assimilation. The authors introduced the H₄F-pathway from *M. extorquens* associated with a glycine cleavage system in an *E. coli* strain auxotrophic for serine, glycine and C1 moieties ($\Delta serA\Delta kbl\Delta ltaE\Delta aceA$). It has been previously described that even if growth could not be supported on formate, the latter compound could be still assimilated in *E. coli* via the H₄F-dependent pathway (Tashiro et al., 2018). Kim et al optimized the operation of this pathway first on formate. Then, they optimized the strain using ALE before implementing the Mdh in the strain. The selected strain was able to grow on formate and CO₂ at a growth rate of $\mu = 0.086 \text{ h}^{-1}$. In this strain, to be assimilated methanol is oxidised to formaldehyde by Mdh and then formaldehyde is oxidized to formate *via* the native GSH-dependent pathway of *E. coli* encoded by the operon *frmRAB*. Several Mdhs were tested and only the Mdh from *B. stearothermophilus* supported growth on methanol and CO₂ at a growth rate of $\mu = 0.013 \text{ h}^{-1}$. To confirm that the methanol pathway was functional, they confirmed that *frmA* deletion blocked growth on methanol. By using ¹³C-methanol, they showed that methanol carbon atoms were recovered in proteinogenic amino acids. The labelling pattern found was the same than when the strain grew on ¹³C-formate and ¹²CO₂.

Among all the attempts to engineer a synthetic methylotrophy in the platform host organisms cited above, it is very interesting to notice that only one succeeded to achieve pure synthetic methylotrophy using a natural cyclic pathway (Chen et al., 2020) while another one achieved to co-utilize methanol with CO₂ using a synthetic linear methylotrophic pathway (Kim

et al., 2020). In both examples, rational and evolutionary approaches have been used. Even if different NAD-dependent Mdhs were used, both are exhibiting improved kinetic parameters for methanol oxidation (i.e. engineered version of Mdh from *C. necator* in (Chen et al., 2020) and Mdh from *B. stearothermophilus* in (Kim et al., 2020)). Moreover, in both cases, auxotrophic strains were first designed prior to applying ALE. On one hand ALE was used to isolate an evolved strain exhibiting a new phenotype (i.e. synthetic methylotrophic strain) (Chen et al., 2020) while on the other hand ALE enabled to select an evolved strain with improved growth characteristics (Kim et al., 2020). Similarly, in both cases, mutations in genes using redox cofactors were found, confirming that the cellular redox balance is a key determinant for the successful implementation of a C1-assimilation pathway.

IV. Bacterial microcompartments to spatially organize prokaryotes

Until recently, synthetic biologists had left out the exploration of the spatial dimension for metabolic engineering in prokaryotes, as they do not have any particular cell-organization. However, the spatial organization of enzymes offers great advantages (Figure 11) (Lee et al., 2012). Indeed, thanks to the spatial organisation, substrate channelling is enhanced as enzymes and substrates are located in the same space. The local concentrations of substrates are increased, hence allowing a higher conversion rate via the encapsulated-enzymes compared to the cytoplasmic ones. Moreover, thanks to the barrier procured by the shell of the microcompartment, volatile or toxic intermediates are sequestered, protecting cells from their possible damages (Sampson and Bobik, 2008) while the production of side compounds by competing reactions is prevented.

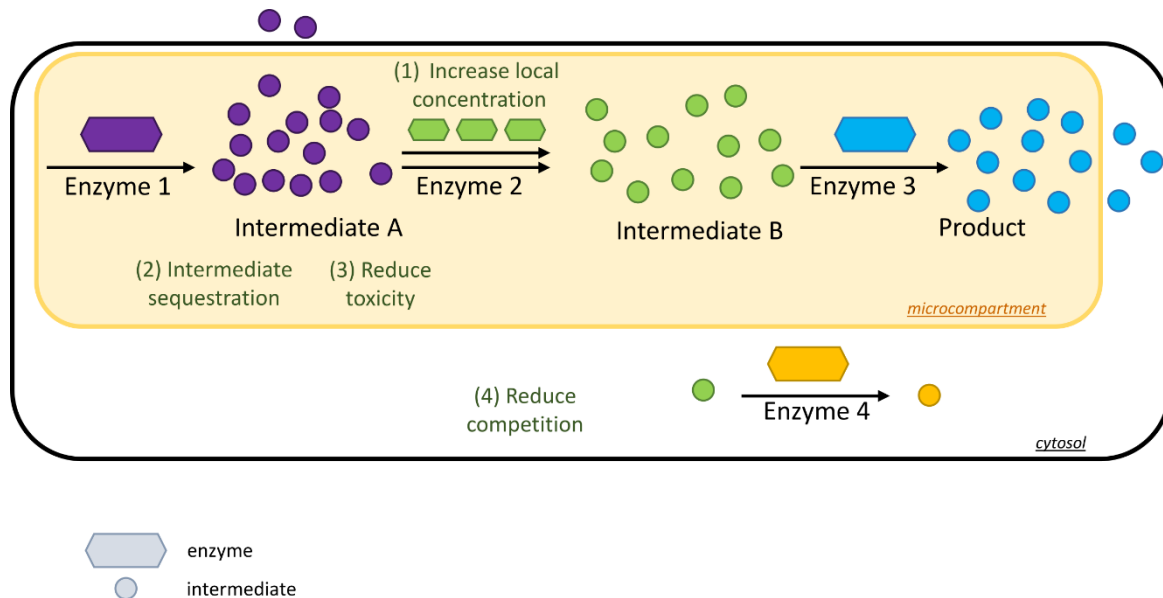


Figure 11: Advantages given by the compartmentalization of metabolic pathways. Adapted from (Chen and Silver, 2012).

A. Bacterial microcompartments (BMCs)

The compartmentalization of the major cellular functions is a key evolutionary event in eukaryotic organisms. This cell organization have long time been thought to be specific of this domain. However, in the 50's, for the first time, some bacterial microcompartments (BMCs) were found in prokaryotic cells. Unlike lipid-based eukaryotic organelles, BMCs are semipermeable proteinaceous organelles. They are located in the cytoplasm and have a diameter of 100-200 nm (BOBIK et al., 1999). BMCs are large protein complexes comprising thousands of shell proteins with an overall mass between 100 and 600 MDa (Cheng et al., 2008). BMC presence in prokaryotes is not anecdotal. A study revealed that around 17% of bacteria contain a BMC locus in their genome and that microcompartments are found in 23 different phyla (Axen et al., 2014). BMCs are classified in two main categories: BMCs involved in anabolism named carboxysomes and BMCs involved in catabolism named metabolosomes.

1. Anabolic BMCs: carboxysomes

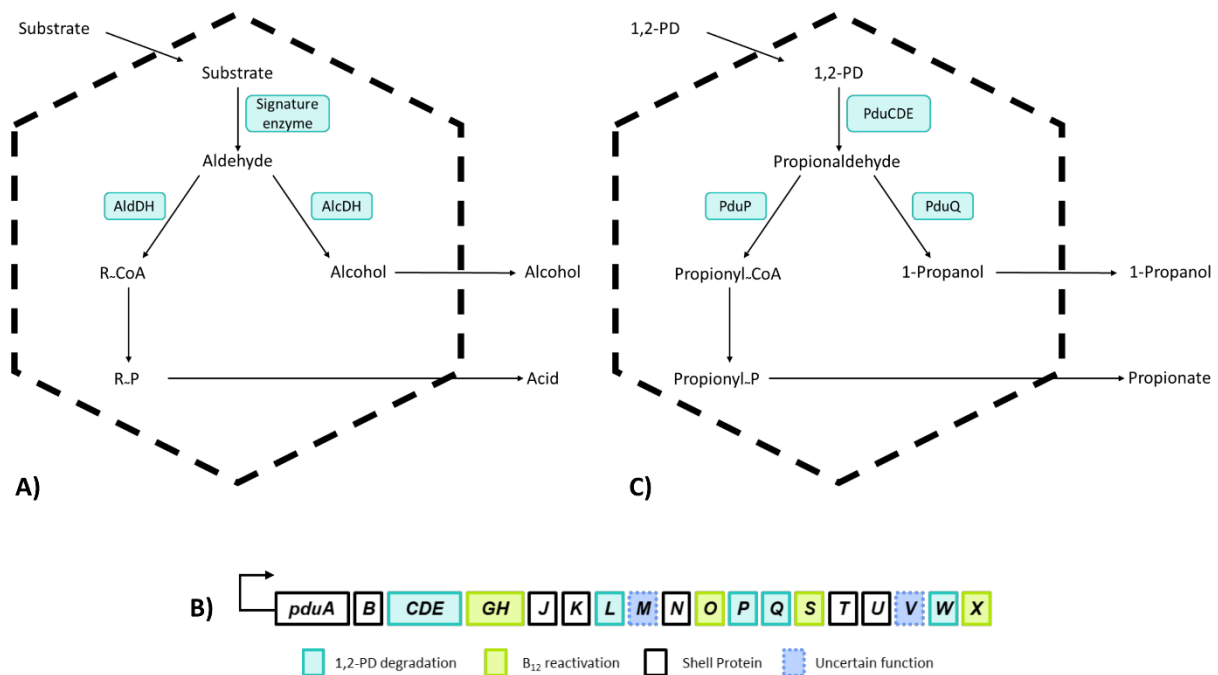
Carboxysomes are involved in carbon fixation in chemoautotrophs and cyanobacteria (see part I.C.1) (Turmo et al., 2017). Carbonic anhydrase and RuBisCO are both encapsulated in the carboxysome and respectively generates CO_2 and fixes CO_2 . RuBisCO is a bifunctional enzyme,

if it uses O_2 as a substrate instead of CO_2 , only one molecule of 3PG is produced with one molecule of phosphoglycolate. The latter molecule cannot be used in CBB cycle and is recycled via photorespiration. Carboxysomal shell proteins are suggested to act as a diffusion barrier by retaining CO_2 in the lumen of the BMC and by blocking O_2 to reach the RuBisCO inside the BMC, the local concentration of CO_2 inside the BMC is increased, enabling RuBisCO carbon fixation activity to be more efficient (Yeates et al., 2008).

Two classes of carboxysomes are found in nature depending on the form of encapsulated RuBisCO: a first class containing type 1A RuBisCO are defined as α -carboxysomes and a second class containing type 1B RuBisCO as β -carboxysomes (Yeates et al., 2008). α - and β -carboxysomes do not have the same genetic organization: genes encoding for α -carboxysomes are found in a single operon whereas genes encoding for β -carboxysomes are found in multiple gene clusters in the genome (Cannon et al., 2001). Both α - and β -carboxysomes form icosahedral structures but β -carboxysomes have a larger structure (175-500 nm in diameter) than α -carboxysomes (average size of 120 nm) (Rae et al., 2013).

2. Catabolic BMCs: metabolosomes

Metabolosomes are degradative BMCs enabling the assimilation of a wide range of substrates acting either as carbon or nitrogen sources, including 1,2-propanediol (Pdu BMC) (BOBIK et al., 1999), ethanolamine (Eut BMC) (KOFOID et al., 1999), choline (Herring et al., 2018), rhamnose and fucose (Petit et al., 2013). No matter the type of metabolosome, the substrate is first converted into an aldehyde by a “signature enzyme”. This aldehyde is then detoxified into its alcohol and acid derivatives (Figure 12A).



 Bacterial Microcompartment

Figure 12: Metabolic processes within metabolosomes. (A) Schematic representation of encapsulated pathways within metabolosomes. (B) Native Pdu operon from *Citrobacter freundii*. (C) 1,2-propanediol degradation pathway within the Pdu BMC. Adapted from (Kerfeld and Erbilgin, 2015).

AlcDH, alcohol dehydrogenase ; *AldDH*, aldehyde dehydrogenase. 1,2-PD, 1,2-propanediol.

Among all the metabolosomes, the Pdu BMC is the best characterized thanks to recombinant systems and *in vivo* studies (Parsons et al., 2008; Sampson and Bobik, 2008). 1,2-propanediol (1,2-PD) is a fermentative product of fucose and rhamnose degradation (Badía et al., 1985). In *Citrobacter freundii*, genes involved in 1,2-PD utilization are organized in an operon of 21 genes (Figure 12B). Seven genes encode for enzymes involved in 1,2-PD degradation (*pduCDELPQW*), five genes are involved B₁₂ regeneration (*pduGHOSX*) and seven genes encode for eight shell proteins (*(pduABJKNU)*, PduB' being an isoform of PduB) forming a polyhedral structure. Two additional genes *pocR* and *pduF*, coding respectively for a regulator and a transporter, are located upstream of the operon (Havemann and Bobik, 2003). Bobik et al determined that the diameter of the native Pdu BMC of *Salmonella enterica*

was between 100-200 nm (BOBIK et al., 1999). Few years later, Mayer et al determined that the diameter of purified Pdu BMCs from *C. freundii* is around 125 nm (Mayer et al., 2016).

1,2-PD is first converted to propionaldehyde by the B₁₂-dependent dioldehydratase PduCDE (Bobik et al., 1997). Propionaldehyde, being a toxic intermediate in the 1,2-PD degradation pathway, is either converted to propionyl-CoA by propionyl-CoA dehydrogenase (PduP) or to 1-propanol by 1-propanol dehydrogenase (PduQ) (Figure 12C). Then, propionyl-CoA is converted to propionyl-phosphate by the phosphotranscyclase (PduL) to finally be converted to propionate by a propionate kinase (PduW). Any disruption in the BMC shell results in an accumulation of propionaldehyde leading to a growth arrest (Cheng et al., 2011).

The Eut BMC is also well described (Chowdhury et al., 2014). Genes for ethanolamine degradation are organized in an operon of 17 genes: five genes encode for Eut shell proteins, eight genes encode for enzymes involved in ethanolamine degradation, one gene is involved in B₁₂ assimilation and recycling, one encodes for a transcription factor, one for a chaperone and one for an ethanolamine permease. Ethanolamine degradation is similar to 1,2-PD degradation. A signature enzyme, a B₁₂-dependent ethanolamine ammonia lyase, converts ethanolamine into acetaldehyde and ammonia. Acetaldehyde is a volatile and toxic compound which is then converted either to ethanol or to Ac-CoA, then to acetyl-phosphate, which is ultimately converted to acetate.

B. BMCs characteristics

1. Shell assembly

BMC shell is composed of building blocks named shell proteins. Shell proteins are classified in three categories: hexamers (BMC-H) and trimers (BMC-T) that are both forming hexagonal tiles and pentamers (BMC-P) (Table 5). BMC-H shell proteins are composed of six identical subunits containing one BMC domain (Pfam: PF00936) with a concave and convex sides (Kerfeld, 2005). BMC-T shell proteins are pseudo-hexameric proteins composed of three identical subunits containing two fused PF00936 domains. BMC-P shell proteins are composed of five identical subunits containing a single BMC domain (PF03319). BMC-P proteins are thought to cap the vertices of the microcompartment (Tanaka et al., 2010). BMC-H and BMC-T contain both a central pore in the middle of their structure (Crowley et al., 2010). Shell proteins forming hexagonal tiles are the major components of the BMC shell (Havemann

and Bobik, 2003). Based on structural and sequence analysis data, hexameric and pentameric BMC domains were shown to be strongly conserved among the BMC subtypes (Frank et al., 2013).




Name	BMC-H	BMC-T	BMC-P
Representation			
Function	Facet	Facet	Vertices
Protein domain	1 PF00936 domain	2 fused PF00936 domains	1 PF03319 domain
Pore size	4-7 Å	~12 Å	Unknown
Examples	PduA, J, K, U	PduB, T	PduN

Table 5: Classification of BMC shell proteins. Adapted from (Lee et al., 2018).

Despite the high degree of conservation between BMC shell proteins, the protein cargo can be very different from a BMC to another one. While the knowledge on shell proteins is increasing, the study of the BMC assembly is just starting. Studies were first done on carboxysomes and two types of assembly are proposed (Kerfeld et al., 2018). In β -carboxysomes, the core assembles first via protein domain interaction or most commonly *via* aggregation of proteins carrying an encapsulation peptide and then the shell is formed around the protein cargo (Kerfeld and Erbilgin, 2015). Moreover, it was shown that encapsulation peptides were interacting with shell proteins (Fan et al., 2012; Kinney et al., 2012). In the other type of assembly, found in α -carboxysomes, core and shell proteins are formed concomitantly. To date, metabolosome formation has not been elucidated but some data suggest that it follows the same assembly mechanism as β -carboxysomes. In the other type of assembly, found in α -carboxysomes, core and shell proteins are formed concomitantly.

2. Protein encapsulation

Proteins are encapsulated into BMCs thanks to encapsulation peptides that are located on either the N- or C- terminus. The lengths of these peptides ranges from 15 to 20 amino acids. They form amphipathic helices that interact with shell proteins via a common hydrophobic motif (Jakobson et al., 2015; Lawrence et al., 2014). It was shown that the encapsulation peptide of the PduP protein (P18) is 18 amino acids long and interacts with PduK shell protein (Lawrence et al., 2014). Moreover, encapsulation peptides can be used to encapsulate non-native proteins such as GFP inside BMCs, by adding the N-terminal 18 amino acids of PduD (D18), P18 or the N-terminal 20 amino acids of PduL (L20) on the N-terminal sequence of the GFP (Fan and Bobik, 2011; Fan et al., 2010; Liu et al., 2015). This result opens the way of BMCs refunctioning by addressing non-native enzymes and by extending the pathway inside them. However, one has to keep in mind that the nature of the encapsulation peptide can have an effect on the activity of the encapsulated protein. Indeed, some enzyme activities were affected by the addition of the encapsulation peptides whereas others were affected by the nature of the encapsulation peptide (P18 or D18). For some others the encapsulation peptide addition did not have an effect (Lee et al., 2016). This illustrates the complexity of predicting the effect of the addition of encapsulation peptides on encapsulated protein activity.

3. Pores and transport

The protein shell acts like a barrier between the lumen of the BMC and the cytosolic environment. Because of the close association between the shell proteins, the only channel for metabolites to access the lumen of the BMC is by diffusion via the central pore of the hexamers and pseudo-hexamers.

Some pores, such as the ones found in BMC-H proteins, are assumed to be selectively permeable to substrates and products (Chowdhury et al., 2015). Thanks to crystallographic structures, it was possible to determine that BMC-H proteins had a pore diameter of ~ 6 Å (Crowley et al., 2010). In the PduA pore, numerous hydrogen-bonds donors and acceptors are found whose arrangements are suggested to ease the entry of the polar 1,2-PD molecule and to keep the toxic propionaldehyde inside the Pdu BMC (Chowdhury et al., 2015).

Crystallography studies of BMC-T proteins revealed that these proteins can undergo conformational changes. They can contain a large pore (12-14 Å) depending on their

conformation: open or closed (Larsson et al., 2017). Ligands were found to bind to the smaller pore and were suggested to act as allosteric regulators for the pore opening (Thompson et al., 2015) or to reflect a specific metabolites selection (Pang et al., 2012). The open pore conformation is assumed to allow the entry of larger molecules in the BMC lumen to regenerate the encapsulated cofactors. However, the open conformation theory is still debated in the community as it was shown that the Eut BMC could function with its private pool of cofactors (Huseby and Roth, 2013).

C. BMCs for metabolic engineering in prokaryotes

Synthetic biologists have looked for new ways to implement metabolic pathways in production hosts. Clearly, BMCs are a challenging solution but they give access to a wide range of possibilities as they act as metabolic modules that can be customized. By giving them new functions thanks to the encapsulation of non-native pathways, it is possible to take advantage of the enhancement of the substrate channelling and cell protection from toxic intermediates given by BMCs.

1. Recombinant production of BMCs in heterologous hosts

To give new functions to BMCs for industrial or commercial purposes, one of the first milestones is to recombinantly produce them in production hosts as *E. coli* or *C. glutamicum*. The first step was to transfer the complete *pdu* operon from *C. freundii* and express it in *E. coli* (Parsons et al., 2008). Then, few years later, a recombinant Pdu BMC was produced by only expressing seven shell proteins, i.e. PduA-B-B'-J-K-N-U, described as an empty Pdu BMC (i.e. eBMC) (Parsons et al., 2010a). PduA, B, J, K and N were found to be essential in the formation of the eBMC while PduT and U appeared to be non-essential. Although PduU is unessential in the eBMC formation, its presence generates larger BMC structures (Parsons et al., 2010a). In a recent study, it was shown that the level of protein aggregation correlates with the size of Pdu BMCs while eBMCs formed smaller microcompartments (Juodeikis et al., 2020). Even though encapsulation peptides play a key role in the assembly of BMCs, the fact that it is possible to produce eBMCs with only shell proteins shows that a high degree of self-assembly is existing (Parsons et al., 2010a). Recombinant production of the Pdu BMC was also achieved in *C. glutamicum* by overexpressing *pduABJKNUT* operon from *C. freundii* (Huber et al.).

By expressing the 10-genes operon from *Halothiobacillus neapolitanus*, α -carboxysomes were produced in *E. coli*. Structures similar to native carboxysomes were visualized by transmission electron microscopy (TEM) and CO₂ fixation was achieved *in vitro* on purified carboxysomes (Bonacci et al., 2012). Based on this work, α -carboxysome production was also achieved in *C. glutamicum* and RuBisCO was found active in crude extracts (Baumgart et al., 2017).

Several attempts to recombinantly produce Eut BMCs in *E. coli* have been made (Choudhary et al., 2012; Held et al., 2016; Quin et al., 2016). Even though, the structures produced were claimed to be BMCs, it is still argued in the community. Indeed, according to some authors, these structures are most likely in-vivo aggregates and the single hexameric shell structures described are most likely to be membranous vesicles.

2. Redesign of eBMCs by the encapsulation of non-native proteins

In 2010, Fan et al showed that it was possible to target non-native proteins (here GFP) into the Pdu BMC by fusing them with an encapsulation peptide (Fan et al., 2010). This proof of concept added to the possibility to produce recombinantly eBMCs in industrial hosts opened the way to “plug and play” with BMCs to create specific metabolic modules (Figure 13). In 2014, an ethanol bioreactor was constructed in *E. coli* by co-expressing an eBMC with a pyruvate carboxylase and alcohol dehydrogenase tagged with encapsulation peptides (Lawrence et al., 2014). By using such ethanol bioreactors, more ethanol was produced compared to the untagged enzymes with the eBMC. Later, a 5-step pathway to produce 1,2-propanediol (1,2-PD) from glycerol was introduced in *E. coli* by targeting four enzymes into eBMCs (Lee et al., 2016). The strain expressing the tagged enzymes produced more 1,2-PD than the strain expressing the untagged enzymes. Interestingly, the strain expressing both the tagged enzymes and the eBMC produced less 1,2-PD than the one expressing only the tagged enzymes. In this strain, the production of 1,2-PD improved by 245%. It is most likely that thanks to the nucleation role of the encapsulation peptides, enzymes were co-localized as a protein aggregate enabling the enhancement of the pathway efficiency.

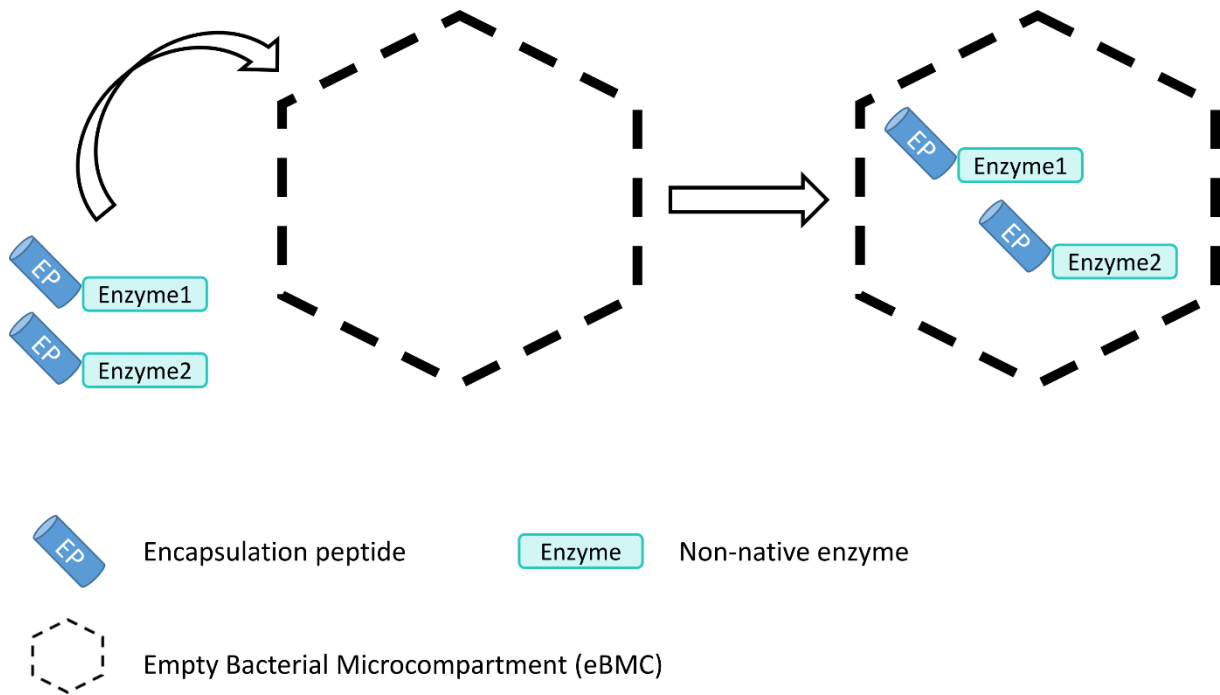


Figure 13: Redesign of eBMC by the addition of encapsulation peptides to non-native enzymes.

Without a doubt, the redesign of eBMCs gives new possibilities in synthetic biology for the implementation of pathways with toxic or volatile intermediates or for the enhancement of the substrate channelling. However, two bottlenecks may prevent scientists to reach their goal. The first one is that the addition of an encapsulation peptide can affect the activity of the enzymes to be encapsulated. Right now, it is very difficult to predict the behaviour of a newly tagged enzyme. Indeed, it was shown that despite the level of similarity between the same enzyme coming from different organisms, the enzyme activity was not impacted in the same way by the addition of the different encapsulation peptides tested (Juodeikis et al., 2020). The second problem is the possible low encapsulation efficiency of the non-native enzymes (Choudhary et al., 2012; Parsons et al., 2010b).

V. Aims and outline of the thesis

By using synthetic and system biology, this thesis aims to functionally integrate and compartmentalize a new synthetic methylotrophic pathway in *E. coli*, which has never been studied before (Figure 14). The methanol utilization (MUT) pathway is composed of a bacterial methanol dehydrogenase (Mdh) and a yeast dihydroxyacetone synthase (Das). To achieve the thesis goal, in **Chapter 2** the metabolic network of the microbial chassis *E. coli* was rationally engineered to fit with the methylotrophic properties. In **Chapter 3**, the MUT pathway was designed and functionally implemented into the modified chassis. Finally, in **Chapter 4**, the MUT pathway was spatially optimized by encapsulating it into an eBMC to engineer a methylotrophic BMC (mBMC).

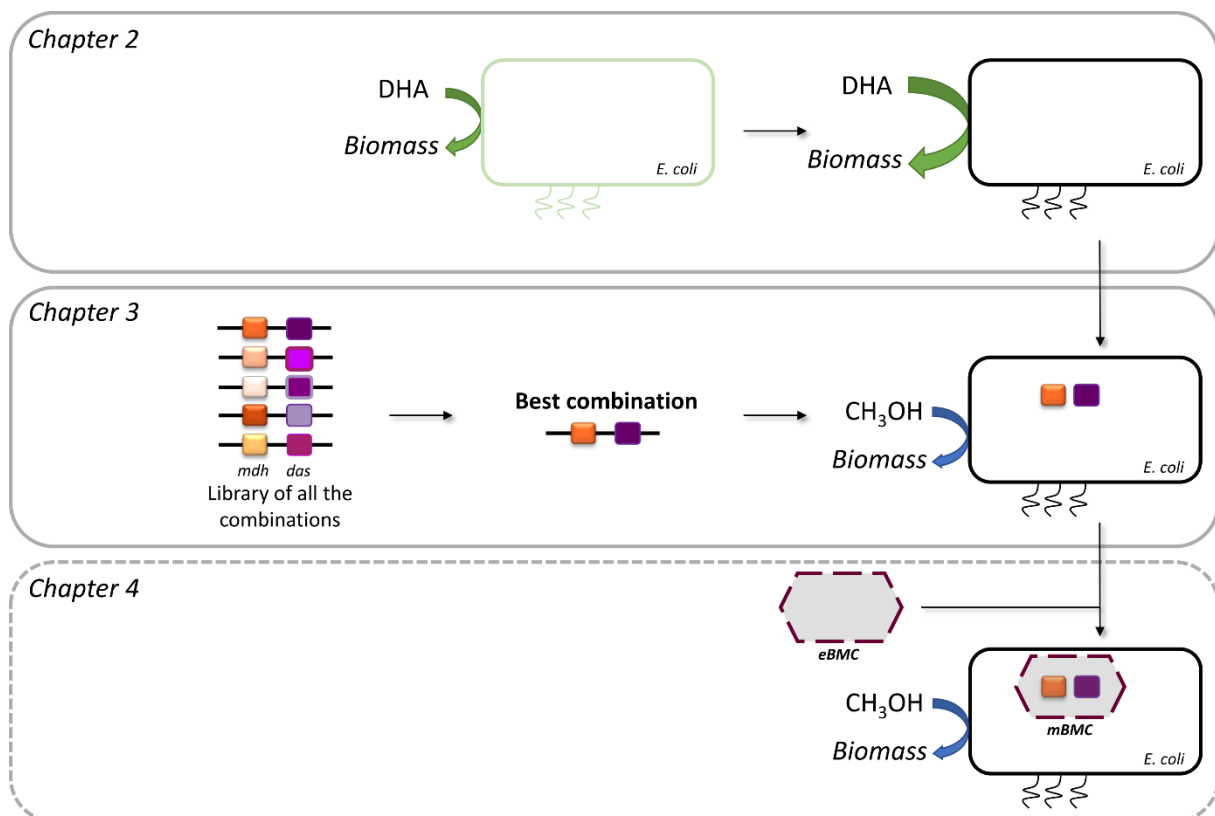


Figure 14: Schematic overview of the thesis strategy to achieve synthetic compartmentalized methylotrophy in *E. coli*.

Full line represents achieved work and dashed line represents work still on going.

eBMC, empty bacterial microcompartment; *mBMC*, methylotrophic bacterial microcompartment. *Das*, dihydroxyacetone synthase; *Mdh*, methanol dehydrogenase. *DHA*, dihydroxyacetone.

In **Chapter 2**, the native metabolism of *E. coli* grown on DHA was investigated. Indeed, one key intermediate of the MUT pathway is DHA, which is involved in both C1 assimilation and Xu5P recycling processes. This compound is a natural carbon source for *E. coli* but its utilization remains unclear. Therefore, we hypothesized that a better understanding of DHA metabolism might help optimizing methanol assimilation. Constraint-based modelling showed that *E. coli* growth on DHA was suboptimal. We identified that optimal growth on DHA was impaired by chemical, hierarchical and probably also by allosteric constraints. Functional analysis showed that DHA was mainly assimilated via the dihydroxyacetone kinase pathway. In addition, two other pathways, predicted by the model to be involved in DHA metabolism (i.e. glycerol and F6P aldolase pathways), turned out to not be fully activated during growth on DHA. By overexpressing some genes involved in these two pathways, constraints were released, and the optimal growth predicted by the model was restored.

In **Chapter 3**, the MUT pathway was built and tested in *E. coli*. The MUT pathway was built by a combinatorial library of *mdh* and *das* encoding genes from native and non-native methylotrophs (bacteria and yeast). All the selected sequences were assembled in a library of 266 combinations of genes (i.e. 14 for *mdh* * 19 *das*) and incorporated into *E. coli*. Methanol assimilation was followed using dynamic ¹³C-labelling experiment. The highest level of methanol conversion into central metabolism intermediates (e.g. 22% at the PEP-level), was obtained using a variant composed of a Mdh from *Acetobacter gertneri* and a codon-optimized version of Das from *P. angusta*. Finally, the activity of the MUT pathway was further improved by engineering strategic metabolic targets identified in **Chapter 3**. The final synthetic strain incorporated 1.5 to 5.9 times more methanol into intracellular metabolites and proteinogenic amino acids than the starting strain did.

In **Chapter 4**, we went one step further by encapsulating the best combination of Mdh and Das identified in **Chapter 3** into a bacterial microcompartment (BMC). First, the effect of the addition of two encapsulation peptides on Mdh and Das activities was studied to choose the one affecting the less their enzymatic activities. Tagged Mdh and Das were expressed together with the eBMCs resulting in methylotrophic BMCs (mBMCs). TEM analysis demonstrated the mBMCs formation in *E. coli*. Then, ¹H-NMR analysis on purified mBMCs showed that DHA was produced from methanol and Xu5P. Finally, we observed that methanol conversion was enhanced in the mBMCs compared to non-encapsulated enzymes.

VI. References

- Arfman, N., Watling, E.M., Clement, W., van Oosterwijk, R.J., de Vries, G.E., Harder, W., Attwood, M.M., and Dijkhuizen, L. (1989). Methanol metabolism in thermotolerant methylotrophic *Bacillus* strains involving a novel catabolic NAD-dependent methanol dehydrogenase as a key enzyme. *Arch. Microbiol.* *152*, 280–288.
- Arfman, N., Hektor, H.J., Bystrykh, L.V., Govorukhina, N.I., Dijkhuizen, L., and Frank, J. (1997). Properties of an NAD(H)-Containing Methanol Dehydrogenase and its Activator Protein from *Bacillus methanolicus*. *Eur J Biochem* *244*, 426–433.
- Axen, S.D., Erbilgin, O., and Kerfeld, C.A. (2014). A Taxonomy of Bacterial Microcompartment Loci Constructed by a Novel Scoring Method. *PLoS Comput Biol* *10*, e1003898.
- Bélangier, L., Figueira, M.M., Bourque, D., Morel, L., Béland, M., Laramée, L., Groleau, D., and Miguez, C.B. (2004). Production of heterologous protein by *Methylobacterium extorquens* in high cell density fermentation. *FEMS Microbiology Letters* *231*, 197–204.
- Badía, J., Ros, J., and Aguilar, J. (1985). Fermentation mechanism of fucose and rhamnose in *Salmonella typhimurium* and *Klebsiella pneumoniae*. *Journal of Bacteriology* *161*, 435–437.
- Baker, S.C., Ferguson, S.J., Ludwig, B., Page, M.D., Richter, O.-M.H., and van Spanning, R.J.M. (1998). Molecular Genetics of the Genus *Paracoccus*: Metabolically Versatile Bacteria with Bioenergetic Flexibility. *Microbiology and Molecular Biology Reviews* *62*, 1046–1078.
- Baumann, K., Carnicer, M., Dragosits, M., Graf, A.B., Stadlmann, J., Jouhten, P., Maaheimo, H., Gasser, B., Albiol, J., Mattanovich, D., et al. (2010). A multi-level study of recombinant *Pichia pastoris* in different oxygen conditions. *BMC Syst Biol* *4*, 141.
- Baumgart, M., Huber, I., Abdollahzadeh, I., Gensch, T., and Frunzke, J. (2017). Heterologous expression of the *Halothiobacillus neapolitanus* carboxysomal gene cluster in *Corynebacterium glutamicum*. *Journal of Biotechnology* *258*, 126–135.
- Becker, J., and Wittmann, C. (2015). Advanced Biotechnology: Metabolically Engineered Cells for the Bio-Based Production of Chemicals and Fuels, Materials, and Health-Care Products. *Angew. Chem. Int. Ed.* *54*, 3328–3350.
- Bennett, R.K., Gonzalez, J.E., Whitaker, W.B., Antoniewicz, M.R., and 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.
- Bennett, R.K., Dillon, M., Gerald Har, J.R., Agee, A., von Hagel, B., Rohlhill, J., Antoniewicz, M.R., and Papoutsakis, E.T. (2020). Engineering *Escherichia coli* for methanol-dependent growth on glucose for metabolite production. *Metabolic Engineering* *60*, 45–55.
- Bobik, T.A., Xu, Y., Jeter, R.M., Otto, K.E., and Roth, J.R. (1997). Propanediol utilization genes (*pdu*) of *Salmonella typhimurium*: three genes for the propanediol dehydratase. *Journal of Bacteriology* *179*, 6633–6639.
- BOBIK, T.A., HAVEMANN, G.D., BUSCH, R.J., WILLIAMS, D.S., and ALDRICH, H.C. (1999). The Propanediol Utilization (*pdu*) Operon of *Salmonella enterica* Serovar Typhimurium LT2 Includes Genes Necessary for Formation of Polyhedral Organelles Involved in Coenzyme B₁₂-Dependent 1,2-Propanediol Degradation[†]. *J. BACTERIOL.* *181*, 9.
- Bogorad, I.W., Chen, C.-T., Theisen, M.K., Wu, T.-Y., Schlenz, A.R., Lam, A.T., and Liao, J.C. (2014). Building carbon–carbon bonds using a biocatalytic methanol condensation cycle. *Proc Natl Acad Sci USA* *111*, 15928–15933.
- Bonacci, W., Teng, P.K., Afonso, B., Niederholtmeyer, H., Grob, P., Silver, P.A., and Savage, D.F. (2012). Modularity of a carbon-fixing protein organelle. *Proceedings of the National Academy of Sciences* *109*, 478–483.

- Bourque, D., Pomerleau, Y., and Groleau, D. High-cell-density production of poly- γ -hydroxybutyrate (PHB) from methanol by *Methylobacterium extorquens*: production of high-molecular-mass PHB. 10.
- Bozdog, A., Komives, C., and Flickinger, M.C. (2015). Growth of *Bacillus methanolicus* in 2 M methanol at 50 °C: the effect of high methanol concentration on gene regulation of enzymes involved in formaldehyde detoxification by the ribulose monophosphate pathway. *J Ind Microbiol Biotechnol* 42, 1027–1038.
- Brautaset, T., Jakobsen, O.M., Flickinger, M.C., Valla, S., and Ellingsen, T.E. (2004). Plasmid-Dependent Methylotrophy in Thermotolerant *Bacillus methanolicus*. *Journal of Bacteriology* 186, 1229–1238.
- Brautaset, T., Jakobsen, Ø.M., Degnes, K.F., Netzer, R., Nærdal, I., Krog, A., Dillingham, R., Flickinger, M.C., and Ellingsen, T.E. (2010). *Bacillus methanolicus* pyruvate carboxylase and homoserine dehydrogenase I and II and their roles for l-lysine production from methanol at 50°C. *Appl Microbiol Biotechnol* 87, 951–964.
- Cannon, G.C., Bradburne, C.E., Aldrich, H.C., Baker, S.H., Heinhorst, S., and Shively, J.M. (2001). Microcompartments in Prokaryotes: Carboxysomes and Related Polyhedra. *Appl. Environ. Microbiol.* 67, 5351–5361.
- Chen, A.H., and Silver, P.A. (2012). Designing biological compartmentalization. *Trends in Cell Biology* 22, 662–670.
- Chen, C.-T., Chen, F.Y.-H., Bogorad, I.W., Wu, T.-Y., Zhang, R., Lee, A.S., and Liao, J.C. (2018). Synthetic methanol auxotrophy of *Escherichia coli* for methanol-dependent growth and production. *Metabolic Engineering* 49, 257–266.
- Chen, F.Y.-H., Jung, H.-W., Tsuei, C.-Y., and Liao, J.C. (2020). Converting *Escherichia coli* to a Synthetic Methylotroph Growing Solely on Methanol. *Cell* 182, 933-946.e14.
- Chen, Y., Crombie, A., Rahman, M.T., Dedysh, S.N., Liesack, W., Stott, M.B., Alam, M., Theisen, A.R., Murrell, J.C., and Dunfield, P.F. (2010). Complete Genome Sequence of the Aerobic Facultative Methanotroph *Methylocella silvestris* BL2. *Journal of Bacteriology* 192, 3840–3841.
- Cheng, S., Liu, Y., Crowley, C.S., Yeates, T.O., and Bobik, T.A. (2008). Bacterial microcompartments: their properties and paradoxes. *BioEssays* 30, 1084–1095.
- Cheng, S., Sinha, S., Fan, C., Liu, Y., and Bobik, T.A. (2011). Genetic Analysis of the Protein Shell of the Microcompartments Involved in Coenzyme B12-Dependent 1,2-Propanediol Degradation by *Salmonella*. *Journal of Bacteriology* 193, 1385–1392.
- Chistoserdova, L. (1998). C1 Transfer Enzymes and Coenzymes Linking Methylotrophic Bacteria and Methanogenic Archaea. *Science* 281, 99–102.
- Chistoserdova, L. (2011). Modularity of methylotrophy, revisited: Modularity of methylotrophy, revisited. *Environmental Microbiology* 13, 2603–2622.
- Chistoserdova, L., Laukel, M., Portais, J.-C., Vorholt, J.A., and Lidstrom, M.E. (2004). Multiple Formate Dehydrogenase Enzymes in the Facultative Methylotroph *Methylobacterium extorquens* AM1 Are Dispensable for Growth on Methanol. *Journal of Bacteriology* 186, 22–28.
- Chistoserdova, L., Crowther, G.J., Vorholt, J.A., Skovran, E., Portais, J.-C., and Lidstrom, M.E. (2007). Identification of a Fourth Formate Dehydrogenase in *Methylobacterium extorquens* AM1 and Confirmation of the Essential Role of Formate Oxidation in Methylotrophy. *Journal of Bacteriology* 189, 9076–9081.
- Chistoserdova, L., Kalyuzhnaya, M.G., and Lidstrom, M.E. (2009). The Expanding World of Methylotrophic Metabolism. *Annual Review of Microbiology* 63, 477–499.
- Choi, Y.J., Gringorten, J.L., Belanger, L., Morel, L., Bourque, D., Masson, L., Groleau, D., and Miguez, C.B. (2008). Production of an Insecticidal Crystal Protein from *Bacillus thuringiensis* by the Methylotroph *Methylobacterium extorquens*. *Applied and Environmental Microbiology* 74, 5178–5182.
- Chou, A., Clomburg, J.M., Qian, S., and Gonzalez, R. (2019). 2-Hydroxyacyl-CoA lyase catalyzes acyloin condensation for one-carbon bioconversion. *Nat Chem Biol* 15, 900–906.

- Choudhary, S., Quin, M.B., Sanders, M.A., Johnson, E.T., and Schmidt-Dannert, C. (2012). Engineered Protein Nano-Compartments for Targeted Enzyme Localization. *PLoS ONE* 7, e33342.
- Chowdhury, C., Sinha, S., Chun, S., Yeates, T.O., and Bobik, T.A. (2014). Diverse Bacterial Microcompartment Organelles. *Microbiol Mol Biol Rev* 78, 438–468.
- Chowdhury, C., Chun, S., Pang, A., Sawaya, M.R., Sinha, S., Yeates, T.O., and Bobik, T.A. (2015). Selective molecular transport through the protein shell of a bacterial microcompartment organelle. *Proc Natl Acad Sci USA* 112, 2990–2995.
- Crowley, C.S., Cascio, D., Sawaya, M.R., Kopstein, J.S., Bobik, T.A., and Yeates, T.O. (2010). Structural Insight into the Mechanisms of Transport across the *Salmonella enterica* Pdu Microcompartment Shell. *J. Biol. Chem.* 285, 37838–37846.
- Crowther, G.J., Kosaly, G., and Lidstrom, M.E. (2008). Formate as the Main Branch Point for Methylophilic Metabolism in *Methylobacterium extorquens* AM1. *Journal of Bacteriology* 190, 5057–5062.
- Dai, Z., Gu, H., Zhang, S., Xin, F., Zhang, W., Dong, W., Ma, J., Jia, H., and Jiang, M. (2017). Metabolic construction strategies for direct methanol utilization in *Saccharomyces cerevisiae*. *Bioresource Technology* 245, 1407–1412.
- Delépine, B., López, M.G., Carnicer, M., Vicente, C.M., Wendisch, V.F., and Heux, S. (2020). Charting the Metabolic Landscape of the Facultative Methylophilic *Bacillus methanolicus*. 5, 16.
- Denby, K.J., Iwig, J., Bisson, C., Westwood, J., Rolfe, M.D., Sedelnikova, S.E., Higgins, K., Maroney, M.J., Baker, P.J., Chivers, P.T., et al. (2016). The mechanism of a formaldehyde-sensing transcriptional regulator. *Sci Rep* 6, 38879.
- Dragosits, M., and Mattanovich, D. (2013). Adaptive laboratory evolution – principles and applications for biotechnology. *Microb Cell Fact* 12, 64.
- Erb, T.J., Berg, I.A., Brecht, V., Muller, M., Fuchs, G., and Alber, B.E. (2007). Synthesis of C5-dicarboxylic acids from C2-units involving crotonyl-CoA carboxylase/reductase: The ethylmalonyl-CoA pathway. *Proceedings of the National Academy of Sciences* 104, 10631–10636.
- Erythropel, H.C., Zimmerman, J.B., de Winter, T.M., Petitjean, L., Melnikov, F., Lam, C.H., Lounsbury, A.W., Mellor, K.E., Janković, N.Z., Tu, Q., et al. (2018). The Green ChemistREE: 20 years after taking root with the 12 principles. *Green Chem.* 20, 1929–1961.
- Ettwig, K.F., Butler, M.K., Le Paslier, D., Pelletier, E., Mangenot, S., Kuypers, M.M.M., Schreiber, F., Dutilh, B.E., Zedelius, J., de Beer, D., et al. (2010). Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. *Nature* 464, 543–548.
- Fan, C., and Bobik, T.A. (2011). The N-Terminal Region of the Medium Subunit (PduD) Packages Adenosylcobalamin-Dependent Diol Dehydratase (PduCDE) into the Pdu Microcompartment. *Journal of Bacteriology* 193, 5623–5628.
- Fan, C., Cheng, S., Liu, Y., Escobar, C.M., Crowley, C.S., Jefferson, R.E., Yeates, T.O., and Bobik, T.A. (2010). Short N-terminal sequences package proteins into bacterial microcompartments. *Proceedings of the National Academy of Sciences* 107, 7509–7514.
- Fan, C., Cheng, S., Sinha, S., and Bobik, T.A. (2012). Interactions between the termini of lumen enzymes and shell proteins mediate enzyme encapsulation into bacterial microcompartments. *Proceedings of the National Academy of Sciences* 109, 14995–15000.
- Fan, L., Wang, Y., Tuyishime, P., Gao, N., Li, Q., Zheng, P., Sun, J., and Ma, Y. (2018). Engineering Artificial Fusion Proteins for Enhanced Methanol Bioconversion. *ChemBioChem* 19, 2465–2471.
- Fischer, J.E., and Glieder, A. (2019). Current advances in engineering tools for *Pichia pastoris*. *Current Opinion in Biotechnology* 59, 175–181.
- Frank, S., Lawrence, A.D., Prentice, M.B., and Warren, M.J. (2013). Bacterial microcompartments moving into a synthetic biological world. *Journal of Biotechnology* 163, 273–279.

Ganesh, I., Vidhya, S., Eom, G.T., and Hong, S.H. (2017). Construction of Methanol-Sensing *Escherichia coli* by the Introduction of a *Paracoccus denitrificans* MxaY-Based Chimeric Two-Component System. *Journal of Microbiology and Biotechnology* 27, 1106–1111.

Goodwin, P.M., and Anthony, C. (1998). The Biochemistry, Physiology and Genetics of PQQ and PQQ-containing Enzymes. In *Advances in Microbial Physiology*, (Elsevier), pp. 1–80.

Gutiérrez, J., Bourque, D., Criado, R., Choi, Y.J., Cintas, L.M., Hernández, P.E., and Míguez, C.B. (2005). Heterologous extracellular production of enterocin P from *Enterococcus faecium* P13 in the methylotrophic bacterium *Methylobacterium extorquens*. *FEMS Microbiology Letters* 248, 125–131.

Hagemeyer, C.H., Chistoserdova, L., Lidstrom, M.E., Thauer, R.K., and Vorholt, J.A. (2000). Characterization of a second methylene tetrahydromethanopterin dehydrogenase from *Methylobacterium extorquens* AM1: Dehydrogenase from M. extorquens AM1. *European Journal of Biochemistry* 267, 3762–3769.

Havemann, G.D., and Bobik, T.A. (2003). Protein Content of Polyhedral Organelles Involved in Coenzyme B12-Dependent Degradation of 1,2-Propanediol in *Salmonella enterica* Serovar Typhimurium LT2. *Journal of Bacteriology* 185, 5086–5095.

He, H., Höper, R., Dodenhöft, M., Marlière, P., and Bar-Even, A. (2020). An optimized methanol assimilation pathway relying on promiscuous formaldehyde-condensing aldolases in *E. coli*. *Metabolic Engineering* 60, 1–13.

Heggeset, T.M.B., Krog, A., Balzer, S., Wentzel, A., Ellingsen, T.E., and Brautaset, T. (2012). Genome Sequence of Thermotolerant *Bacillus methanolicus*: Features and Regulation Related to Methylophony and Production of L-Lysine and L-Glutamate from Methanol. *Applied and Environmental Microbiology* 78, 5170–5181.

Held, M., Kolb, A., Perdue, S., Hsu, S.-Y., Bloch, S.E., Quin, M.B., and Schmidt-Dannert, C. (2016). Engineering formation of multiple recombinant Eut protein nanocompartments in *E. coli*. *Sci Rep* 6, 24359.

Hennig, G., Haupka, C., Brito, L.F., Rückert, C., Cahoreau, E., Heux, S., and Wendisch, V.F. (2020). Methanol-Essential Growth of *Corynebacterium glutamicum*: Adaptive Laboratory Evolution Overcomes Limitation due to Methanethiol Assimilation Pathway. *IJMS* 21, 3617.

Herring, T.I., Harris, T.N., Chowdhury, C., Mohanty, S.K., and Bobik, T.A. (2018). A Bacterial Microcompartment Is Used for Choline Fermentation by *Escherichia coli* 536. *J Bacteriol* 200, e00764-17, /jlb/200/10/e00764-17.atom.

Hu, B., and Lidstrom, M.E. (2014). Metabolic engineering of *Methylobacterium extorquens* AM1 for 1-butanol production. *Biotechnol Biofuels* 7, 156.

Huber, I., Palmer, D.J., Ludwig, K.N., Brown, I., Warren, M.J., and Frunzke, J. Construction of recombinant Pdu metabolosome shells for small molecule production in *Corynebacterium glutamicum*. *ACS Synthetic Biology* 40.

Huseby, D.L., and Roth, J.R. (2013). Evidence that a Metabolic Microcompartment Contains and Recycles Private Cofactor Pools. *Journal of Bacteriology* 195, 2864–2879.

Irla, M., Nærdal, I., Brautaset, T., and Wendisch, V.F. (2017). Methanol-based γ -aminobutyric acid (GABA) production by genetically engineered *Bacillus methanolicus* strains. *Industrial Crops and Products* 106, 12–20.

Jakobsen, O.M., Benichou, A., Flickinger, M.C., Valla, S., Ellingsen, T.E., and Brautaset, T. (2006). Upregulated Transcription of Plasmid and Chromosomal Ribulose Monophosphate Pathway Genes Is Critical for Methanol Assimilation Rate and Methanol Tolerance in the Methylotrophic Bacterium *Bacillus methanolicus*. *Journal of Bacteriology* 188, 3063–3072.

Jakobson, C.M., Kim, E.Y., Slininger, M.F., Chien, A., and Tullman-Ercek, D. (2015). Localization of Proteins to the 1,2-Propanediol Utilization Microcompartment by Non-native Signal Sequences Is Mediated by a Common Hydrophobic Motif. *Journal of Biological Chemistry* 290, 24519–24533.

Jordà, J., Jouhten, P., Cámara, E., Maaheimo, H., Albiol, J., and Ferrer, P. (2012). Metabolic flux profiling of recombinant protein secreting *Pichia pastoris* growing on glucose:methanol mixtures. *Microb Cell Fact* 11, 57.

Jordà, J., de Jesus, S.S., Peltier, S., Ferrer, P., and Albiol, J. (2014). Metabolic flux analysis of recombinant *Pichia pastoris* growing on different glycerol/methanol mixtures by iterative fitting of NMR-derived ¹³C-labelling data from proteinogenic amino acids. *New Biotechnology* 31, 120–132.

Jungo, C., Schenk, J., Pasquier, M., Marison, I.W., and von Stockar, U. (2007). A quantitative analysis of the benefits of mixed feeds of sorbitol and methanol for the production of recombinant avidin with *Pichia pastoris*. *Journal of Biotechnology* 131, 57–66.

Juodeikis, R., Lee, M.J., Mayer, M., Mantell, J., Brown, I.R., Verkade, P., Woolfson, D.N., Prentice, M.B., Frank, S., and Warren, M.J. (2020) Effect of metabolosome encapsulation peptides on enzyme activity, coaggregation, incorporation, and bacterial microcompartment formation. *MicrobiologyOpen*, e1010.

Kalyuzhnaya, M.G., Hristova, K.R., Lidstrom, M.E., and Chistoserdova, L. (2008). Characterization of a Novel Methanol Dehydrogenase in Representatives of Burkholderiales: Implications for Environmental Detection of Methylophony and Evidence for Convergent Evolution. *J Bacteriol* 190, 3817–3823.

Keltjens, J.T., Pol, A., Reimann, J., and Op den Camp, H.J.M. (2014). PQQ-dependent methanol dehydrogenases: rare-earth elements make a difference. *Appl Microbiol Biotechnol* 98, 6163–6183.

Kerfeld, C.A. (2005). Protein Structures Forming the Shell of Primitive Bacterial Organelles. *Science* 309, 936–938.

Kerfeld, C.A., and Erbilgin, O. (2015). Bacterial microcompartments and the modular construction of microbial metabolism. *Trends in Microbiology* 23, 22–34.

Kerfeld, C.A., Aussignargues, C., Zarzycki, J., Cai, F., and Sutter, M. (2018). Bacterial microcompartments. *Nature Reviews Microbiology* 16, 277–290.

Khadem, A.F., Pol, A., Wieczorek, A., Mohammadi, S.S., Francoijs, K.-J., Stunnenberg, H.G., Jetten, M.S.M., and Op den Camp, H.J.M. (2011). Autotrophic Methanotrophy in *Verrucomicrobia*: *Methylacidiphilum fumarolicum* SolV Uses the Calvin-Benson-Bassham Cycle for Carbon Dioxide Fixation. *Journal of Bacteriology* 193, 4438–4446.

Kiefer, P., Delmotte, N., and Vorholt, J.A. (2011). Nanoscale Ion-Pair Reversed-Phase HPLC–MS for Sensitive Metabolome Analysis. *Anal. Chem.* 83, 850–855.

Kim, S., Lindner, S.N., Aslan, S., Yishai, O., Wenk, S., Schann, K., and Bar-Even, A. (2020). Growth of *E. coli* on formate and methanol via the reductive glycine pathway. *Nat Chem Biol.*

Kinney, J.N., Salmeen, A., Cai, F., and Kerfeld, C.A. (2012). Elucidating Essential Role of Conserved Carboxysomal Protein CcmN Reveals Common Feature of Bacterial Microcompartment Assembly. *J. Biol. Chem.* 287, 17729–17736.

van der Klei, I.J., Yurimoto, H., Sakai, Y., and Veenhuis, M. (2006). The significance of peroxisomes in methanol metabolism in methylophonyic yeast. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1763, 1453–1462.

KOFOID, E., RAPPLEYE, C., STOJILJKOVIC, I., and ROTH, J. (1999). The 17-Gene Ethanolamine (eut) Operon of *Salmonella typhimurium* Encodes Five Homologues of Carboxysome Shell Proteins. *J. BACTERIOL.* 181, 13.

Koopman, F.W., de Winde, J.H., and Ruijssenaars, H.J. (2009). C1 compounds as auxiliary substrate for engineered *Pseudomonas putida* S12. *Appl Microbiol Biotechnol* 83, 705–713.

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

Larsson, A.M., Hasse, D., Valegård, K., and Andersson, I. (2017). Crystal structures of β -carboxysome shell protein CcmP: ligand binding correlates with the closed or open central pore. *Journal of Experimental Botany* 68, 3857–3867.

Laukel, M., Chistoserdova, L., Lidstrom, M.E., and Vorholt, J.A. (2003). The tungsten-containing formate dehydrogenase from *Methylobacterium extorquens* AM1: Purification and properties. *Eur J Biochem* 270, 325–333.

Lawrence, A.D., Frank, S., Newnham, S., Lee, M.J., Brown, I.R., Xue, W.-F., Rowe, M.L., Mulvihill, D.P., Prentice, M.B., Howard, M.J., et al. (2014). Solution Structure of a Bacterial Microcompartment Targeting Peptide and Its Application in the Construction of an Ethanol Bioreactor. *ACS Synthetic Biology* 3, 454–465.

Lee, H., DeLoache, W.C., and Dueber, J.E. (2012). Spatial organization of enzymes for metabolic engineering. *Metabolic Engineering* 14, 242–251.

Lee, M.J., Brown, I.R., Juodeikis, R., Frank, S., and Warren, M.J. (2016). Employing bacterial microcompartment technology to engineer a shell-free enzyme-aggregate for enhanced 1,2-propanediol production in *Escherichia coli*. *Metabolic Engineering* 36, 48–56.

Lee, M.J., Palmer, D.J., and Warren, M.J. (2018). Biotechnological Advances in Bacterial Microcompartment Technology. *Trends in Biotechnology*.

Leßmeier, L., Pfeifenschneider, J., Carnicer, M., Heux, S., Portais, J.-C., and Wendisch, V.F. (2015). Production of carbon-13-labeled cadaverine by engineered *Corynebacterium glutamicum* using carbon-13-labeled methanol as co-substrate. *Applied Microbiology and Biotechnology* 99, 10163–10176.

Liang, W.-F., Cui, L.-Y., Cui, J.-Y., Yu, K.-W., Yang, S., Wang, T.-M., Guan, C.-G., Zhang, C., and Xing, X.-H. (2017). Biosensor-assisted transcriptional regulator engineering for *Methylobacterium extorquens* AM1 to improve mevalonate synthesis by increasing the acetyl-CoA supply. *Metabolic Engineering* 39, 159–168.

Liu, Y., Jorda, J., Yeates, T.O., and Bobik, T.A. (2015). The PduL Phosphotransacylase Is Used To Recycle Coenzyme A within the Pdu Microcompartment. *J. Bacteriol.* 197, 2392–2399.

Liu, Y., Tu, X., Xu, Q., Bai, C., Kong, C., Liu, Q., Yu, J., Peng, Q., Zhou, X., Zhang, Y., et al. (2018). Engineered monoculture and co-culture of methylotrophic yeast for de novo production of monacolin J and lovastatin from methanol. *Metabolic Engineering* 45, 189–199.

Looser, V., Bruhlmann, B., Bumbak, F., Stenger, C., Costa, M., Camattari, A., Fotiadis, D., and Kovar, K. (2015). Cultivation strategies to enhance productivity of *Pichia pastoris*: A review. *Biotechnology Advances* 33, 1177–1193.

Lu, X., Liu, Y., Yang, Y., Wang, S., Wang, Q., Wang, X., Yan, Z., Cheng, J., Liu, C., Yang, X., et al. (2019). Constructing a synthetic pathway for acetyl-coenzyme A from one-carbon through enzyme design. *Nat Commun* 10, 1378.

Lüers, G.H., Advani, R., Wenzel, T., and Subramani, S. (1998). The *Pichia pastoris* dihydroxyacetone kinase is a PTS1-containing, but cytosolic, protein that is essential for growth on methanol. *Yeast* 14, 759–771.

Matsushita, K., Arents, J.C., Bader, R., Yamada, M., Adachi, O., and Postma, P.W. (1997). *Escherichia coli* is unable to produce pyrroloquinoline quinone (PQQ). *Microbiology* 143, 3149–3156.

Mayer, M.J., Juodeikis, R., Brown, I.R., Frank, S., Palmer, D.J., Deery, E., Beal, D.M., Xue, W.-F., and Warren, M.J. (2016). Effect of bio-engineering on size, shape, composition and rigidity of bacterial microcompartments. *Scientific Reports* 6.

Meyer, F., Keller, P., Hartl, J., Gröninger, O.G., Kiefer, P., and Vorholt, J.A. (2018). Methanol-essential growth of *Escherichia coli*. *Nature Communications* 9.

Mo, X.-H., Zhang, H., Wang, T.-M., Zhang, C., Zhang, C., Xing, X.-H., and Yang, S. (2020). Establishment of CRISPR interference in *Methylobacterium extorquens* and application of rapidly mining a new phytoene desaturase involved in carotenoid biosynthesis. *Appl Microbiol Biotechnol*.

Mokhtari-Hosseini, Z.B., Vasheghani-Farahani, E., Shojaosadati, S.A., Karimzadeh, R., and Heidarzadeh-Vazifekhoran, A. (2009). Effect of feed composition on PHB production from methanol by HCDC of *Methylobacterium extorquens* (DSMZ 1340). *J. Chem. Technol. Biotechnol.* 84, 1136–1139.

Müller, J.E.N., Meyer, F., Litsanov, B., Kiefer, P., Potthoff, E., Heux, S., Quax, W.J., Wendisch, V.F., Brautaset, T., Portais, J.-C., et al. (2015a). Engineering *Escherichia coli* for methanol conversion. *Metabolic Engineering* 28, 190–201.

Müller, J.E.N., Meyer, F., Litsanov, B., Kiefer, P., and Vorholt, J.A. (2015b). Core pathways operating during methylotrophy of *Bacillus methanolicus* MGA3 and induction of a bacillithiol-dependent detoxification pathway upon formaldehyde stress: Core pathways operating during methylotrophy. *Molecular Microbiology* 98, 1089–1100.

Müller, J.E.N., Heggeset, T.M.B., Wendisch, V.F., Vorholt, J.A., and Brautaset, T. (2015c). Methylotrophy in the thermophilic *Bacillus methanolicus*, basic insights and application for commodity production from methanol. *Applied Microbiology and Biotechnology* 99, 535–551.

Naerdal, I., Pfeifenschneider, J., Brautaset, T., and Wendisch, V.F. (2015). Methanol-based cadaverine production by genetically engineered *Bacillus methanolicus* strains: Cadaverine production by *Bacillus methanolicus*. *Microbial Biotechnology* 8, 342–350.

Nagy, P.L., Marolewski, A., Benkovic, S.J., and Zalkin, H. (1995). Formyltetrahydrofolate hydrolase, a regulatory enzyme that functions to balance pools of tetrahydrofolate and one-carbon tetrahydrofolate adducts in *Escherichia coli*. *Journal of Bacteriology* 177, 1292–1298.

Olah, G.A. (2013). Towards Oil Independence Through Renewable Methanol Chemistry. *Angew. Chem. Int. Ed.* 52, 104–107.

Orita, I., Sakamoto, N., Kato, N., Yurimoto, H., and Sakai, Y. (2007). Bifunctional enzyme fusion of 3-hexulose-6-phosphate synthase and 6-phospho-3-hexuloisomerase. *Appl Microbiol Biotechnol* 76, 439–445.

Orita, I., Nishikawa, K., Nakamura, S., and Fukui, T. (2014). Biosynthesis of polyhydroxyalkanoate copolymers from methanol by *Methylobacterium extorquens* AM1 and the engineered strains under cobalt-deficient conditions. *Appl Microbiol Biotechnol* 98, 3715–3725.

Ozimek, P., van Dijk, R., Latchev, K., Gancedo, C., Wang, D.Y., van der Kleij, I.J., and Veenhuis, M. (2003). Pyruvate Carboxylase Is an Essential Protein in the Assembly of Yeast Peroxisomal Oligomeric Alcohol Oxidase. *MBoC* 14, 786–797.

Pang, A., Liang, M., Prentice, M.B., and Pickersgill, R.W. (2012). Substrate channels revealed in the trimeric *Lactobacillus reuteri* bacterial microcompartment shell protein PduB. *Acta Crystallogr D Biol Crystallogr* 68, 1642–1652.

Parsons, J.B., Dinesh, S.D., Deery, E., Leech, H.K., Brindley, A.A., Heldt, D., Frank, S., Smales, C.M., Lünsdorf, H., Rambach, A., et al. (2008). Biochemical and Structural Insights into Bacterial Organelle Form and Biogenesis. *Journal of Biological Chemistry* 283, 14366–14375.

Parsons, J.B., Frank, S., Bhella, D., Liang, M., Prentice, M.B., Mulvihill, D.P., and Warren, M.J. (2010a). Synthesis of Empty Bacterial Microcompartments, Directed Organelle Protein Incorporation, and Evidence of Filament-Associated Organelle Movement. *Molecular Cell* 38, 305–315.

Parsons, J.B., Lawrence, A.D., McLean, K.J., Munro, A.W., Rigby, S.E.J., and Warren, M.J. (2010b). Characterisation of PduS, the pdu Metabolosome Corrin Reductase, and Evidence of Substructural Organisation within the Bacterial Microcompartment. *PLoS ONE* 5, e14009.

Peña, D.A., Gasser, B., Zanghellini, J., Steiger, M.G., and Mattanovich, D. (2018). Metabolic engineering of *Pichia pastoris*. *Metabolic Engineering* 50, 2–15.

Petit, E., LaTouf, W.G., Coppi, M.V., Warnick, T.A., Currie, D., Romashko, I., Deshpande, S., Haas, K., Alvelo-Maurosa, J.G., Wardman, C., et al. (2013). Involvement of a Bacterial Microcompartment in the Metabolism of Fucose and Rhamnose by *Clostridium phytofermentans*. *PLoS ONE* 8, e54337.

Peyraud, R., Kiefer, P., Christen, P., Massou, S., Portais, J.-C., and Vorholt, J.A. (2009). Demonstration of the ethylmalonyl-CoA pathway by using ¹³C metabolomics. *Proceedings of the National Academy of Sciences* 106, 4846–4851.

Peyraud, R., Schneider, K., Kiefer, P., Massou, S., Vorholt, J.A., and Portais, J.-C. (2011). Genome-scale reconstruction and system level investigation of the metabolic network of *Methylobacterium extorquens* AM1. *BMC Systems Biology* 5, 189.

Peyraud, R., Kiefer, P., Christen, P., Portais, J.-C., and Vorholt, J.A. (2012). Co-Consumption of Methanol and Succinate by *Methylobacterium extorquens* AM1. *PLoS ONE* 7, e48271.

Pfeifenschneider, J., Markert, B., Stolzenberger, J., Brautaset, T., and Wendisch, V.F. (2020). Transaldolase in *Bacillus methanolicus*: biochemical characterization and biological role in ribulose monophosphate cycle. *BMC Microbiol* 20, 63.

Price, J.V., Chen, L., Whitaker, W.B., Papoutsakis, E., and Chen, W. (2016). Scaffoldless engineered enzyme assembly for enhanced methanol utilization. *Proceedings of the National Academy of Sciences* 113, 12691–12696.

Puxbaum, V., Mattanovich, D., and Gasser, B. (2015). Quo vadis? The challenges of recombinant protein folding and secretion in *Pichia pastoris*. *Appl Microbiol Biotechnol* 99, 2925–2938.

Quin, M.B., Perdue, S.A., Hsu, S.-Y., and Schmidt-Dannert, C. (2016). Encapsulation of multiple cargo proteins within recombinant Eut nanocompartments. *Appl Microbiol Biotechnol* 100, 9187–9200.

Rae, B.D., Long, B.M., Badger, M.R., and Price, G.D. (2013). Functions, Compositions, and Evolution of the Two Types of Carboxysomes: Polyhedral Microcompartments That Facilitate CO₂ Fixation in Cyanobacteria and Some Proteobacteria. *Microbiology and Molecular Biology Reviews* 77, 357–379.

Reaser, B.C., Yang, S., Fitz, B.D., Parsons, B.A., Lidstrom, M.E., and Synovec, R.E. (2016). Non-targeted determination of ¹³C-labeling in the *Methylobacterium extorquens* AM1 metabolome using the two-dimensional mass cluster method and principal component analysis. *Journal of Chromatography A* 1432, 111–121.

Rohlfhill, J., Sandoval, N.R., and Papoutsakis, E.T. (2017). Sort-Seq Approach to Engineering a Formaldehyde-Inducible Promoter for Dynamically Regulated *Escherichia coli* Growth on Methanol. *ACS Synthetic Biology* 6, 1584–1595.

Rohlfhill, J., Gerald Har, J.R., Antoniewicz, M.R., and Papoutsakis, E.T. (2020). Improving synthetic methylotrophy via dynamic formaldehyde regulation of pentose phosphate pathway genes and redox perturbation. *Metabolic Engineering* 57, 247–255.

Roth, T.B., Woolston, B.M., Stephanopoulos, G., and Liu, D.R. (2019). Phage-Assisted Evolution of *Bacillus methanolicus* Methanol Dehydrogenase 2. *ACS Synth. Biol.* 8, 796–806.

Russmayer, H., Buchetics, M., Gruber, C., Valli, M., Grillitsch, K., Modarres, G., Guerrasio, R., Klavins, K., Neubauer, S., and Drexler, H. (2015). Systems-level organization of yeast methylotrophic lifestyle. *BMC Biology* 13, 80.

Sahm, H., and Wagner, F. (1973). Microbial Assimilation of Methanol. The Ethanol- and Methanol-Oxidizing Enzymes of the Yeast *Candida boidinii*. *Eur J Biochem* 36, 250–256.

Sampson, E.M., and Bobik, T.A. (2008). Microcompartments for B₁₂-Dependent 1,2-Propanediol Degradation Provide Protection from DNA and Cellular Damage by a Reactive Metabolic Intermediate. *JB* 190, 2966–2971.

Schneider, K., Peyraud, R., Kiefer, P., Christen, P., Delmotte, N., Massou, S., Portais, J.-C., and Vorholt, J.A. (2012). The Ethylmalonyl-CoA Pathway Is Used in Place of the Glyoxylate Cycle by *Methylobacterium extorquens* AM1 during Growth on Acetate. *Journal of Biological Chemistry* 287, 757–766.

Schrader, J., Schilling, M., Holtmann, D., Sell, D., Filho, M.V., Marx, A., and Vorholt, J.A. (2009). Methanol-based industrial biotechnology: current status and future perspectives of methylotrophic bacteria. *Trends in Biotechnology* 27, 107–115.

Schultenkämper, K., Brito, L.F., López, M.G., Brautaset, T., and Wendisch, V.F. (2019). Establishment and application of CRISPR interference to affect sporulation, hydrogen peroxide detoxification, and mannitol catabolism in the methylotrophic thermophile *Bacillus methanolicus*. *Appl Microbiol Biotechnol* 103, 5879–5889.

Shah, A.A., Hasan, F., Hameed, A., and Ahmed, S. (2008). Biological degradation of plastics: A comprehensive review. *Biotechnology Advances* 26, 246–265.

Sheehan, M.C., Bailey, C.J., Dowds, B.C.A., and McConnell, D.J. (1988). A new alcohol dehydrogenase, reactive towards methanol, from *Bacillus stearothermophilus*. *Biochemical Journal* 252, 661–666.

Siegel, J.B., Smith, A.L., Poust, S., Wargacki, A.J., Bar-Even, A., Louw, C., Shen, B.W., Eiben, C.B., Tran, H.M., Noor, E., et al. (2015). Computational protein design enables a novel one-carbon assimilation pathway. *Proc Natl Acad Sci USA* 112, 3704–3709.

Simakov, D.S.A. (2017). *Renewable Synthetic Fuels and Chemicals from Carbon Dioxide* (Cham: Springer International Publishing).

Sonntag, F., Buchhaupt, M., and Schrader, J. (2014). Thioesterases for ethylmalonyl-CoA pathway derived dicarboxylic acid production in *Methylobacterium extorquens* AM1. *Appl Microbiol Biotechnol* 98, 4533–4544.

Sonntag, F., Kroner, C., Lubuta, P., Peyraud, R., Horst, A., Buchhaupt, M., and Schrader, J. (2015). Engineering *Methylobacterium extorquens* for de novo synthesis of the sesquiterpenoid α -humulene from methanol. *Metabolic Engineering* 32, 82–94.

Tan, X., Titorenko, V.I., van der Klei, I.J., Sulter, G.J., Haima, P., Waterham, H.R., Evers, M., Harder, W., Veenhuis, M., and Cregg, J.M. (1995). Characterization of peroxisome-deficient mutants of *Hansenula polymorpha*. *Current Genetics* 28, 248–257.

Tanaka, S., Sawaya, M.R., and Yeates, T.O. (2010). Structure and Mechanisms of a Protein-Based Organelle in *Escherichia coli*. *Science* 327, 81–84.

Tashiro, Y., Hirano, S., Matson, M.M., Atsumi, S., and Kondo, A. (2018). Electrical-biological hybrid system for CO₂ reduction. *Metabolic Engineering* 47, 211–218.

Thompson, M.C., Cascio, D., Leibly, D.J., and Yeates, T.O. (2015). An allosteric model for control of pore opening by substrate binding in the EutL microcompartment shell protein: Pore Regulation by Substrate Binding. *Protein Science* 24, 956–975.

Tomàs-Gamisans, M., Ferrer, P., and Albiol, J. (2016). Integration and Validation of the Genome-Scale Metabolic Models of *Pichia pastoris*: A Comprehensive Update of Protein Glycosylation Pathways, Lipid and Energy Metabolism. *PLoS ONE* 11, e0148031.

Turmo, A., Gonzalez-Esquer, C.R., and Kerfeld, C.A. (2017). Carboxysomes: metabolic modules for CO₂ fixation. *FEMS Microbiology Letters* 364.

Tuyishime, P., and Sinumvayo, J.P. (2020). Novel outlook in engineering synthetic methylotrophs and formatotrophs: a course for advancing C₁-based chemicals production. *World J Microbiol Biotechnol* 36, 118.

Tuyishime, P., Wang, Y., Fan, L., Zhang, Q., Li, Q., Zheng, P., Sun, J., and Ma, Y. (2018). Engineering *Corynebacterium glutamicum* for methanol-dependent growth and glutamate production. *Metabolic Engineering* 49, 220–231.

Vartiainen, E., Blomberg, P., Ilmén, M., Andberg, M., Toivari, M., and Penttilä, M. (2019). Evaluation of synthetic formaldehyde and methanol assimilation pathways in *Yarrowia lipolytica*. *Fungal Biol Biotechnol* 6, 27.

Villadsen, J., Nielsen, J., and Lidén, G. (2011). *Bioreaction Engineering Principles* (Boston, MA: Springer US).

Vorholt, J. (2002). Cofactor-dependent pathways of formaldehyde oxidation in methylotrophic bacteria. *Archives of Microbiology* 178, 239–249.

Vorholt, J.A., Chistoserdova, L., Lidstrom, M.E., and Thauer, R.K. (1998). The NADP-Dependent Methylene Tetrahydromethanopterin Dehydrogenase in *Methylobacterium extorquens* AM1. *Journal of Bacteriology* 180, 5351–5356.

Wang, X., Wang, Y., Liu, J., Li, Q., Zhang, Z., Zheng, P., Lu, F., and Sun, J. (2017). Biological conversion of methanol by evolved *Escherichia coli* carrying a linear methanol assimilation pathway. *Bioresources and Bioprocessing* 4.

Wang, Y., Fan, L., Tuyishime, P., Zheng, P., and Sun, J. (2020). Synthetic Methylophony: A Practical Solution for Methanol-Based Biomanufacturing. *Trends in Biotechnology* S0167779919303087.

Wegner, E.H. (1983) (54) BOCHEMICAL CONVERSIONS BY YEAST FERMENTATION AT HIGH CELL DENSITIES. 16.

Weinhandl, K., Winkler, M., Glieder, A., and Camattari, A. (2014). Carbon source dependent promoters in yeasts. *Microb Cell Fact* 13, 5.

Whitaker, W.B., Jones, J.A., Bennett, R.K., Gonzalez, J.E., Vernacchio, V.R., Collins, S.M., Palmer, M.A., Schmidt, S., Antoniewicz, M.R., Koffas, M.A., et al. (2017). Engineering the biological conversion of methanol to specialty chemicals in *Escherichia coli*. *Metabolic Engineering* 39, 49–59.

Windass, J.D., Worsley, M.J., Pioli, E.M., Pioli, D., Barth, P.T., Atherton, K.T., Dart, E.C., Byrom, D., Powell, K., and Senior, P.J. (1980). Improved conversion of methanol to single-cell protein by *Methylophilus methylotrophus*. *Nature* 287, 396–401.

Witthoff, S., Schmitz, K., Niedenführ, S., Nöh, K., Noack, S., Bott, M., and Marienhagen, J. (2015). Metabolic Engineering of *Corynebacterium glutamicum* for Methanol Metabolism. *Applied and Environmental Microbiology* 81, 2215–2225.

Woolston, B.M., Roth, T., Kohale, I., Liu, D.R., and Stephanopoulos, G. (2017). Development of a formaldehyde biosensor with application to synthetic methylophony. *Biotechnology and Bioengineering*.

Woolston, B.M., King, J.R., Reiter, M., Van Hove, B., and Stephanopoulos, G. (2018). Improving formaldehyde consumption drives methanol assimilation in engineered *E. coli*. *Nature Communications* 9.

Wu, T.-Y., Chen, C.-T., Liu, J.T.-J., Bogorad, I.W., Damoiseaux, R., and Liao, J.C. (2016). Characterization and evolution of an activator-independent methanol dehydrogenase from *Cupriavidus necator* N-1. *Appl Microbiol Biotechnol* 100, 4969–4983.

Yeates, T.O., Kerfeld, C.A., Heinhorst, S., Cannon, G.C., and Shively, J.M. (2008). Protein-based organelles in bacteria: carboxysomes and related microcompartments. *Nat Rev Microbiol* 6, 681–691.

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

Yurimoto, H., Kato, N., and Sakai, Y. (2009). Genomic organization and biochemistry of the ribulose monophosphate pathway and its application in biotechnology. *Appl Microbiol Biotechnol* 84, 407–416.

Chapter 2

Chemical and metabolic controls on dihydroxyacetone metabolism lead to a suboptimal growth of *E. coli*.

Camille Peiro¹, Pierre Millard¹, Alessandro de Simone¹, Edern Cahoreau¹, Lindsay Peyriga¹, Brice Enjalbert¹, Stéphanie Heux¹

¹ LISBP, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France

Applied and Environmental Microbiology, 2019, Vol 85, Issue 15, e00768-19

Abstract

In this work, we shed light on the metabolism of dihydroxyacetone (DHA), a versatile, ubiquitous and important intermediate for various chemicals in industry, by analysing its metabolism at the system level in *E. coli*. Using constraints based modelling, we show that the growth of *E. coli* on DHA is suboptimal and we identify the potential causes. NMR analysis shows that DHA is degraded non-enzymatically into substrates known to be unfavourable to high growth rates. Transcriptomic analysis reveals that DHA promotes genes involved in biofilm formation, which may reduce the bacterial growth rate. Functional analysis of the genes involved in DHA metabolism proves that under the aerobic conditions used in this study, DHA is mainly assimilated via the dihydroxyacetone kinase pathway. In addition, these results show that the alternative routes of DHA assimilation (i.e. the glycerol and fructose-6-phosphate aldolase pathways) are not fully activated under our conditions because of anaerobic mediated hierarchical control. These pathways are therefore certainly unable to sustain fluxes as high as the ones predicted *in silico* for optimal aerobic growth on DHA. Overexpressing some of the genes in these pathways releases these constraints and restores the predicted optimal growth on DHA.

Introduction

Dihydroxyacetone (DHA) is an attractive molecule that is used as a final product in a wide range of industries (i.e. the food, cosmetic and pharmaceutical industries) (Ciriminna et al., 2018; Hekmat, 2003; Stanko, 1990) or as a growth substrate for various microorganisms (Gonzalez et al., 2008; Molin et al., 2003; Streekstra, 1987). DHA is a ubiquitous molecule found in all domains. A physiological product of the body naturally present in human urine (Bales et al., 1984), it also plays a role in the osmoregulation of yeast and algae (Akhtar et al., 1997). DHA is produced as an intermediate of various metabolic pathways. In methylotrophic yeast, DHA is a key intermediate in methanol assimilation (Ahmad et al., 2014; O'connor, 1980) while in bacteria, DHA is produced by aldol cleavage of the glycolytic intermediary fructose 6-phosphate (Sánchez-Moreno et al., 2012) or by oxidation of glycerol, the latter process being the basis for the current industrial production of DHA by *Gluconobacter oxydans* (Deppenmeier et al., 2002). DHA is also a highly reactive molecule and its accumulation is presumed to be toxic. Indeed, DHA can react with DNA and proteins in Maillard-type reactions and thus alter cell heredity (Benov and Beema, 2003; Tessier et al., 2003). In addition, DHA is unstable and can be converted non-enzymatically into several molecular species, among which highly toxic methylglyoxal (Lip et al., 2013; Maksimović et al., 2006; Subedi et al., 2008; Yaylayan et al., 1999).

Escherichia coli can metabolise DHA aerobically through at least three different metabolic pathways: (i) the dihydroxyacetone kinase (DAK) pathway, (ii) the glycerol (GLD) pathway and (iii) the fructose-6-phosphate (FSA) pathway (Subedi et al., 2008). The DAK pathway is named after dihydroxyacetone kinase, encoded by the *dhaKLM* operon. This operon is controlled by DhaR, a transcription factor activated by DHA (Bächler et al., 2005). DhaKLM is composed of three subunits (DhaK, DhaL and DhaM) and phosphorylates DHA to dihydroxyacetone phosphate (DHAP), an intermediate in the glycolytic pathway. DhaKLM has a high affinity for DHA ($K_m < 6\mu\text{M}$) and a catalytic constant (K_{cat}) of 4.8 s^{-1} (Garcia-Alles et al., 2004). This kinase resembles a phosphotransferase system (PTS) that uses phosphoenolpyruvate (PEP) as a phosphoryl donor (Jin and Lin, 1984). The DhaM subunit is first multiphosphorylated before the DhaL subunit uses the phosphate to convert ADP into ATP. Finally, this ATP is used directly by the DhaK subunit to phosphorylate DHA and form DHAP (Gutknecht et al., 2001). The fact

that *dhaK* transposon insertion prevents the growth of *E. coli* on DHA (Paulsen et al., 2000) indicates that DhaKLM is essential for this process (Paulsen et al., 2000).

The GLD pathway involves glycerol dehydrogenase (GldA) and glycerol-3-phosphate kinase (GlpK). Hydroxyacetone induces *gldA* expression in the stationary phase (Truniger and Boos, 1994) while the transcription of *glpK* is regulated by catabolite repression and by glycerol and glycerol-3-phosphate via the transcriptional repressor GlpR (Freedberg and Lin, 1973). During anaerobic glycerol fermentation, GldA enables the formation of DHA from glycerol (Gonzalez et al., 2008) whereas in aerobic conditions, GldA enables the reverse process, i.e. the formation of glycerol from DHA (Subedi et al., 2008; Tang et al., 1979). GldA has a greater affinity for DHA ($K_m = 0.3$ mM) than for glycerol ($K_m = 56$ mM) (Subedi et al., 2008) but the catalytic rate constants are close ($K_{cat} = 17.2$ and 22.4 s⁻¹, respectively). It has been proposed that the primary *in vivo* role of GldA is the removal of surplus dihydroxyacetone to limit its toxic effects (Subedi et al., 2008). However, DhaKLM and GldA have been shown to constitute a fermentative route for the conversion of glycerol to glycolytic intermediates (Durnin et al., 2009), demonstrating that GldA is also involved in fermentation of glycerol (Gonzalez et al., 2008). Once formed from DHA, glycerol is phosphorylated by GlpK to form glycerol-3-phosphate, which is central for lipid biosynthesis. GlpK has been shown to play a crucial role in DHA assimilation in various microorganisms (Jin et al., 1982; Kremer and Hansen, 1987; Ouellette et al., 2013; Weinhouse and Benziman, 1976). In *Haloferax volcanii* for instance, the deletion of *glpK* has a higher impact on growth than does *dhaKLM* deletion and mutants with both genes deleted do not grow on DHA (Ouellette et al., 2013). It has been hypothesized that in these organisms, GlpK can either phosphorylate DHA into DHAP or control the transport of DHA into the cells (Jin et al., 1982; Ouellette et al., 2013). In *E. coli*, *in vitro* results indicate that GlpK is able to use DHA as a substrate with a K_m value of 0.5 mM (Hayashi and Lin, 1967); however, its *in vivo* implication in DHA metabolism has never been studied.

The central enzyme in the FSA pathway is fructose-6-aldolase, which condenses DHA with GAP to form fructose-6-phosphate (F6P). Two different genes in *E. coli* (*fsaA* and *fsaB* or *mipB* and *talC*) code for FsaA and FsaB, respectively. These two enzymes share 70% identity and have similar affinities for DHA ($K_m = 32$ mM for FsaA and 27 mM for FsaB), but FsaA has a higher catalytic constant (k_{cat}) for DHA (116 s⁻¹ vs 41 s⁻¹) (Sánchez-Moreno et al., 2012). The

physiological role of *FsaA* in *E. coli* (Schurmann, 2001) and its regulation remain unknown. However, *fsaB* is in the same operon as *gldA* and *ptsA*, suggesting that *fsaB* may be metabolically associated with DHA (Subedi et al., 2008). The role of this operon and its regulation are not completely clear. However, *PtsA* has been identified as a component of a putative PTS system involved in anaerobic fructose catabolism (Reizer et al., 1995). In addition, higher expression of *gldA*, *fsaA* and *fsaB* has been observed in wild type *E. coli* grown on glucose during a transition from aerobic to anaerobic cultivation (Rolfe et al., 2012).

While several pathways can potentially support DHA catabolism in *E. coli*, which ones are actually involved during growth on DHA remains unclear, and their regulation is poorly characterised. This missing information would help unravel the physiological significance of this versatile, ubiquitous and high-potential metabolite. The objective of our study was to investigate DHA metabolism in *E. coli* using system level analysis. We first predicted the metabolic fate of DHA using a genome scale model of *E. coli* to evaluate its metabolic capabilities. We then used transcriptomic analysis to identify ‘genes that matter’ for growth on DHA. Based on these results, we characterized the physiology effect of overexpression and deletion of genes involved in DHA metabolism and we propose mechanisms that would explain the fate of DHA in *E. coli*.

Material and methods

Bacterial strains and plasmids

Escherichia coli K-12 BW25113 was selected as the model wild-type strain. Single deletion mutants were taken from the Keio collection (Baba et al., 2006) and the $\Delta dhaKLM$ strain was constructed using the gene deletion method described previously (Datsenko and Wanner, 2000). The *dhaKLM* operon was replaced by a CmR cassette using the primers listed in Table 1. All the antibiotic resistant cassettes were removed by FLP recombination. The plasmids listed in Table 2 were purchased from Baseclear (Leiden, Netherland). Cells were transformed according to the rubidium chloride protocol (Green and Rogers, 2013). All the strains are listed in Table 2 and genetic modifications were checked by PCR.

Plasmid, bacterial strain or primer	Description (genotype and/or relevant characteristics) or sequence of oligonucleotide primer	Reference or source
E. coli bacterial strains		
WT	BW25113: <i>rrnB3 ΔlacZ4787 hsdR514 Δ(araBAD)567 Δ(rhaBAD)568 rph-1</i>	(Baba et al., 2006)
ΔfsaA	BW25113 ΔfsaA::frr	This study
ΔfsaB	BW25113 ΔfsaB::frr	This study
ΔgldA	BW25113 ΔgldA::frr	This study
ΔglpK	BW25113 ΔglpK::frr	This study
ΔdhaKLM	BW25113 ΔdhaKLM::frr	This study
ΔptsA::kan	BW25113 ΔptsA::kan	(Baba et al., 2006)
ΔptsA	BW25113 ΔptsA::frr	This study
WT+p131	BW25113 + pSEVA131	This study
WT+p234	BW25113 + pSEVA234	This study
fsaA+++	BW25113 + pSEVA234-fsaA	This study
fsaB+++	BW25113 + pSEVA131-fsaB	This study
gldA+++	BW25113 + pSEVA131-gldA	This study
glpK+++	BW25113 + pSEVA131-glpK	This study
dhaKLM+++	BW25113 + pSEVA131-dhaKLM	This study
ΔfsaA+++	DfsaA + pSEVA234-fsaA	This study
ΔfsaB+++	DfsaB + pSEVA131-fsaB	This study
ΔgldA+++	DgldA + pSEVA131-gldA	This study
ΔglpK+++	DglpK + pSEVA131-glpK	This study
ΔdhaKLM+++	DdhaKLM + pSEVA131-dhaKLM	This study
Plasmids		
pSEVA131	Medium-copy-number, <i>lacIq</i> / <i>P_{trc}</i> promoter, pBBR1 ori, AmpR. Original pSEVA131 plasmid does not contain a promoter.	(Muller et al., 2015)
pSEVA234	Medium-copy-number, <i>lacIq</i> / <i>P_{trc}</i> promoter, pBBR1 ori, KmR.	(Silva-Rocha et al., 2013)
pSEVA131-dhaKLM	Derivative of pSEVA-131 containing dhaKLM operon. Used to overexpress dhaKLM in BW25113	This study
pSEVA131-fsaB	Derivative of pSEVA-131 containing fsaB gene. Used to overexpress fsaB in BW25113	This study
pSEVA131-gldA	Derivative of pSEVA-131 containing gldA gene. Used to overexpress gldA in BW25113	This study
pSEVA131-glpK	Derivative of pSEVA-131 containing glpK gene. Used to overexpress glpK in BW25113	This study
pSEVA234-fsaA	Derivative of pSEVA-234 containing fsaA gene. Used to overexpress fsaA in BW25113	This study
Primers		
dhaKLM_knockout_F	CGTGTTCGTTGAACATCATCCATGCCCTACCGTAATTGCTGG AGCAAATAGTGTAGGCTGGAGCTGCTTC	
dhaKLM_knockout_R	CATCAGAACGATGCCATCCGAACAGTGGCTTAACCCTGAC GGTTGAAACGCATATGAATATCCTCCTTAG	
glpK_knockout_F	TCCTTCAGAACAAAAGCTTCGCTGTAATATGACTACGGG ACAATTAA	
glpK_knockout_R	ACGTGTAGGCTGGAGCTGCTTC ACGTTTCGGGACTACGGATGCGGCATAAACGCTTCATTC GGCATTACACATATGAATATCCTCCTTAG	
Cm_F	AATCGTCGTGGTATTCACTCC	

Table 1: Bacterial strains, plasmids, and primers used in this study

Growth conditions

E. coli was cultivated overnight at 37°C with agitation at 220 rpm in LB broth (10 g tryptone, 5 g yeast extract, 10 g NaCl per liter) with appropriate antibiotics (100 µg/mL ampicillin or 50 µg/mL kanamycin) and 0.25 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) if needed. The optical density at 600 nm (OD_{600nm}) was measured by spectrophotometry and 5 mL samples of preculture were centrifuged at 8,000g for 5 minutes. The pellets were resuspended in M9 medium (3.48 g/L Na₂HPO₄•12H₂O, 0.606 g/L KH₂PO₄, 0.102 g/L NaCl, 0.408 g/L NH₄Cl, 0.49 g/L MgSO₄, 4.38 mg/L CaCl₂, 15 mg/L Na₂EDTA•2H₂O, 4.5 mg/L ZnSO₄•7H₂O, 0.3 mg/L CoCl₂•6H₂O, 1 mg/L MnCl₂•4H₂O, 1 mg/L H₃BO₃, 0.4 mg/L Na₂MoO₄•2H₂O, 3 mg/L FeSO₄•7H₂O, 0.3 mg/L CuSO₄•5H₂O, 0.1 g/L thiamine) with no carbon source.

The cultures were incubated at 37°C, 220 rpm in 250 ml baffled shake flasks containing 50 mL of M9 medium supplemented with DHA (Merck, 99.9% purity) at a final concentration of 15 mM. Cells were inoculated at an initial OD_{600nm} of 0.05 and growth was analysed every 2 h for 48 h using a Thermo Genesys6 spectrophotometer (Thermo Scientific). 750 µL of medium was collected and centrifuged at 8,000g for 3 minutes. 250 µL of supernatant was stored at -20°C before NMR or HPLC analysis.

Sampling, RNA extraction and microarray procedures

Cells were grown in duplicate in the same conditions as described above. When an OD_{600nm} around 0.5 was reached (See Figure S3), the cells were harvested and immediately frozen in liquid nitrogen. At such OD cells are in exponential growth phase and there is DHA and formate present in the medium for all of them (See Figure S3). For the WT strain, as the sampling time occurs during the night, the cultivation was diluted the next day in a fresh modified M9 medium @ OD=0.2 and sampling was done at OD-600 nm of around 0.5. Total RNA was extracted following the Qiagen RNeasy MiniKit procedure and quantified using a Nanodrop® spectrophotometer. Double-stranded complementary DNA (cDNA) synthesis and array processing were performed using the Agilent Technologies One-Color Microarray-Based gene Expression Analysis protocol. The images were analysed with the software DEVA (v1.2.1). All array procedures were performed using the GeT-Biopuces platform (<http://get.genotoul.fr/>). For the wild-type strain, gene expression of batch cultures with DHA was expressed relative to the ratio of gene expression of chemostatic growth with glucose at low ($\mu = 0.1 \text{ h}^{-1}$) and

high ($\mu = 0.6 \text{ h}^{-1}$) growth rates (data from reference (Esquerre et al., 2015)). This allowed to focus only on gene whose level of expression change in relation to the nature of the substrate while avoiding to highlight genes whose expression change because of the very different growth rate. For the $\Delta dhaKLM$ and $\Delta ptsa::kan$ mutants, gene expression on DHA was compared to the wild type's gene expression on DHA. Gene ontology analyses were performed using Ecocyc (Keseler et al., 2017). Gene expression data have been deposited in the ArrayExpress database at EMBL-EBI (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-7666.

NMR and HPLC analysis

NMR analysis were performed on an Avance III 800 MHz spectrometer (Bruker, Rheinstetten, Germany) equipped with a 5 mm QPCI cryogenic probe head at 280 K. Supernatants were analysed by quantitative ^1H 1D-NMR at 280 K using a zgpr30 sequence with water presaturation. A total of 32 scans were accumulated after 8 dummy scans. The time domain function (the FID) was converted to the frequency domain function (the spectrum) by Fourier transform. The phase of the spectra was adjusted manually, the baseline was corrected automatically and the spectra were aligned using the signal from 3-trimethylsilylpropionic-2,2,3,3-d₄ acid sodium salt (TSP-d₄), an internal standard, with the Bruker software TopSpin (v3.5). Propane-1,2,2,3-tetrol was identified using 2D heteronuclear single quantum correlation spectroscopy (HSQC) and heteronuclear multiple bond correlation spectroscopy (HMBC) spectra. The HSQC experiment was acquired with 2K x 1024 points for a spectral width of 13.95 x 140 ppm in the ^1H and ^{13}C dimensions, respectively. It was processed with 4k x 1024 points. The HMBC experiment was acquired with 2K x 1024 points, for a spectral width of 13.95 x 220 ppm in the ^1H and ^{13}C dimensions, respectively. It was processed with 8K x 1024 points. The HSQC and HMBC data were acquired with 16 scans per increment. Both 2D spectra were calibrated in the frequency domain by setting the peak from TSP-d₄ to 0 ppm in both dimensions. The spectra were processed using Topspin 3.5.

For transcriptomic analysis, HPLC analysis were performed using a column made of an H^+ chromatography resin (Zorbax-C18 from Agilent Technologies). A solution of sulfuric acid (5% v/v) was used as the eluent, at a flow rate of 0.5 mL min^{-1} and a volume of injection of 20 μL . The oven temperature was set at 48°C. Ranges of standards of glycerol, formate, and DHA

were realised in order to quantify extracellular metabolites. Retention times in min for each compounds were the following: Glycerol 16.6 (refractometry detection); DHA 16.3 (UV detection) and Formate 16.9 (UV detection). Calibration curves of these three compounds were established and used to calculate their concentrations in the culture supernatants.

Calculation of growth, substrate uptake and degradation rates

We developed a mathematical model to infer quantitative growth and exchange flux information from the measured time-dependent concentrations of biomass and extracellular metabolites. The general model, which accounts for the non-enzymatic degradation of substrates or products, relates changes of concentrations to fluxes using the following system of ordinary differential equations,

$$\frac{dX}{dt} = \mu \cdot X \quad (\text{eq. 1})$$

$$\frac{dM_i}{dt} = -k \cdot M_i + X \cdot q_{M_i} \quad (\text{eq. 2})$$

where X is the biomass concentration ($\text{g}_{\text{DW}} \text{L}^{-1}$), μ is the growth rate (h^{-1}), and M_i is the concentration of exchanged metabolite i (mmol L^{-1}) with a degradation constant k (h^{-1}) and exchange flux q_{M_i} ($\text{mmol g}_{\text{DW}}^{-1} \text{h}^{-1}$). Integrating equations 1-2 provides the following analytical functions:

$$X(t) = X_0 \cdot e^{\mu \cdot t} \quad (\text{eq. 3})$$

$$M_i(t) = q_{M_i} \cdot \frac{X_0}{\mu + k} \cdot (e^{\mu \cdot t} - e^{-k \cdot t}) + M_i^0 \cdot e^{-k \cdot t} \quad (\text{eq. 4})$$

This formalism has been implemented in PhysioFit, a flexible R program that allows growth rates and exchange fluxes to be quantified by fitting time-variations of extracellular metabolite and biomass concentrations using the *nlsic* algorithm (Sokol et al., 2012). PhysioFit includes options to account (or not) for the degradation of extracellular compounds and a lag phase before cells start to grow, and implements sensitivity analyses to evaluate the precision of the estimated fluxes. PhysioFit is provided open source at <https://github.com/MetaSys-LISBP/PhysioFit>.

A conversion factor was used to obtain cellular dry weights (DWs) from OD_{600nm} ($0.37 \text{ g}_{DW}/OD_{600nm}$) and calculate specific DHA uptake rates ($\text{mmol g}_{DW}^{-1}\text{h}^{-1}$), formate production yields ($\text{mmol g}_{DW}^{-1}\text{h}^{-1}$) and biomass yields (g mol^{-1}). Experiments were triplicated to calculate averages and standard deviations.

In silico analysis of DHA metabolism

In silico analyses of DHA metabolism were carried out with the *E. coli* genome scale metabolic model iJO1366 ((Orth et al., 2011), Biomedels identifier MODEL1108160000) constrained with experimental uptake fluxes of the wild-type strain (DHA = $5.2 \text{ mmol g}_{DW}^{-1} \text{ h}^{-1}$, formate = $3.2 \text{ mmol g}_{DW}^{-1} \text{ h}^{-1}$, glycolate = $1.0 \text{ mmol g}_{DW}^{-1} \text{ h}^{-1}$, and acetate = $0.1 \text{ mmol g}_{DW}^{-1} \text{ h}^{-1}$), using Sybil package (v2.1.2, (Gelius-Dietrich et al., 2013)) of R environment (v3.2.4, (R Development Core Team, 2009; Team, 2015)). The model and all the scripts used to run calculations are available at https://github.com/MetaSys-LISBP/Peiro_2019 under GPLv3 open source license to ensure reproducibility and reusability.

Results and discussion

DHA metabolism in *E. coli* involves a complex network of chemical and biological processes

We carried out cultivation experiments to study the global behaviour of *E. coli* during growth on DHA, monitoring the concentrations of biomass and extracellular metabolites as a function of time. DHA is an unstable compound that can be interconverted into different forms when dissolved in water (Yaylayan et al., 1999) or autoxidized by Fenton's reaction to form glycolate, other short-chained carbohydrates and organic acids upon incubation (Lip et al., 2013; Maksimović et al., 2006). We therefore determined beforehand whether non-enzymatic conversions of DHA could occur under our conditions. Non-inoculated M9 minimal medium containing 15 mM DHA was incubated at 37°C for 48 h, and samples were collected after 0, 24 and 48 h of incubation and analysed by ¹H-NMR (Figure 1). Since DHA is a symmetrical molecule and all the detectable protons from the two CH₂ moieties are equivalent, the ¹H-NMR spectrum of pure DHA is expected to contain a singlet at 4.4 ppm. However, several signals were detected in the medium at time 0 (Figure 1A). None of these peaks were observed in a sample containing M9 medium without DHA (Figure 1A) suggesting that they originated from the DHA solution. 2D-NMR analysis (Figure S1) shows that the signal at 3.6 ppm corresponds to a hydrated form of DHA, propane-1,2,2,3-tetrol (or dihydroxyacetone monomer hydrate), as previously reported (Yaylayan et al., 1999). Under our conditions, the proportions of non-hydrated and hydrated DHA were 65 ± 2% and 21 ± 2 %, respectively. The remaining 14 ± 2% corresponds to the signals at 4.8, 4.7 (both singlets) and 3.7 (AB system) ppm, all originating from the same unidentified molecule X, which appeared after freezing/thawing the samples. Unfortunately, we were unable to identify the compound corresponding to these peaks.

Upon incubation of the non-inoculated medium, the peaks from DHA, propane-1,2,2,3-tetrol and molecule X decreased in intensity while additional signals appeared compared with the initial spectrum (Figure 1B). This phenomenon was influenced by the amount of salts (i.e. Na₂HPO₄, KH₂PO₄, NaCl and NH₄Cl) in the medium. At the salt concentration typically used for the M9 medium (Heux et al., 2011), 82 % of DHA was degraded after 48 h while only 34 % was degraded when Na₂HPO₄, KH₂PO₄, NaCl and NH₄Cl were diluted five times (data not shown). A 5X reduced salt concentration was thus used for all the subsequent experiments. In this modified M9 medium, DHA was mainly converted into formate, glycolate and acetate (Figure

1B and Figure S2). After 39 h of incubation, 27% of the DHA was converted into formate (36%), glycolate (26%) and acetate (1.3%). The nonlinear DHA concentration profile was fitted assuming first order kinetics. The excellent agreement between experimental and fitted curves (Pearson's $r > 0.99$) validates the proposed degradation kinetics; the estimated degradation rate constant is $0.0086 \pm 0.0010 \text{ h}^{-1}$, corresponding to a DHA half-life of 116 h (Figure S2).

In the inoculated medium, the peaks from both forms of DHA decreased monotonously over time (data not shown). The proportions of both forms remained constant during the experiment, suggesting that chemical equilibrium is faster than DHA uptake. No accumulation of glycolate, formate or acetate was observed, indicating that these products of non-enzymatic DHA degradation are co-consumed with DHA by wild-type *E. coli*. In order to infer quantitative flux information from these data, we developed a R program, PhysioFit, that is a mathematical model to estimate the growth rate and exchange (uptake or production) fluxes. This model, which takes into account non-enzymatic degradation of substrates, is described in detail in the *Methods* section. This program calculated a growth rate of $0.15 \pm 0.03 \text{ h}^{-1}$ with a DHA uptake rate of $5.2 \pm 0.19 \text{ mmol g}_{\text{DW}}^{-1} \text{ h}^{-1}$ (corresponding to $15.6 \text{ Cmmol g}_{\text{DW}}^{-1} \text{ h}^{-1}$). Based on the non-enzymatic conversion yields of DHA, we estimated that DHA accounted for 74.3 % of *E. coli* carbon uptake (expressed in Cmmol), with the remaining carbon taken up as formate (15.3%), glycolate (9.7%) and acetate (0.7%). This represent a specific uptake rate of respectively 3.2, 1.0 and 0.1 $\text{mmol g}_{\text{DW}}^{-1} \text{ h}^{-1}$ (i.e. 3.2, 2.0 and 0.2 $\text{Cmmol g}_{\text{DW}}^{-1} \text{ h}^{-1}$).

Overall, these results show that an additional set of reactions (i.e. non-enzymatic conversion of DHA) occur under these condition (Figure 2) which have to be taken into account when studying the metabolism of DHA.

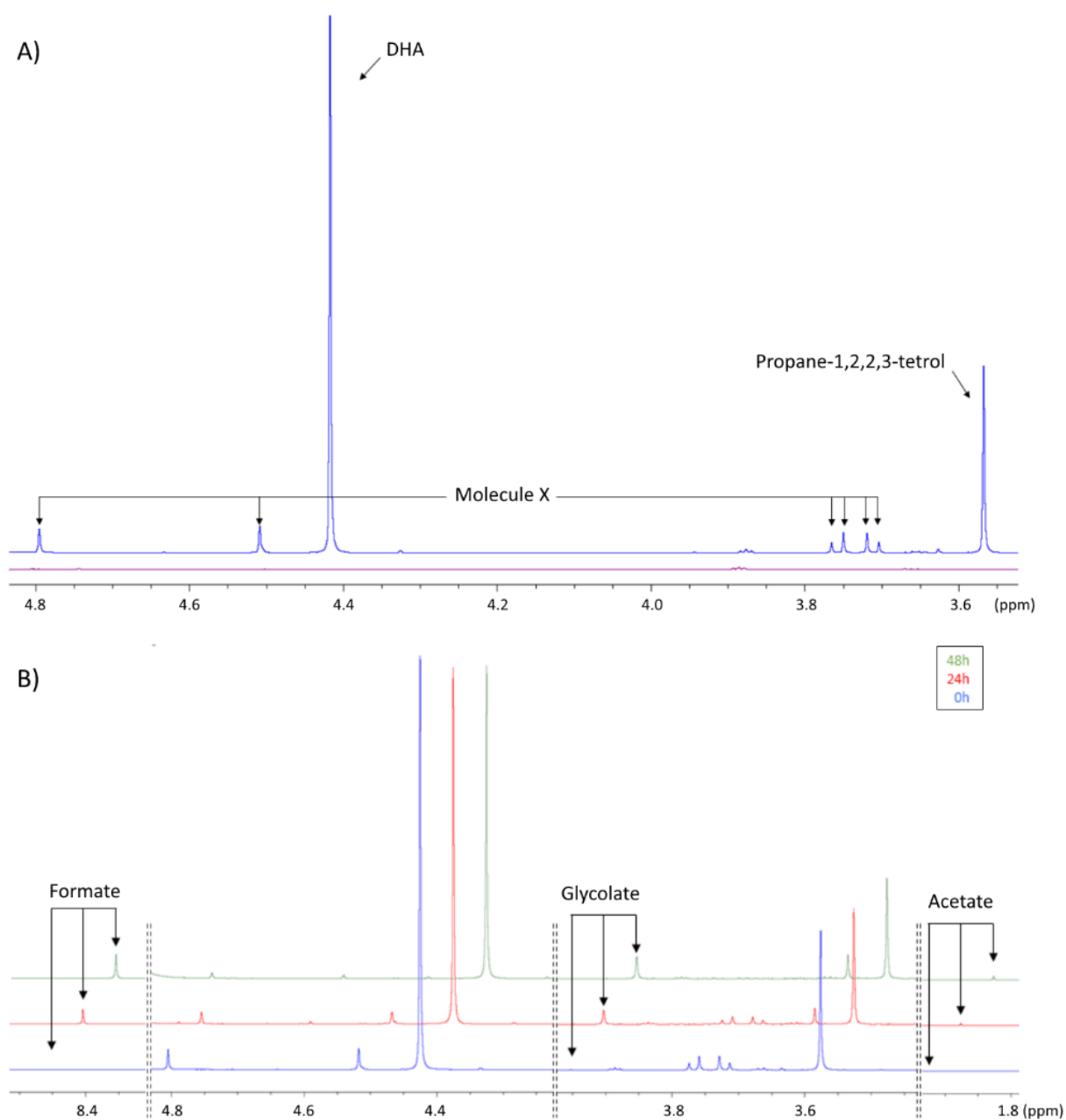


Figure 1: Non-enzymatic transformation of DHA in M9 medium with 5X diluted salts. (A) NMR spectra overlay of non incubated modified M9 medium with (blue) or without (purple) 15 mM of DHA; (B) kinetics of non enzymatic transformation of DHA in modified M9 medim after 0 h (blue), 24 h (red) and 48 h (green) of incubation at 37°C. The three spectra are shown with a vertical step of 10% and a horizontal offstep of 0.05 ppm.

***E. coli* growth on DHA is robust but suboptimal**

Constraint-based metabolic models provide an attractive starting point for studying poorly characterised metabolisms because they can predict growth rates and metabolite flows through a metabolic network with minimal *a priori* knowledge. We performed flux balance analysis (FBA) using a genome scale *E. coli* model (iJO1366, (Orth et al., 2011)) to simulate the fate of DHA (Figure 2). The only constraints used were the DHA, formate, glycolate and acetate uptake rates of wild type *E. coli*, calculated as detailed above from experiment. The model predicted a growth rate of 0.28 h^{-1} , i.e. almost 2 times higher than the experimentally observed growth rate (0.15 h^{-1}). Since FBA predictions assume optimal growth under stoichiometric and maximum uptake constraints, these results suggest that for some reason the system was not operating optimally *in vivo*. For optimal growth, the predicted model has 63% of the DHA converted into DHAP via the DAK based pathway and 36 % via the FSA pathway. The GLD based pathway is only used to feed glycerol-3-phosphate for biomass synthesis (Figure 2). Flux balance analysis computes an optimal objective value and a flux state that are consistent with that objective (and all the imposed constraints). While the objective value (i.e. the growth rate under our conditions) is unique, this can typically be supported by multiple flux states in genome-scale models. For this reason, we performed flux variability analysis (FVA) to find the minimum and maximum flux for each reaction in the network while maintaining growth at 95% of its optimal value (Mahadevan, 2003). The results indicate that the fluxes through the GLD, DAK and FSA based pathways offer a wide range of possible values, including opposite fluxes (i.e. reverse reactions). These data suggest that, based solely on stoichiometric constraints, optimal growth does not in theory depend on any particular DHA metabolic pathway. Overall, these data demonstrate flexible use of all the DHA catabolic pathways and suggest that this metabolism in *E. coli* is highly robust.

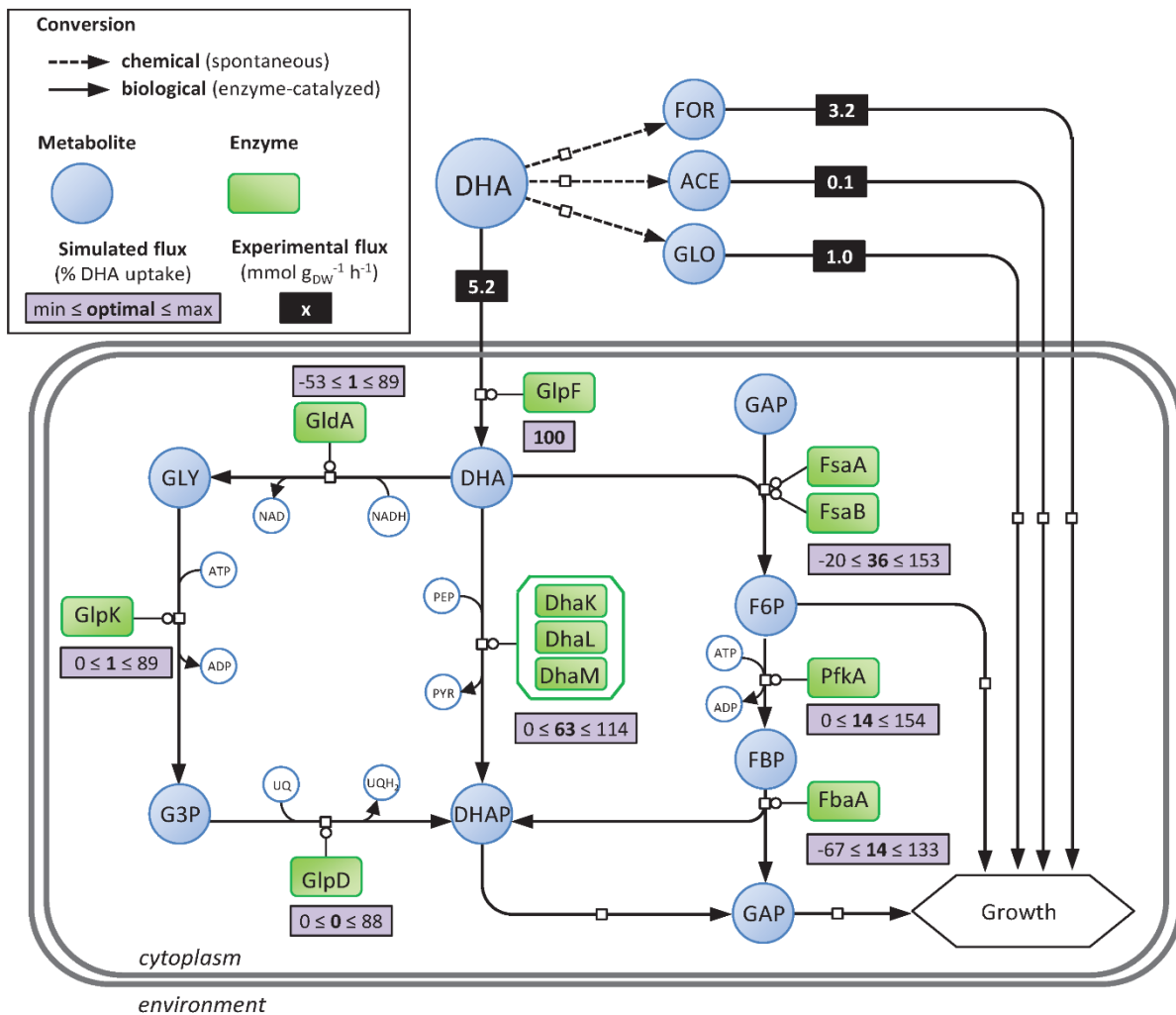


Figure 2: Experimental and simulated fluxes through DHA metabolism of *E. coli*.

Grey rectangles give the optimal and the ranges of simulated fluxes. Optimal simulated fluxes were obtained using flux balance analysis constrained with experimental uptake fluxes of the wild-type strain grown on modified M9 media containing 15 mM of DHA (i.e. values given in black rectangle). Ranges of simulated fluxes were obtained using flux variability analysis when growth rate is constrained to 95% of the optimal value and with experimental uptake fluxes. Experimental flux values are given in mmol g_{DW}⁻¹ h⁻¹ and simulated flux values are given in % relative to DHA uptake rate.

Dihydroxyacetone kinase enzyme (DhaK, DhaL and DhaM), fructose-6-phosphate aldolase enzymes (FsaA and FsaB), glycerol dehydrogenase enzyme (GldA), glycerol-3-phosphate dehydrogenase enzyme (GlpD), glycerol kinase enzyme (GlpK), fructose-bisphosphate aldolase enzyme (FbaA), 6-phosphofructokinase enzyme (PfkA), dihydroxyacetone (DHA), formate (FOR), acetate (ACE), glycolate (GLO), dihydroxyacetone phosphate (DHAP), fructose-6-phosphate (F6P), glyceraldehyde-3-phosphate (GAP), glycerol (GLY), glycerol-3-phosphate (G3P), phosphoenolpyruvate (PEP), pyruvate (PYR), ubiquinone (UQ), ubiquinol (UQH₂).

DHA-induced genes involved in a biofilm growth state

In order to gain insight into the regulation of DHA metabolism, we compared the global transcriptional responses of wild-type *E. coli* grown on DHA (Figure S3) and glucose medium. This revealed that several genes related to mobility, adherence, and biofilm formation and stress were expressed differently between DHA and glucose growth (Table 2 and Data File S1). For instance, the *csg* genes encoding for curli fibers, which promote cell adhesion during biofilm formation (Barnhart and Chapman, 2006), were expressed at a higher level (by a factor 10 on average) during DHA growth. Conversely, genes involved in motility and flagella assembly were less expressed (e.g. by factors of 25 and 12 fold on average for the *flg* and *fli* genes, respectively) (Table 2 and Data File S1), indicating that motility might be repressed on DHA. Genes related to lipopolysaccharide (LPS) synthesis and were expressed by a factor 4 on DHA (Table 2 and Data File S1). LPSs are a major component of the outer leaflet of the outer membrane of most Gram-negative bacteria, contributing to their structural integrity. LPSs govern many of the biological interactions between cells and their environment (Zhang et al., 2013) and play a crucial role in the anchoring process (Jacques, 1996). Taken together, these data suggest that *E. coli* switches from planktonic growth on glucose to biofilm growth on DHA. This might contribute for the low growth rate of *E. coli* on DHA since cells in biofilms are known to grow more slowly than in the planktonic mode (Jefferson, 2004).

Over-expressed	GO concerned	P-values	Gene examples
Cell wall and adhesion	GO:0009103 - lipopolysaccharide biosynthetic process	2.66E-14	<i>waao, wbbI, arrn, wcaB, waab, waal, waaY, waaU, rfbA, rfbC</i>
	GO:0022610 - biological adhesion	8.85E-12	<i>dgcz, eljA, yadK, yehA, pgaC, yeeI, csqB, csqF, csqE, csqG</i>
Response to host	GO:0043711 - pilus organization	4.79E-10	<i>htrE, yehB, ybgQ, elfC, sfmtD, ydeT, yqiG, fimD, sfmtC, elfD</i>
	GO:0006952 - defense response	4.50E-07	<i>casE, casD, casC, casB, casA, abpA, abpB, rzpD, rrrQ, rzpR</i>
Response to acidic pH	GO:0009243 - O antigen biosynthetic process	4.86E-05	<i>rfbC, rfbD, rfbB, rfbA, wbbI, rfbX</i>
	GO:0010447 - response to acidic pH	6.07E-05	<i>cadB, yjaA, hdeA, iraM, oxc, frc, glsA, yagU, evgS, hdeB</i>
Copper ion homeostasis Genes involved in the glucarate catabolic process	GO:0006878 - cellular copper ion homeostasis	7.59E-04	<i>cusA, cusB, cusC, cusF</i>
	GO:0019394 - glucarate catabolic process	1.67E-04	<i>garKLR, gudD</i>
Under-expressed	GO concerned	P-values	Gene examples
Locomotion	GO:0040011 - locomotion	1.28E-24	<i>filI, flgB, flgC, flgG, flgD, flgE, flgF, motA, fliA, flig</i>
	GO:0045333 - cellular respiration	6.09E-15	<i>frdB, hyaA, hyaB, pfID, fdoG, fdoH, frdA, frdC, frdD, hybB</i>
Cellular respiration	GO:0019321 - pentose metabolic process	2.11E-10	<i>rbsD, rbsK, xylA, araD, araA, xylF, xylG, xylH, rhaD</i>
	GO:0016052 - carbohydrate catabolic process	3.19E-9	<i>treC, lacZ, melA, mals, glcA, hybA, gatY, uidA, gntK, gntP, gntT</i>
Primary metabolism	GO:0008643 - carbohydrate transport	4.84E-09	<i>kdgT, gntP, gntT, ptsG, yjaO, melB, rbsA, srlB, frwC, treB</i>
	GO:0071941 - nitrogen cycle metabolic process	2.45E-08	<i>argG, argI, argF, carA, ygeW, napC, nirB, narG, nirD, nirC</i>
	GO:0006099 - tricarboxylic acid cycle	5.99E-07	<i>sdhD, sdhC, sdhB, acnA, sucA, fumB, mdh, sdhA, sucD, sucC</i>

Table 2: Functional classification of genes with statistically significant decreases and increases in mRNA level in BW25113 *E. coli* strain in M9-DHA medium compared to M9-glucose medium (Esquerre et al., 2015).

The categories of orthologous genes (COG) were used for grouping.

The DAK pathway is central for DHA metabolism but not essential

The regulation of *dhaKLM* expression (Bächler et al., 2005) and the properties of the enzyme (Garcia-Alles et al., 2004) have led to the suggestion that the DAK pathway is important and may even be essential for growth on DHA in *E. coli* (Paulsen et al., 2000). To test this hypothesis, we investigated the expression levels of the genes involved in DHA metabolism in wild-type and $\Delta dhaKLM$ strains grown on DHA (Figure 3 and Figure S3) and studied the functional characteristics of DhaKLM (Figure 4). Comparing the expression levels of wild-type strains grown on DHA vs glucose shows that the genes involved in glucarate metabolism (i.e. the *gar* and *gud* operons) are upregulated by about a factor of 5 on average (Figure 3 and Data File S1) while those involved in glycolate metabolism (i.e. the *glc* operon) were not. As shown previously, DHA is spontaneously converted into glycolate, which is supposed to induce the *glc* genes (Pellicer et al., 1999) however glycolate can also be recognized by the same transporter as glucarate (Robertson et al., 1980) and catabolised using part of its metabolic pathway. The gene encoding the glycerol facilitator GlpF, a nonspecific channel protein capable of transporting straight-chain carbon compounds such as DHA, was upregulated about 3-fold, consistent with its putative role in DHA transport (Paulsen et al., 2000). As expected, both the *dhaKLM* operon and its transcription factor *dhaR* were strongly upregulated (Figure 3 and Data File S1). However, the expression levels of *glpK*, *fsaA* and *fsaB* were not affected, and *gldA* was downregulated by a factor of 2.5 (Figure 3). Transcriptome analysis of the $\Delta dhaKLM$ and wild-type strains (Figure 3) revealed that none of the genes of the FSA and GLD pathways were upregulated in $\Delta dhaKLM$, demonstrating that there is no compensatory activation of these pathways. Interestingly, the *hyc* and *hyp* operon and Fhl encoded enzymes, involved in the production of dihydrogen and carbon dioxide in formate metabolism (McDowall et al., 2014) and activated by formate, were up-regulated by a factor of 10 (Figure 3 and Data File S1). This points to the presence of formate in the medium, which is consistent with the non-enzymatic degradation of DHA into formate and suggests that Fhl encoded enzymes allow the $\Delta dhaKLM$ strain to convert formate from the medium and use the hydrogen as an electron donor for respiration (Unden et al., 2014). Deletion of *dhaKLM* resulted in a significant and strong reduction of the growth rate and a shift in the substrate profiles compared with the control strain (Figure S3 and 4A) while complementation of the deleted genes restored normal growth (Figure S4). In this strain furthermore, the specific DHA uptake rate was strongly reduced. When *dhaKLM* was overexpressed (i.e. *dhaKLM+++*), the

specific DHA uptake rate increased slightly, by 15% compared with the wild type, but the growth rate did not (Figure 4). Taken together, these data support the hypothesis that DhaKLM plays a crucial role in DHA utilisation, in agreement with the transcriptomics data. However, under our conditions, and contrary to previous observations in *E. coli* (Paulsen et al., 2000), inactivation of *dhaKLM* was not lethal. This is because the metabolism of DHA is shifted to formate, glycolate and acetate metabolism (due to the instability of DHA) in this strain. Based on the non-enzymatic conversion yields of DHA into formate, glycolate and acetate determined above, we estimate that formate accounted for 38.5% of the $\Delta dhaKLM$ strain's carbon uptake (expressed in Cmmol), while DHA, glycolate and acetate accounted for 34, 25.7 and 1.6%, respectively.

Overall these data show that *dhaKLM* is the predominantly expressed operon on DHA, while the genes encoding for the GLD and FSA pathways are not transcriptionally activated, even to rescue the functionality of a strain deleted for *dhaKLM*. Instead, this strain survives on DHA's degradation products, leaving the alternative catabolic pathways dormant.

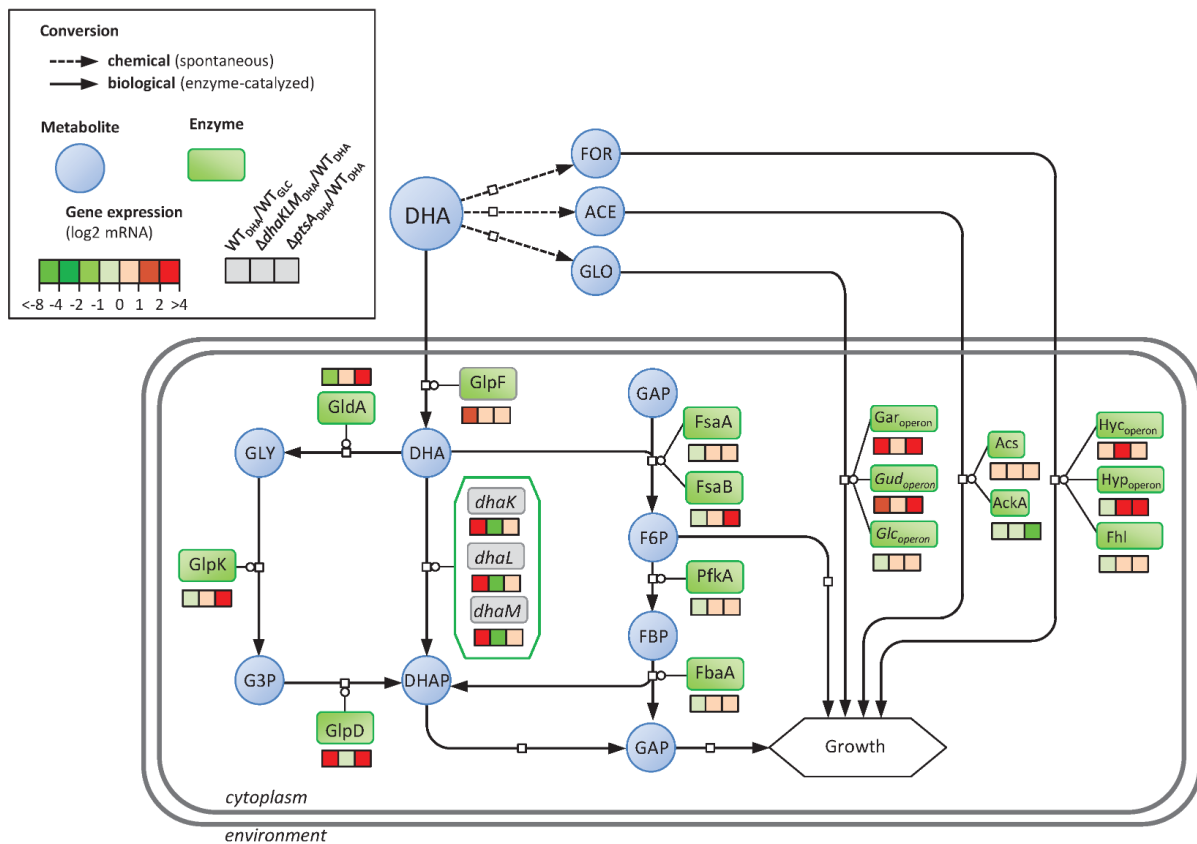


Figure 3: Change in gene expression in DHA metabolism of wild type strain, in $\Delta dhaKLM$ strain and $\Delta ptsA$ strain cultured on M9-DHA medium.

For the wild type strain (WT_{DHA}), the fold changes in gene expression was calculated in reference to the expression in an *E. coli* strain cultured in chemostat in M9-glucose medium at 0.1 h⁻¹ (WT_{GLC}). For the $\Delta dhaKLM$ ($\Delta dhaKLM_{DHA}$) and the $\Delta ptsA$ ($\Delta ptsA_{DHA}$) strains the fold changes in gene expression were calculated in reference to expression in an *E. coli* strain cultured in M9-DHA medium. The *ptsA* gene in the $\Delta ptsA$ mutant is replaced by a kanamycin cassette leading to the overexpression of *gldA* and *fsaB* which are part of the same operon. For all the experiments, n = 2 biological replicates and value are given in Log₂.

Dihydroxyacetone kinase gene (*dhaK*, *dhaL* and *dhaM*), fructose-6-phosphate aldolase genes (*fsaA* and *fsaB*), glycerol dehydrogenase gene (*gldA*), glycerol-3-phosphate dehydrogenase gene (*glpD*), glycerol kinase gene (*glpK*), fructose-bisphosphate aldolase gene (*fbaA*), 6-phosphofruktokinase gene (*pfkA*), glucarate operon (*gar*_{operon} and *gud*_{operon}), acetyl-CoA synthetase gene (*acs*), acetate kinase (*ackA*), formate hydrogenlyase system (*hyc*_{operon}, *hyp*_{operon} and *fhl*), enzyme I (*ptsI*), dihydroxyacetone (DHA), formate (FOR), acetate (ACE), glycolate (GLO), dihydroxyacetone phosphate (DHAP), fructose-6-phosphate (F6P), glyceraldehyde-3-phosphate (GAP), glycerol (GLY), glycerol-3-phosphate (G3P).

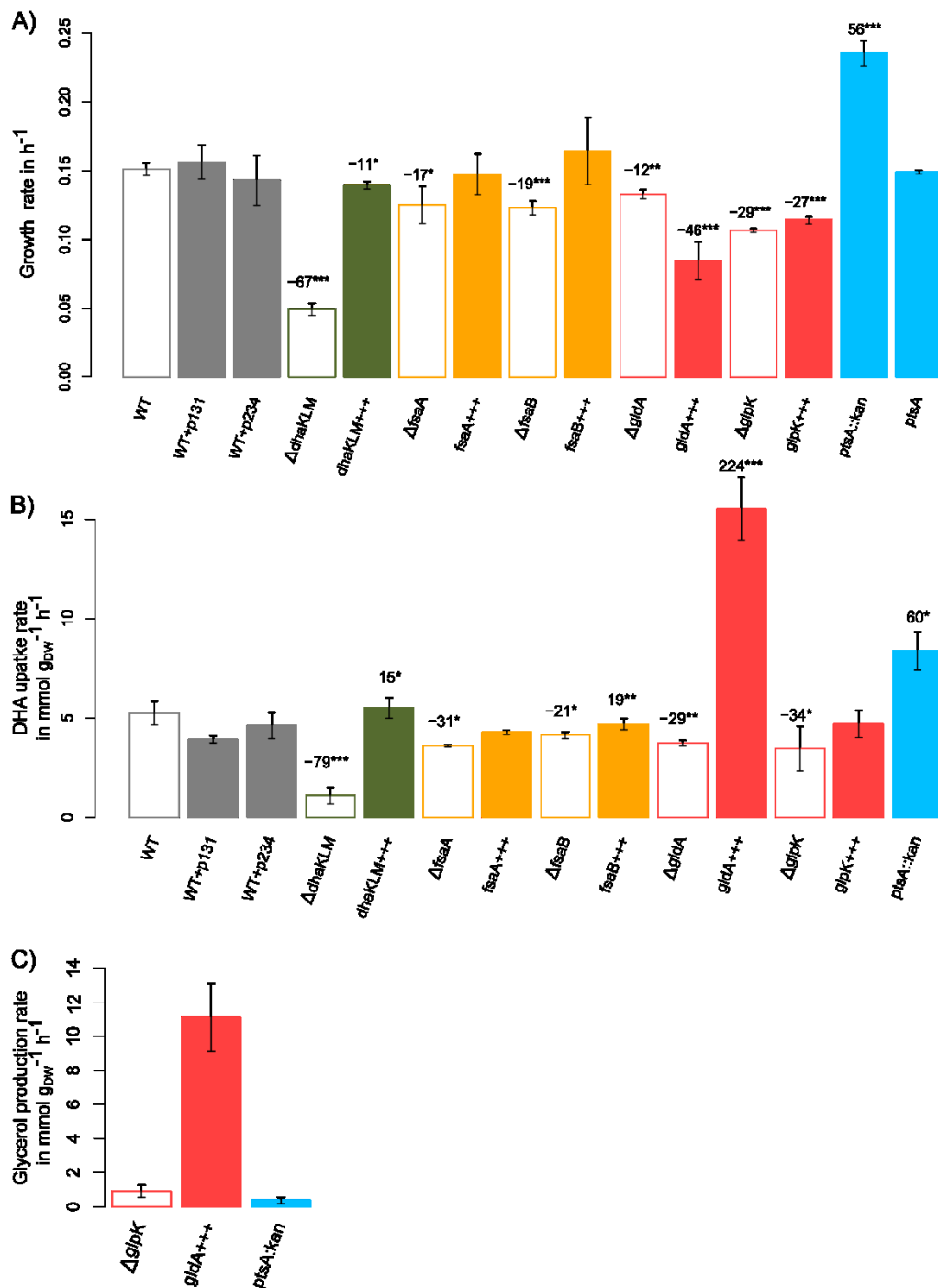


Figure 4: Growth rate (A), specific DHA uptake rate (B) and specific glycerol production rate (C) of the strains deleted (Δ) and overexpressing (+++) genes involved in DHA metabolism (Table 1).

Cells were grown on modified M9 medium with 15 mM of DHA at 37°C and 250 rpm shaking. Growth rate is given in h^{-1} . DHA uptake rate and glycerol production rate are given in $mmol\ g_{DW}^{-1}\ h^{-1}$. Physiofit was used to estimate growth and exchanges rates. Data shown represent mean and standard deviation of mean ($n=3$). P values were obtained using a t-test comparing mutants strains with the control WT (***) $p < 0.001$, (**) $p < 0.01$, (*) $p < 0.05$). Please refer to table 1 for details on each strain.

The GLD and FSA pathways are functionally involved in DHA metabolism

To assess whether the GLD and FSA pathways are functionally involved in DHA metabolism even if their transcription is not activated on DHA, we performed functional analysis of the genes involved in these pathways. The results obtained for all the deleted or overexpressing strains (Table 1) grown on minimal medium containing DHA are shown in Figure 4. Inactivation of *gldA*, *glpK*, *fsaA* and *fsaB*, resulted in significantly slower growth compared with the control strain (Figure 4A) and complementation of the deleted genes restored the growth rate (Figure S4). After $\Delta dhAKLM$, the gene deletion that had the greatest impact on the growth rate and the specific DHA uptake rate was $\Delta glpK$, reducing them by 29 and 34%, respectively (Figure 4A&B). In *H. volcanii* (Ouellette et al., 2013), GlpK has a central role in DHA metabolism because it can directly phosphorylate DHA to DHAP using ATP. If GlpK could act as an ATP-dependent DHA kinase in *E. coli*, its overexpression would improve the fitness of the cells. Indeed, the conversion of DHAP to pyruvate generates two ATPs and only one PEP, and a doubling of the level of phosphate donors (ATP molecules) in the cytoplasm should support higher ATP-dependent DHA kinase activity compared with PEP-dependent DHA kinase, and thus better growth. This has indeed been observed recently in an *E. coli* strain that overexpresses an ATP-dependent DHA kinase during anaerobic growth on DHA (Wang et al., 2018). However, this was not observed under our aerobic conditions – overexpressing *glpK* (i.e. *glpK+++*) decreased the growth rate (Figure 4A) –, suggesting that GlpK does not act *in vivo* as an ATP-dependent DHA kinase enzyme. In addition, the production of glycerol observed in the $\Delta glpK$ strain (Figure 4C) clearly indicates that GlpK is involved in glycerol phosphorylation rather than in DHA phosphorylation. Results obtained for the strain overexpressing *gldA* (i.e. *gldA+++*) also support this hypothesis. In this strain, glycerol was produced at almost the same rate as DHA was taken up (Figure 4C) and started to be consumed only when DHA was exhausted (Figure S5). This suggests that DHA may have an effect *in vivo* on glycerol utilization. In *E. coli*, the catalytic activity of GlpK is inhibited allosterically by enzyme IIA from the glucose PTS system (de Boer et al., 1986). It could also be inhibited allosterically by enzyme IIA from the DHA PTS system (i.e. DhaM) (Gutknecht et al., 2001). Deletion of *gldA* (i.e. $\Delta gldA$) reduced both the growth and the specific DHA uptake rate (Figure 4A & B), most probably because the production of glycerol-3-phosphate, which is essential for biomass synthesis, is provided by just one pathway (i.e. from DHAP via glycerol-3-phosphate dehydrogenase encoded by *gpsA*). Overexpressing *gldA* more than doubled the

specific DHA uptake rate (Figure 4B) which is consistent with its high affinity for DHA. However, the growth rate was not improved (Figure 4A) because of a bottleneck at the GlpK level possibly due to allosteric regulation by DHA.

Finally the deletions *fsaA* and *fsaB* (i.e. $\Delta fsaA$ and $\Delta fsaB$), reduced both the growth rate and the specific DHA uptake rate (Figure 4A&B). While their overexpression (i.e. *fsaA*+++ and *fsaB*+++) did not have a significant effect on the growth rate compared to the wild type, *fsaB* overexpression was associated with a significant increase in the DHA uptake rate. This demonstrates that FsaA and FsaB are involved in DHA utilisation and suggests that the concentration of FsaB could be limited in the wild type strain.

Overall, these results show that GLD and FSA are both functionally involved in DHA metabolism but that their action is restricted. A number of strands of evidence suggest that *fsaB* and *gldA* are both anaerobic genes whose transcription is activated under anaerobic conditions (Durnin et al., 2009; Reizer et al., 1995; Rolfe et al., 2012). This would explain why these genes were not highly expressed in our aerobic conditions and thus why these pathways were only weakly used for DHA assimilation.

Overexpressing the GLD and FSA pathways lead to optimal growth on DHA

We used a *ptsA* mutant from the KEIO collection (Baba et al., 2006) to test whether optimal DHA growth could be restored by overexpressing the FSA- and GLD-based pathways. The *ptsA* gene in this mutant is replaced by a kanamycin cassette leaving *gldA* and *fsaB* expression under the control of the kanamycin promoter. As a result, *gldA* and *fsaB* were overexpressed 28- and 123-fold, respectively, compared with the wild-type strain (Figure 2, Figure S3 and Data File S1). Interestingly, in the $\Delta ptsA::kan$ mutant, *glpK* was also up-regulated (by a factor of 10) making the GLD pathway fully activated. Furthermore, the $\Delta ptsA::kan$ strain grew twice as fast as the wild type, at close to the *in silico*-predicted optimal rate on DHA (Figure 4A). The specific DHA uptake rate was likewise increased by 60% (Figure 4B). Removing the kanamycin cassette of the $\Delta ptsA::kan$ strain (i.e. $\Delta ptsa$) decreased the growth rate to a value similar to the one measured for the wild-type strain on DHA (Figure 4A). This means that PtsA is not involved itself in the overexpression of genes encoding for the GLD and FSA pathways. Overall, these data are consistent with our previous results that overexpressing *fsaB* increases the DHA

uptake rate. However, glycerol was only slightly accumulated in the $\Delta ptsA::kan$ strain (Figure 4C) in contrast to what has been observed in the strain overexpressing *gldA*. This is consistent with the upregulation of *glpK* along with genes involved in *sn*-glycerol 3-phosphate (G3P) catabolism (i.e. *glpABC* and *glpQ*) and transport (*glpT*) (Table 3 and Data File S1) and thus, with the use of the GLD pathway. The remaining question is why *glpK* is upregulated in the $\Delta ptsA::kan$. This cannot only be due to the overexpression of *gldA*, since *glpK* was not upregulated in the wild type strain that only overexpressed *gldA*.

In the $\Delta ptsA::kan$ strain, genes coding for the DAK pathway were not up-regulated while genes encoding the gluconate pathway (used for glycolate assimilation) were highly expressed (7-fold on average) (Figure 2 and Data File S1). Surprisingly, several genes associated with anaerobic or O₂-limited conditions were up-regulated (Table 3 and Data File S1). Several operons encoding enzymes in anaerobic respiratory chains (Tseng et al., 1996; Uden et al., 2014), i.e. (i) anaerobic dehydrogenases (GlpABC, FdnGHI, HyaABC, HybABC); (ii) anaerobic terminal reductases (NarG, NapABCGH, NirBD, FrdABCD, DmsABC) and (iii) micro-aerobic terminal oxidase (CydAB), were upregulated. All these dehydrogenases are theoretically able to transfer electrons to each of the terminal reductases or oxidases, provided the enzymes react with the same type of quinone. This may provide the $\Delta ptsA::kan$ strain with a large variety of respiratory chains (Uden et al., 2014) to fulfill its energy requirements. Secondly, genes encoding enzymes in the tricarboxylic acid cycle and involved in fatty acid degradation were downregulated along with genes involved in putrescine catabolism. These genes are known to be repressed under anaerobic conditions (Gunsalus and Park, 1994; Luchi and Lin, 1991; Partridge et al., 2006). However, no change in the expression of Fnr, the major regulator governing the physiological switch between aerobic and anaerobic growth conditions and controlling the expression of all these genes (Constantinidou et al., 2006; Kang et al., 2005), was observed. Altogether, these data suggest that the $\Delta ptsA::kan$ strain adopts anaerobic growth behavior in a fully aerated medium, taking advantage of all available sources of carbon to grow on DHA while meeting its energetic needs.

In sum, these results demonstrate that optimal growth on DHA can be achieved by releasing hierarchical constraints on DHA metabolism, opening additional routes for its assimilation. However, the regulation mechanisms involved in the activation of anaerobic genes in the $\Delta ptsA::kan$ strain remain to be elucidated.

Over-expressed	GO concerned	P-values	Gene involved
Secondary metabolism	GO:0019563 - glycerol catabolic process	3.82E-11	<i>glcA, hybA, glpK, glpC, glpB, glpD, glpA</i>
	GO:0045333 - cellular respiration	2.90E-19	<i>frdB, dmsA, dmsB, dmsC, hybB, fdnG, fdnH, frdA, frdC, frdD, hybB, hybO, napB, napC, narG, narH, nirB, nirD, napA, glpD, glpA, glpB, glpC, ndh, hybA, hybC, hyaC, yjiI, cydA, cydB</i>
	GO:0071941 - nitrogen cycle metabolic process	1.75E-09	<i>argF, napC, glnL, glnG, nirB, narG, nirD, nirC, narJ, narK, narH, napA, nifA</i>
	GO:0019394 - gluconate catabolic process	7.97E-06	<i>garL, guuD, garR, garK, ,</i>
	GO:0046688 - response to copper ion	1.14E-08	<i>cusS, copA, cusA, cusB, cusC, cusF, cusR, yobA, cueO</i>
Response to stimuli	GO:0009432 - SOS response	1.74E-07	<i>recA, recN, sulA, umuC, umuD, yebG, recX, dinI, lexA, dinD, dinG</i>
Transport	GO:0015886 - heme transport	2.41E-08	<i>ccmA, ccmD, ccmB, ccmC, ccmF, dppF, ccmE, dppC</i>
	GO:0015675 - nickel cation transport	2.05E-05	<i>nikB, nikC, nika, nikE, nikD, nika</i>
	GO:0006857 - oligopeptide transport	8.02E-04	<i>dppC, dppF, oppB, oppC, oppF, oppD, oppA</i>
Under-expressed	GO concerned	P-values	Gene involved
Iron	GO:0055072 - iron ion homeostasis	7.40E-15	<i>fepA, efeB, cirA, yqjH, fhuf, bfr, iscU, fhUA, fhUB, fhUC, fhUD, feoA, fepB, fes, fhUE, fecl, fecR, fepD, fepG, fepC, fju</i>
	GO:0016226 - iron-sulfur cluster assembly	2.05E-06	<i>iscS, iscU, erpA, sufB, sufD, sufA, sufC, sufE, sufS,</i>
Secondary metabolism	GO:0009447 - putrescine catabolic process	3.95E-10	<i>patA, patD, puuE, puuA, puuD, puuP, puuR, puuC, puuB,</i>
	GO:0009065 - glutamine family amino acid catabolic process	1.41E-05	<i>putA, astE, astD, astC, astB, astA, gabD, alda,</i>
	GO:0006790 - sulfur compound metabolic process	7.66E-08	<i>prpB, prpD, prpE, prpR, fadhH, fadI, fadE, fadM, fadD, fadA, fadB, yqeF</i>
	GO:0006099 - tricarboxylic acid cycle	4.17E-09	<i>sdhC, sdhB, acnA, gltA, fumaA, fumC, sdhA, aceB, aceA, aceK, sucA, prpC, prpD</i>
TCA	GO:0009062 - fatty acid catabolic process	4.62E-09	<i>prpB, prpD, prpE, prpR, fadhH, fadI, fadE, fadM, fadD, fadA, fadB, yqeF</i>
Fatty acids	GO:0009712 - catechol-containing compound metabolic process	8.28E-08	<i>entA, entC, entE, entH, entF, entB, fes, paal</i>
	GO:0070814 - hydrogen sulfide biosynthetic process	1.25E-07	<i>astE, astD, astC, astB, astA, gabD</i>
Hydrogen sulfide	GO:0042930 - enterobactin transport	3.10E-06	<i>fepA, fepD, fepG, fepC, fepB, entS</i>
Enterobactine			

Table 3: Functional classification of genes with statistically significant decreases and increases in mRNA level in $\Delta ptsA$ BW25113 *E. coli* strain in M9-DHA medium compared to BW25113 *E. coli* WT strain in M9-DHA medium.

The categories of orthologous genes (COG) were used for grouping.

Conclusion

This study shows that aerobic DHA metabolism in *E. coli* is far from optimal because of chemical, hierarchical and possibly allosteric constraints. However, our results show that removing hierarchical constraints optimizes growth on DHA. Beyond contributing to a better system-level understanding of DHA metabolism, these results are likely to accelerate the development of microbial biocatalysts. Enabling the rational design of strain that use DHA efficiently should improve biotechnological applications involving DHA as an intermediate, a well-known example being the bioconversion of glycerol into valued added chemicals (Volker F. Wendisch, 2011). This should also facilitate research aiming at constructing synthetic methylotrophs (i.e. engineering of non-native methylotrophs for methane and methanol-based production of chemicals) for which of DHA can be an intermediate (Heux et al., 2018; Wang et al., 2018).

References

- Ahmad, M., Hirz, M., Pichler, H., and Schwab, H. (2014). Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. *Appl Microbiol Biotechnol* *98*, 5301-5317.
- Akhtar, N., Blomberg, A., and Adler, L. (1997). Osmoregulation and protein expression in a pbs2delta mutant of *Saccharomyces cerevisiae* during adaptation to hypersaline stress. *FEBS Lett* *403*, 173-180.
- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology* *2*, 2006.0008-2006.0008.
- Bächler, C., Schneider, P., Bähler, P., Lustig, A., and Erni, B. (2005). *Escherichia coli* dihydroxyacetone kinase controls gene expression by binding to transcription factor DhaR. *The EMBO Journal* *24*, 283-293.
- Bales, J.R., Higham, D.P., Howe, I., Nicholson, J.K., and Sadler, P.J. (1984). Use of high-resolution proton nuclear magnetic resonance spectroscopy for rapid multi-component analysis of urine. *Clinical chemistry* *30*, 426-432.
- Barnhart, M.M., and Chapman, M.R. (2006). Curli biogenesis and function. *Annu Rev Microbiol* *60*, 131-147.
- Benov, L., and Beema, A.F. (2003). Superoxide-dependence of the short chain sugars-induced mutagenesis. *Free Radic Biol Med* *34*, 429-433.
- Ciriminna, R., Fidalgo, A., Ilharco, L.M., and Pagliaro, M. (2018). Dihydroxyacetone: An Updated Insight into an Important Bioproduct. *ChemistryOpen* *7*, 233-236.
- Constantinidou, C., Hobman, J.L., Griffiths, L., Patel, M.D., Penn, C.W., Cole, J.A., and Overton, T.W. (2006). A reassessment of the FNR regulon and transcriptomic analysis of the effects of nitrate, nitrite, NarXL, and NarQP as *Escherichia coli* K12 adapts from aerobic to anaerobic growth. *J Biol Chem* *281*, 4802-4815.
- Datsenko, K.A., and 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.
- de Boer, M., Broekhuizen, C.P., and Postma, P.W. (1986). Regulation of glycerol kinase by enzyme IIIIGlc of the phosphoenolpyruvate:carbohydrate phosphotransferase system. *J Bacteriol* *167*, 393-395.
- Deppenmeier, U., Hoffmeister, M., and Prust, C. (2002). Biochemistry and biotechnological applications of *Gluconobacter* strains. *Applied Microbiology and Biotechnology* *60*, 233-242.
- Durnin, G., Clomburg, J., Yeates, Z., Alvarez, P.J., Zygorakis, K., Campbell, P., and Gonzalez, R. (2009). Understanding and harnessing the microaerobic metabolism of glycerol in *Escherichia coli*. *Biotechnol Bioeng* *103*, 148-161.
- Esquerre, T., Moisan, A., Chiapello, H., Arike, L., Vilu, R., Gaspin, C., Coccagn-Bousquet, M., and Girbal, L. (2015). Genome-wide investigation of mRNA lifetime determinants in *Escherichia coli* cells cultured at different growth rates. *BMC Genomics* *16*, 275.
- Freedberg, W.B., and Lin, E.C.C. (1973). Three Kinds of Controls Affecting the Expression of the *glp* Regulon in *Escherichia coli*. *Journal of Bacteriology* *115*, 816-823.
- Garcia-Alles, L.F., Siebold, C., Nyffeler, T.L., Flükiger-Brühwiler, K., Schneider, P., Bürgi, H.-B., Baumann, U., and Erni, B. (2004). Phosphoenolpyruvate- and ATP-Dependent Dihydroxyacetone Kinases: Covalent Substrate-Binding and Kinetic Mechanism[†]. *Biochemistry* *43*, 13037-13045.
- Gelius-Dietrich, G., Desouki, A.A., Fritzemeier, C.J., and Lercher, M.J. (2013). Sybil--efficient constraint-based modelling in R. *BMC systems biology* *7*, 125-125.

Gonzalez, R., Murarka, A., Dharmadi, Y., and Yazdani, S.S. (2008). A new model for the anaerobic fermentation of glycerol in enteric bacteria: Trunk and auxiliary pathways in *Escherichia coli*. *Metabolic Engineering* 10, 234-245.

Green, R., and Rogers, E.J. (2013). Transformation of chemically competent *E. coli*. *Methods in enzymology* 529, 329-336.

Gunsalus, R.P., and Park, S.J. (1994). Aerobic-anaerobic gene regulation in *Escherichia coli*: control by the ArcAB and Fnr regulons. *Res Microbiol* 145, 437-450.

Gutknecht, R., Beutler, R., Garcia-Alles, L.F., Baumann, U., and Erni, B. (2001). The dihydroxyacetone kinase of *Escherichia coli* utilizes a phosphoprotein instead of ATP as phosphoryl donor. *The EMBO Journal* 20, 2480-2486.

Hayashi, S.-I., and Lin, E.C.C. (1967). Purification and properties of glycerol kinase from *Escherichia coli*. *Journal of Biological Chemistry* 242, 1030-1035.

Hekmat, D., Bauer, R., and Fricke, J. (2003). Optimization of the microbial synthesis of dihydroxyacetone from glycerol with *Gluconobacter oxydans*. *Bioprocess Biosyst Eng* 26, 109-116.

Heux, S., Brautaset, T., Vorholt, J.A., Wendisch, V.F., and Portais, J.C. (2018). Synthetic Methylophony: Past, Present, and Future. In *Methane Biocatalysis: Paving the Way to Sustainability*, M.G. Kalyuzhnaya, and X.-H. Xing, eds. (Cham: Springer International Publishing), pp. 133-151.

Heux, S., Philippe, B., and Portais, J.C. (2011). High-throughput workflow for monitoring and mining bioprocess data and its application to inferring the physiological response of *Escherichia coli* to perturbations. *Appl Environ Microbiol* 77, 7040-7049.

Jacques, M. (1996). Role of lipo-oligosaccharides and lipopolysaccharides in bacterial adherence. *Trends Microbiol* 4, 408-409.

Jefferson, K.K. (2004). What drives bacteria to produce a biofilm? *FEMS Microbiol Lett* 236, 163-173.

Jin, R.Z., Forage, R.G., and Lin, E.C. (1982). Glycerol kinase as a substitute for dihydroxyacetone kinase in a mutant of *Klebsiella pneumoniae*. *Journal of bacteriology* 152, 1303-1307.

Jin, R.Z., and Lin, E.C.C. (1984). An inducible phosphoenolpyruvate: dihydroxyacetone phosphotransferase system in *Escherichia coli*. *Microbiology* 130, 83-88.

Kang, Y.S., Weber, K.D., Yu, Q., Kiley, P.J., and Blattner, F.R. (2005). Genome-wide expression analysis indicates that FNR of *Escherichia coli* K-12 regulates a large number of genes of unknown function. *Journal of Bacteriology* 187, 1135-1160.

Keseler, I.M., Mackie, A., Santos-Zavaleta, A., Billington, R., Bonavides-Martinez, C., Caspi, R., Fulcher, C., Gama-Castro, S., Kothari, A., Krummenacker, M., *et al.* (2017). The EcoCyc database: reflecting new knowledge about *Escherichia coli* K-12. *Nucleic Acids Res* 45, D543-D550.

Kremer, D.R., and Hansen, T.A. (1987). Glycerol and Dihydroxyacetone Dissimilation in *Desulfovibrio* Strains. *Arch Microbiol* 147, 249-256.

Lip, H., Yang, K., MacAllister, S.L., and O'Brien, P.J. (2013). Glyoxal and methylglyoxal: autoxidation from dihydroxyacetone and polyphenol cytoprotective antioxidant mechanisms. *Chem Biol Interact* 202, 267-274.

Luchi, S., and Lin, E.C. (1991). Adaptation of *Escherichia coli* to respiratory conditions: regulation of gene expression. *Cell* 66, 5-7.

Mahadevan, R., and Schilling, C.H. (2003). The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metabolic Engineering* 5, 264-276.

- Maksimović, V., Mojović, M., and Vučinić, Ž. (2006). Monosaccharide–H₂O₂ reactions as a source of glycolate and their stimulation by hydroxyl radicals. *Carbohydrate Research* 341, 2360-2369.
- McDowall, J.S., Murphy, B.J., Haumann, M., Palmer, T., Armstrong, F.A., and Sargent, F. (2014). Bacterial formate hydrogenlyase complex. *Proc Natl Acad Sci U S A* 111, E3948-3956.
- Molin, M., Norbeck, J., and Blomberg, A. (2003). Dihydroxyacetone Kinases in *Saccharomyces cerevisiae* Are Involved in Detoxification of Dihydroxyacetone. *Journal of Biological Chemistry* 278, 1415-1423.
- Muller, J.E.N., Meyer, F., Litsanov, B., Kiefer, P., Potthoff, E., Heux, S., Quax, W.J., Wendisch, V.F., Brautaset, T., Portais, J.C., *et al.* (2015). Engineering *Escherichia coli* for methanol conversion. *Metab Eng* 28, 190-201.
- O'connor, M.B. (1980). Pentose Phosphate dependent Fixation of Formaldehyde by Methanol-grown *Hansenula polymorpha* and *Candida boidinii*. *Journal of General Microbiology*, 219-255.
- Orth, J.D., Conrad, T.M., Na, J., Lerman, J.A., Nam, H., Feist, A.M., and Palsson, B.Ø. (2011). A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism--2011. *Molecular systems biology* 7, 535-535.
- Ouellette, M., Makkay, A.M., and Papke, R.T. (2013). Dihydroxyacetone metabolism in *Haloferax volcanii*. *Frontiers in Microbiology* 4.
- Partridge, J.D., Scott, C., Tang, Y., Poole, R.K., and Green, J. (2006). *Escherichia coli* transcriptome dynamics during the transition from anaerobic to aerobic conditions. *J Biol Chem* 281, 27806-27815.
- Paulsen, I.T., Reizer, J., Jin, R.Z., Lin, E.C.C., and Saier Jr, M.H. (2000). Functional genomic studies of dihydroxyacetone utilization in *Escherichia coli*. *Microbiology* 146, 2343–2344.
- Pellicer, M.T., Fernandez, C., Badia, J., Aguilar, J., Lin, E.C.C., and Baldoma, L. (1999). Cross-induction of *glc* and *ace* operons of *Escherichia coli* attributable to pathway intersection - Characterization of the *glc* promoter. *Journal of Biological Chemistry* 274, 1745-1752.
- R Development Core Team (2009). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).
- Reizer, J., Reizer, A., and Saier Jr, M.H. (1995). Novel phosphotransferase system genes revealed by bacterial genome analysis—a gene cluster encoding a unique Enzyme I and the proteins of a fructose-like permease system. *Microbiology* 141, 961–971.
- Roberton, A.M., Sullivan, P.A., Jones-Mortimer, M.C., and Kornberg, H.L. (1980). Two genes affecting glucarate utilization in *Escherichia coli* K12. *J Gen Microbiol* 117, 377-382.
- Rolfe, M.D., Ocone, A., Stapleton, M.R., Hall, S., Trotter, E.W., Poole, R.K., Sanguinetti, G., Green, J., and Sys, M.O.S.C. (2012). Systems analysis of transcription factor activities in environments with stable and dynamic oxygen concentrations. *Open Biol* 2, 120091.
- Sánchez-Moreno, I., Nauton, L., Théry, V., Pinet, A., Petit, J.-L., de Berardinis, V., Samland, A.K., Guérard-Hélaine, C., and Lemaire, M. (2012). FSAB: A new fructose-6-phosphate aldolase from *Escherichia coli*. Cloning, over-expression and comparative kinetic characterization with FSAA. *Journal of Molecular Catalysis B: Enzymatic* 84, 9-14.
- Schurmann, M. (2001). Fructose-6-phosphate Aldolase Is a Novel Class I Aldolase from *Escherichia coli* and Is Related to a Novel Group of Bacterial Transaldolases. *Journal of Biological Chemistry* 276, 11055-11061.
- Silva-Rocha, R., Martinez-Garcia, E., Calles, B., Chavarria, M., Arce-Rodriguez, A., de Las Heras, A., Paez-Espino, A.D., Durante-Rodriguez, G., Kim, J., Nikel, P.I., *et al.* (2013). The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res* 41, D666-675.

- Sokol, S., Millard, P., and Portais, J.C. (2012). *influx_s*: increasing numerical stability and precision for metabolic flux analysis in isotope labelling experiments. *Bioinformatics* 28, 687-693.
- Stanko, R.T., Robertson, R.J., Spina, R.J., Reilly Jr, J.J., Greenawalt, K.D., and Goss, F.L. (1990). Enhancement of arm exercise endurance capacity with dihydroxyacetone and pyruvate. *J Appl Physiol* 68, 119-124.
- Streekstra, H., Teixeira de Mattos, M.J., Neijssel, O.M., Tempest, D.W., (1987). Overflow metabolism during anaerobic growth of *Klebsiella aerogenes* NCTC 418 on glycerol and dihydroxyacetone in chemostat culture. *Arch Microbiol* 147, 268-275.
- Subedi, K.P., Kim, I., Kim, J., Min, B., and Park, C. (2008). Role of GldA in dihydroxyacetone and methylglyoxal metabolism of *Escherichia coli* K12. *FEMS Microbiology Letters* 279, 180-187.
- Tang, C.T., Ruch, F.E., and Lin, C.C. (1979). Purification and properties of a nicotinamide adenine dinucleotide-linked dehydrogenase that serves an *Escherichia coli* mutant for glycerol catabolism. *Journal of bacteriology* 140, 182–187.
- Team, R. (2015). RStudio: Integrated Development for R. (RStudio, Inc., Boston, MA).
- Tessier, F.J., Monnier, V.M., Sayre, L.M., and Kornfield, J.A. (2003). Triosidines: novel Maillard reaction products and cross-links from the reaction of triose sugars with lysine and arginine residues. *The Biochemical journal* 369, 705-719.
- Truniger, V., and Boos, W. (1994). Mapping and cloning of *gldA*, the structural gene of the *Escherichia coli* glycerol dehydrogenase. *J Bacteriol* 176, 1796-1800.
- Tseng, C.P., Albrecht, J., and Gunsalus, R.P. (1996). Effect of microaerophilic cell growth conditions on expression of the aerobic (*cyoABCDE* and *cydAB*) and anaerobic (*narGHJ*, *frdABCD*, and *dmsABC*) respiratory pathway genes in *Escherichia coli*. *J Bacteriol* 178, 1094-1098.
- Uden, G., Steinmetz, P.A., and Degreif-Dunnwald, P. (2014). The Aerobic and Anaerobic Respiratory Chain of *Escherichia coli* and *Salmonella enterica*: Enzymes and Energetics. *EcoSal Plus* 6.
- Volker F. Wendisch, S.N.L.a.T.M.M. (2011). Use of Glycerol in Biotechnological Applications. *Biodiesel*, Gisela Montero and Margarita Stoytcheva, IntechOpen.
- Wang, L., Chauliac, D., Rhee, M.S., Panneerselvam, A., Ingram, L.O., and Shanmugam, K.T. (2018). Fermentation of dihydroxyacetone by engineered *Escherichia coli* and *Klebsiella variicola* to products. *Proc Natl Acad Sci U S A* 115, 4381-4386.
- Weinhouse, H., and Benziman, M. (1976). Phosphorylation of glycerol and dihydroxyacetone in *Acetobacter xylinum* and its possible regulatory role. *Journal of bacteriology* 127, 747–754.
- Yaylayan, V.A., Harty-Majors, S., and Ismail, A.A. (1999). Investigation of DL-glyceraldehyde–dihydroxyacetone interconversion by FTIR spectroscopy. *Carbohydrate Research* 318, 20–25.
- Zhang, G., Meredith, T.C., and Kahne, D. (2013). On the essentiality of lipopolysaccharide to Gram-negative bacteria. *Curr Opin Microbiol* 16, 779-785.

Supplementary data

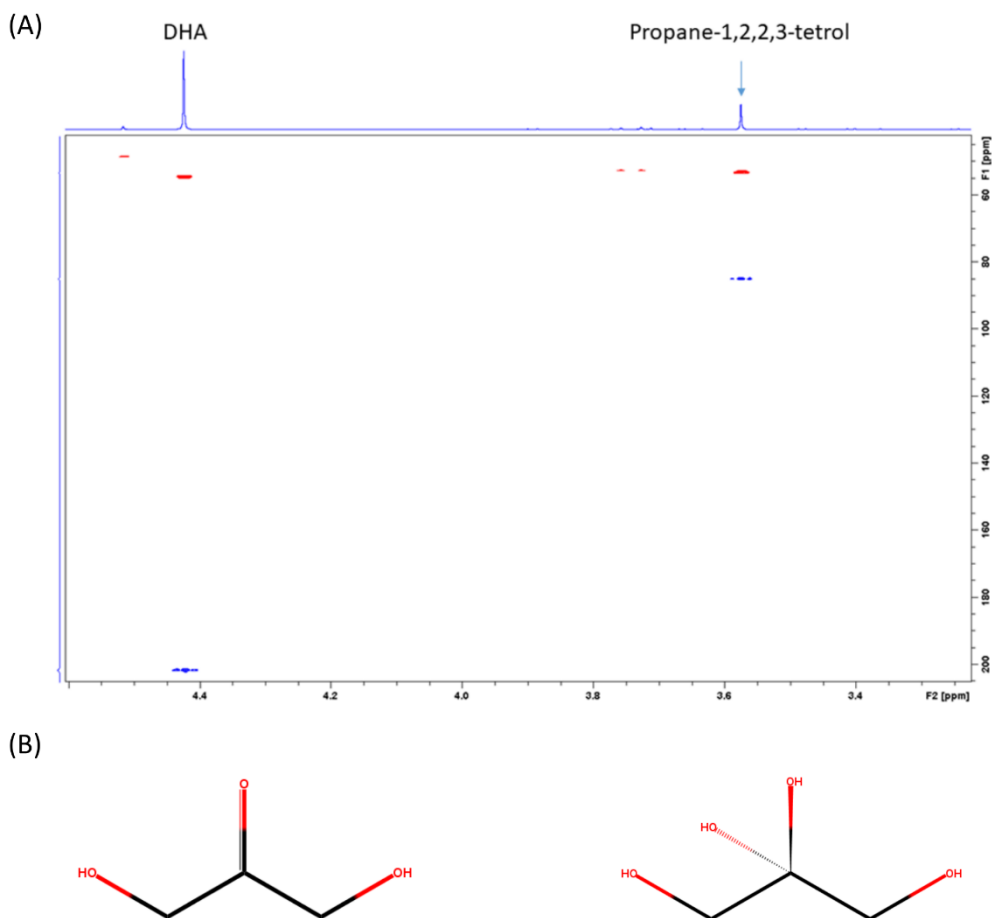


Figure S1: Identification of propane-1,2,2,3-tetrol by 2D-NMR analyses in a sample of DHA in M9 medium with 5X diluted salts. (A) Overlay of Heteronuclear Single Quantum Correlation spectroscopy (HSQC) experiment in red and Heteronuclear Multiple Bonds Correlation spectroscopy (HMBC) experiment in blue. (B) DHA (left) and propane-1,2,2,3-tetrol (right) condensed formula.

The HSQC experiments enables to see one bond correlation between the ^1H and ^{13}C atoms called $^1J_{\text{CH}}$ and the HMBC experiments exhibits $^2J_{\text{CH}}$ to $^3J_{\text{CH}}$, i.e. long distance correlations between a ^1H and a ^{13}C . From HSQC acquisitions (Figure S1A in red), we observed that protons at 4.4 ppm correlates with carbon resonance at 54.4 ppm and ^1H 3.6 ppm signal correlates with 53.0 ppm carbon signal. Knowing that the DHA molecular structure (Figure S1B), we can assume that the peak at 54.4 ppm corresponds to a hydroxyl group. Likewise, we can assume that the peak at 53.0 ppm corresponds also to a hydroxyl group with a very slight deshielded. From the HMBC spectrum (Figure S1A in blue), a correlation peak between ^1H at 4.4 ppm and ^{13}C at 200 ppm was observed corresponding to the ketone group. For the unknown molecule, a correlation peak between ^1H at 3.6 ppm and carbon at 82.9 ppm was observed (Figure S1A). This carbon chemical shift can be due to a deshielded carbon bound to two hydroxyl groups. According to these data, we suggest that the unknown molecule is most probably a derived form of DHA. This is in accordance with the results of Yaylayan and colleagues (Yaylayan et al., 1999) who described that in water DHA can be hydrated into a secondary form i.e. propane-1,2,2,3-tetrol (or dihydroxyacetone monomer hydrate) (Figure S1B).

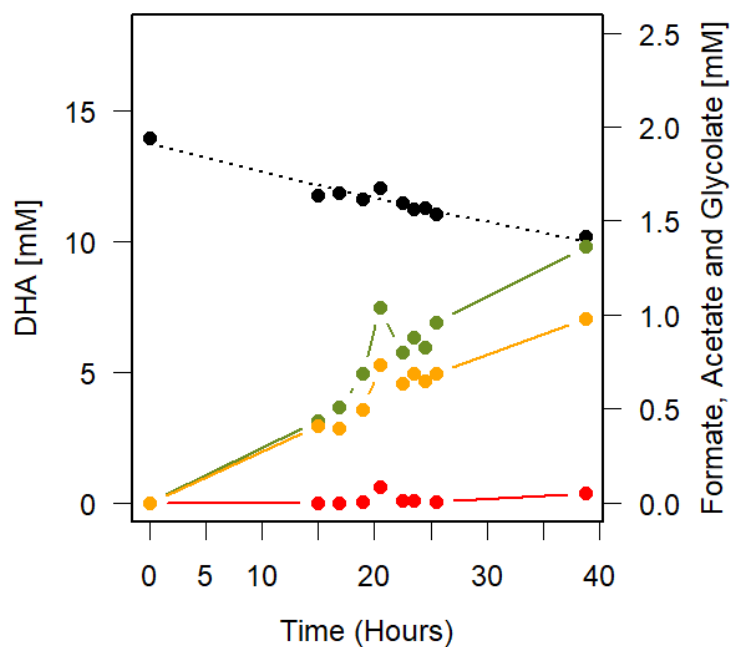


Figure
course

S2: Time analysis of

the fate of DHA in non-inoculated modified M9 medium at 37°C, 250 rpm shaking. Kinetic of degradation of the DHA (experimental (black points) and fitted (dashed black lines) into formate (green), glycolate (orange) and acetate (red). One out of three experiments is shown.

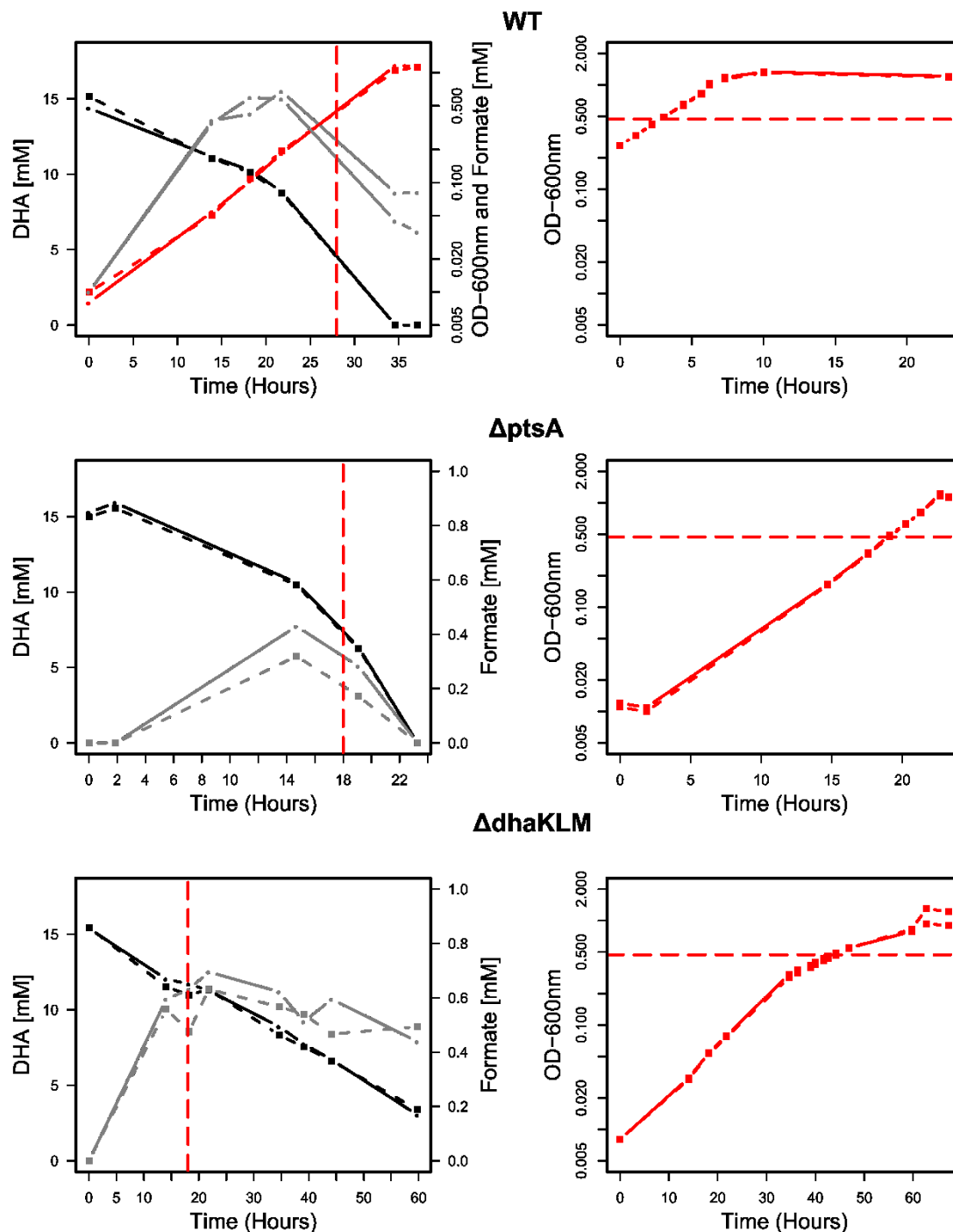


Figure S3: Time course analysis of the fate of DHA, in *E. coli* BW25113 strain (WT) and in the deleted for *ptsA* ($\Delta ptsA$) and *dhaKLM* ($\Delta dhaKLM$) strains. Cells were grown modified M9 medium with 15 mM of DHA at 37°C, 250 rpm shaking. DHA (black), formate (grey) and growth (red). Vertical dashed red lines indicate the sampling time point for transcriptomic analysis. The axis of the growth plot are in logarithmic scale. For the WT strain, the sampling for transcriptomic analysis was done the next day from the same cultivation inoculated in a fresh medium at $OD_{600nm} = 0.2$. ($n=2$ biological replicates). Supernatant were analysed by HPLC and NMR. Results obtain by HPLC are presented. Growth rate and specific DHA uptake rate for each strains are provided on Figure 4.

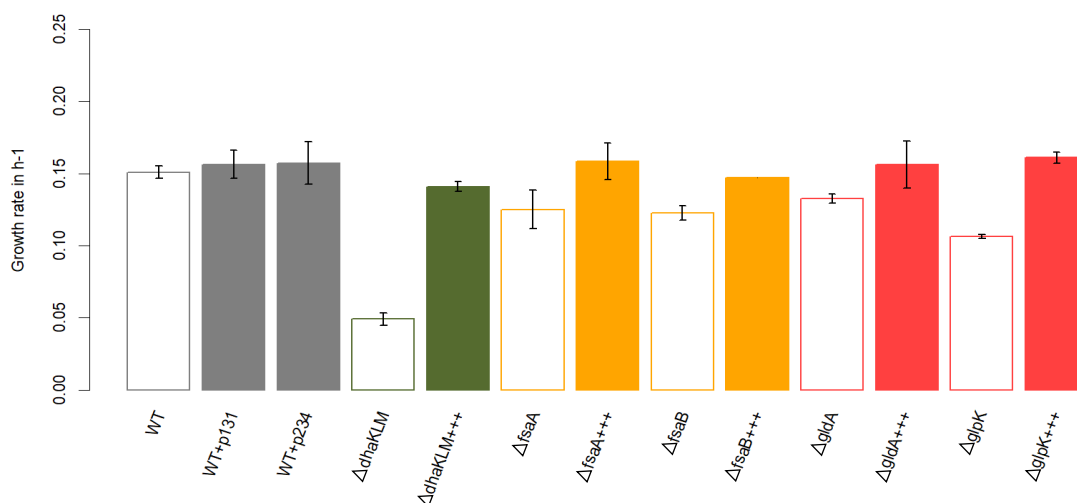


Figure S4: Growth rate of the deleted ($\Delta gene$) and complemented ($\Delta gene+++$) strain. $n = 2$ biological replicates for the complemented strain and 3 biological replicates for the deleted strains. Cells were grown modified M9 medium with 15 mM of DHA at 37°C, 250 rpm shaking. Please refer to Table 2 for details on each strain.

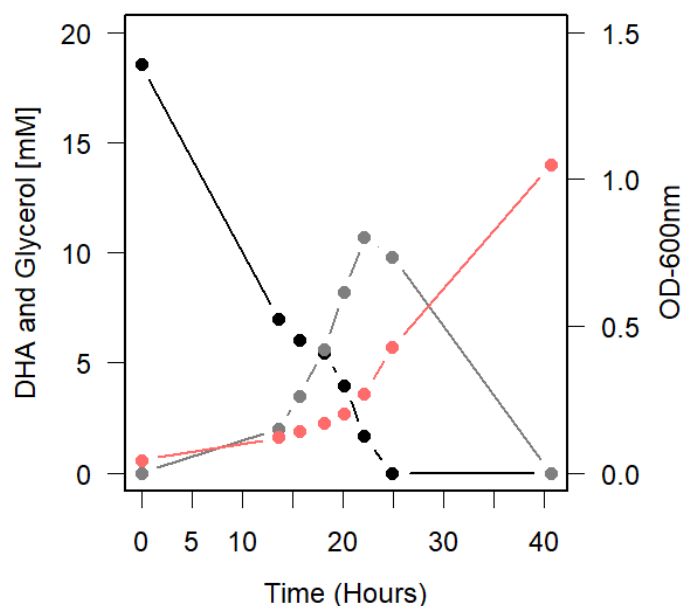


Figure S5: Time course analysis of the fate of DHA in an *E. coli* BW25113 strain overexpressing *gldA*. Cells were grown modified M9 medium with 15 mM of DHA at 37°C, 250 rpm shaking. DHA (black), glycerol (grey) and growth (red). One over three experiment is shown.

Chapter 3

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

Alessandro de Simone¹, Cláudia M. Vicente¹, Camille Peiro¹, Lara Gales^{1,2}, Floriant Bellvert^{1,2}, Brice Enjalbert¹, Stéphanie Heux¹

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

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

Metabolic Engineering, 2020, Vol 61, Pages 315-325

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 ¹³C-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. gernerii* 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*.

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 CO₂. These approaches, which are still under development, can provide a way to recycle emitted CO₂ 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 CO₂, and (iii) the assimilation of one carbon compounds, either formaldehyde or CO₂ 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*.

Material and methods

Reagents

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

Bacterial strains and culture media

All the strains, plasmids, primers and synthetic gene constructs used in this study are listed in Supplementary Table S1. *E. coli* DH5 α was used for plasmid construction and propagation whereas *E. coli* BW25113 was used for methanol assimilation. *E. coli* BW25113 $\Delta\text{frma}::\text{neo}$ was obtained from the Keio collection and the Flp recognition target (FRT)-flanked kanamycin cassette was removed using Flp recombinase from pCP20 plasmid (Cherepanov and Wackernagel, 1995). After recombination, loss of pCP20 was confirmed by re-streaking on ampicillin, and removal of the resistance cassette was confirmed by polymerase chain reaction (PCR). For operon construction, *A. gernerii* Mdh and *P. angusta* Das genes, containing RBS and a 6xHis tag, were amplified from the pSEVA plasmids using primers P91&P92 and P83&P93, respectively. The two fragments, designed to overlap by 35 bp, were joined by overlapping PCR. The complete Mdh-Das operon was subsequently cloned into the pSEVA424 vector using primers P102&P103 and the In-Fusion® HD kit (Takara Bio, Otsu, Japan). The λ red recombination method (Datsenko and Wanner, 2000) was used to generate knockout strains ΔptsA (primers P129&P130) and ΔfrmRAB (primers P145&P146). The introduced antibiotic resistance cassettes were removed using the FRT/FLP recombination system (Cherepanov and Wackernagel, 1995). All constructs were subsequently verified by colony PCR and sequencing (GATC, Konstanz, Germany).

All *E. coli* strains harboring plasmids were propagated in Luria-Bertani (LB) medium or M9 minimal medium containing the appropriate antibiotics. The composition of the M9 minimal medium was as follows (in $\text{g}\cdot\text{L}^{-1}$): 18 Na_2HPO_4 , 3.13 KH_2PO_4 , 0.53 NaCl , 2.11 NH_4Cl , 0.49 $\text{MgSO}_4\cdot 7\text{H}_2\text{O}$, 0.00438 $\text{CaCl}_2\cdot 2\text{H}_2\text{O}$, 0.1 thiamine hydrochloride, trace elements (mg L^{-1}) 15

Na₂EDTA·2H₂O, 4.5 ZnSO₄·7H₂O, 0.3 CoCl₂·6H₂O, 1 MnCl₂, 1 H₃BO₃, 0.4 Na₂MoO₄·2H₂O, 3 FeSO₄·7H₂O, 0.3 CuSO₄·5H₂O. The antibiotics were added when necessary in the following concentrations: ampicillin (Amp, 100 µg/ml), kanamycin (Kan, 50 µg/ml), streptomycin (Strp, 50 µg/ml). The optical density at 600 nm (OD₆₀₀) was measured using a GENESYS 6™ spectrophotometer (Thermo Scientific).

In silico design of the synthetic pathway

The synthetic pathway for methanol assimilation was designed using the software FindPath (Vieira et al., 2014). The workflow starts with the creation of a substrate-associated reaction database based on the literature and available metabolic databases. This database consists of reactions involving the target molecule (in our case, methanol). The database is then converted into a model that is subsequently used to compute elementary flux modes (EFMs), i.e., all the possible flux distributions in a metabolic network under steady state conditions. Among these EFMs, the best pathways are selected and ranked according to their efficiency. Finally, the best module combinations for efficient methanol conversion were identified. In our case, the methanol database encompassed more than 100 reactions steps and 100 metabolic compounds involved in methanol metabolism. For each reaction, the genes, reaction, EC number, KEGG name, localization, and reversibility were reported. Finally, the model was built by bringing together all the reactions along with the transporters and cofactor recycling, i.e. 47 reactions and 114 metabolites. Using this model, 10000 EFMs were generated of which 85 allowed the conversion of methanol into *E. coli* metabolites. From these, 20 efficient EFMs were selected, i.e. those involving a small number of reactions and with low cofactor consumption (ATP, NAD(P)H). The hypothesis was that their introduction into the host would require little genetic effort (the number of genes being correlated with the number of reactions) and would have little or no effect on the host's energy and redox machinery. The reactions composing the 20 EFMs were implemented in a genome scale model *E. coli* (iAF1620). Finally, the biomass yields on methanol of each of the 20 EFMs were simulated using *in-silico* flux balance analysis (FBA).

Library generation by combinatorial assembly

A BLAST search against UniRef50 (Suzek et al., 2014) using *B. methanolicus* PB1 Mdh2 (UniProt ID: I3DVX6) and *P. angusta* Das (UniProt ID: P06834) as query sequences returned two clusters with 177 and 230 members, respectively. The sequence clustering tool H-CD-HIT (Huang et al., 2010) was used to hierarchically merge similar sequences at varying levels of sequence identity. Proteins were first clustered at a high identity (90%) before the non-redundant sequences were further clustered at a low identity (80% and eventually 70%). Among the representatives of the different clusters, we selected 12 putative Mdh variants and 17 putative Das variants from aerobic and mesophilic microorganisms. The corresponding Mdh genes, as well as the *Bacillus stearothermophilus* Mdh (Dowds et al., 1988) and *Pichia angusta* Mox genes (Shleev et al., 2006), were cloned in the expression vector pSEVA424 (Silva-Rocha et al., 2013) between restriction sites AvrII and NotI. The selected Das genes, plus the *Mycobacterium* Das gene (Ro et al., 1997) and an *E. coli* codon-optimized version of P06834, were cloned in the expression vector pSEVA134 (Silva-Rocha et al., 2013) between restriction sites AvrII and SpeI. All the constructs were synthesized and cloned by BaseClear (Leiden, The Netherlands). The same ribosome binding site (RBS) (AGGAGGAAAAACAT) and 6xHis tag was used for all the genes. The two gene libraries were co-transformed in the BW25113 Δ frmA::frr strain using the rubidium chloride method (Green and Rogers, 2013) and plated on LB-Amp-Strp plates (Supplementary Figure S3).

Dynamic ¹³C-labelling incorporation

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

After 90 and 180 min incubation at 30°C and 220 rpm, intracellular metabolites were sampled as follows: 120 μ L of culture was taken and mixed with 1 mL of cold (-20°C) acetonitrile:methanol:water:formic acid (40:40:20:0.1) extraction solution. The samples were

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

After 48 h of incubation at 30°C and 220 rpm, proteinogenic amino acids were sampled as follows: the plates were centrifuged at 4400g for 3 min and the supernatant was removed. To release protein-bound amino acids from cellular proteins, the cell pellets collected were hydrolyzed for 15 h with 6N HCl at 100°C. HCl was evaporated at low pressure (20 mbar, room temperature). Biomass hydrolysates were washed twice in water using the same evaporation method. The dried hydrolysates were resuspended in 200 μ L water and centrifuged. A 10-fold dilution was prepared, and samples were analyzed by LC-HRMS. Proteinogenic amino acids were separated on a Supelco™ HS F5 DISCOVERY column (150 \times 2.1 mm; 5 μ m) equipped with a SUPELGUARD KIT HS F5 guard column (20 \times 2.1 mm; 5 μ m) 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 desolvation temperatures

of 250°C. The spectrometer was operated in full-scan mode at a resolution of 60,000 (400 m/z).

¹³C-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 ¹³C-isotopic enrichment were then determined as follows: $^{13}\text{C}\text{-enrichment (\%)} = \frac{\sum(M_i \cdot i)}{n}$, where n is the number of carbon atoms for the measured fragment and M_i is the corrected abundance of the mass isotopologue.

Supernatant analysis

Metabolite utilization and the production of the synthetic methylotroph were analyzed by quantitative 1D ¹H-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 sodium salt (TSP-d4, 1 mM) as a chemical shift and concentration standard. The concentrations of the different metabolites (xylose, methanol, formate, and acetate) were calculated with the following equation: concentration = integrated peak area * TSP concentration * dilution of the sample / number of protons in the molecule. For xylose, only the peaks corresponding to the anomeric protons were integrated.

Transcriptomic analysis

Cells were grown in flasks of M9 minimal media containing 15 mM xylose with or without 150 mM MeOH. At T1 (OD600 = 1, exponential phase) and T2 (OD600 = 2, stationary phase), 4 mL of each culture was centrifuged for 90 s at 14000 rpm before discarding the supernatant and immediately freezing the pellets in liquid nitrogen. Total RNA was extracted according to

the Qiagen RNAeasy MiniKit procedure and quantified using a Nanodrop® spectrophotometer. Double-stranded complementary DNA (cDNA) synthesis and array processing were performed using the Agilent Technologies One-Color Microarray-Based Gene Expression Analysis protocol. The images were analyzed with the software DEVA (v. 1.2.1). All array procedures were performed using the GeT-Biopuces platform (<http://get.genotoul.fr/>). For each data set, corresponding to time point T1 or T2, the log₂ intensities obtained in the presence of methanol were divided by the log₂ intensities obtained without methanol. These ratios were then normalized by the log median intensity. Genes whose expression level differed by a factor of 2 or more between the two conditions were selected for further analysis. Gene ontology analyses were performed using Ecocyc (<https://ecocyc.org/>). Gene expression data have been deposited in the ArrayExpress database at EMBL-EBI (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-8909.

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).

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.

Results

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 predicted fluxes show that this optimal growth rate is achieved when 16% of the methanol is incorporated into the biomass with the rest being used to recycle the C1 acceptor, Xu5P. No flux through the methanol oxidation pathway (i.e. through FrmA & B) was predicted. In addition, the GSM predicted that Xu5P would be recycled by fructose-6-phosphate aldolase and transaldolase (FSA/TAL pathway variant) rather than by fructose 1,6-bisphosphate aldolase with transaldolase (FBA/TAL variant), or by sedoheptulose biphosphatase (FBA/SBP variant) (Supplementary Figure S1). No ATP is required for Xu5P regeneration in the FSA/TAL variant, while the other two metabolic variants require two ATP molecules (Supplementary Figure S1). In comparison, in the synthetic RuMP based pathway, the C1 acceptor, ribulose-5-phosphate, is recycled using one or two ATP molecules. However, in term of assimilation no ATP is required in the synthetic RuMP based pathway while one equivalent of ATP is needed to phosphorylate DHA into DHAP (Supplementary Figure S1). Considering both recycling and assimilation, the synthetic RuMP pathway requires the same amount of ATP with the FBA/TA variant and one more ATP with

the FBA/SBP variant compared to the FSA/TA variant which can occur only with the proposed pathway (Woolston et al., 2018). Finally, the transketolase activity of Das may contribute to the regeneration of the C1 acceptor, making our pathway partially independent of the pentose phosphate pathway (PPP). This has been demonstrated recently in methylotrophic yeasts (Russmayer et al., 2015).

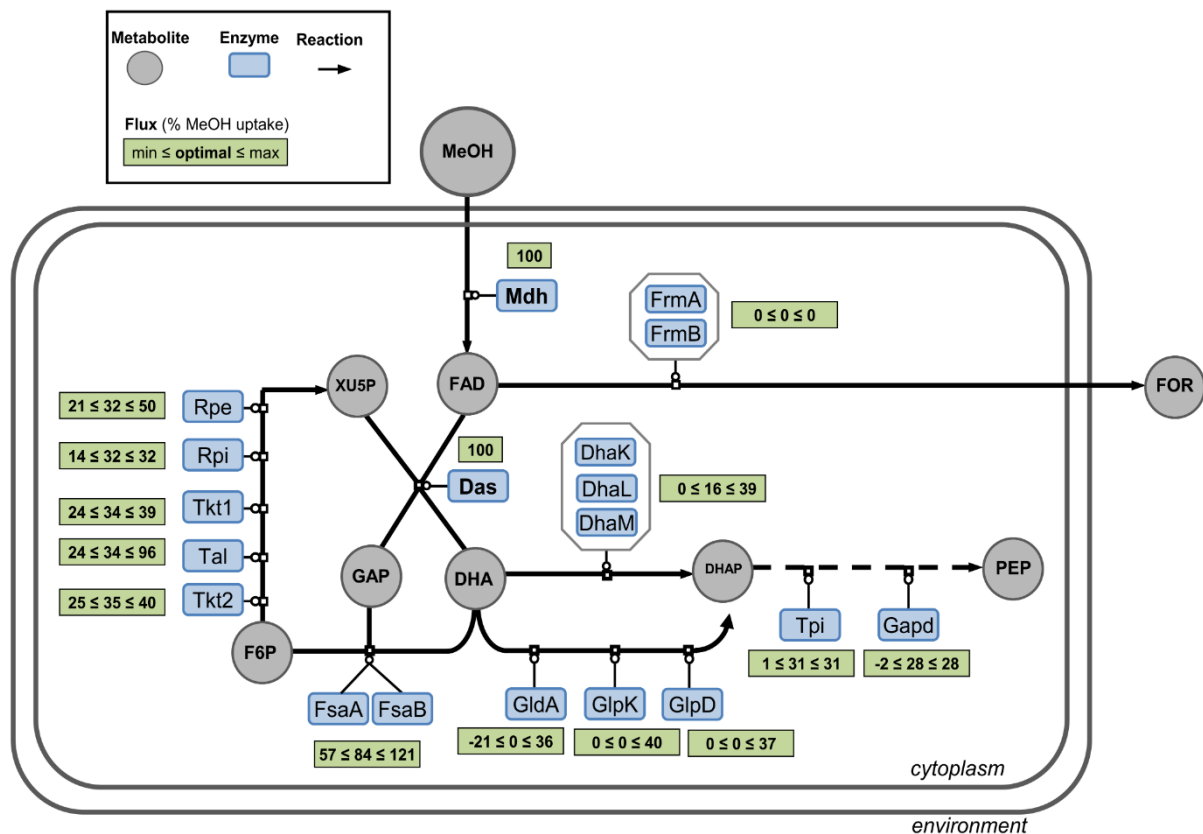


Figure 1: Overview of the synthetic methanol metabolism and its operation in *E. coli*. The new hybrid methanol assimilation pathway comprises a methanol dehydrogenase (Mdh) and a dihydroxyacetone synthase (Das). Green rectangles give the optimal and the ranges of simulated fluxes obtained using flux 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).

Screening for the best matching methylotrophic enzymes

To identify the combination of enzymes that would optimize the *in vivo* activity of the pathway in *E. coli*, we built a combinatorial library of *mdh* and *das* homologues derived from native and non-native methylotrophs (bacteria and yeasts). Starting with the query protein sequences of *Bacillus methanolicus* PB1 Mdh2 and *Pichia angusta* Das, a BLAST search with an identity cut-off at 50% was used with the software CD-HIT to filter out and cluster homologous templates (see Materials and Methods section). This threshold ensured that only functional homologues were identified (Sangar et al., 2007). Some of the 12 prokaryotic Mdh sequences and 17 eukaryotic Das sequences selected in this way belong to genera known to contain methylotrophs, such as *Bacillus*, *Burkholderia* (Chistoserdova et al., 2009), *Acinetobacter* (Del Rocío Bustillos-Cristales et al., 2017), *Pichia* and *Candida* (Supplementary Figure S2). We added an Mdh from *Bacillus stearothersophilus* and a *P. angusta* methanol oxidase (MOX), since both have been reported to have good affinity for methanol (K_ms of 20 mM and 0.4 mM, respectively) (Shleev et al., 2006; Whitaker et al., 2017). Finally we added a Das from *Mycobacterium* with a high affinity for formaldehyde (K_m of 1.86 mM) (Ro et al., 1997) and an *E. coli* codon-optimized version of *P. angusta* Das. The Mdh and Das genes were respectively cloned into the low-copy plasmid pSEVA424 and the middle-copy plasmid pSEVA134 (Silva-Rocha et al., 2013). All the selected sequences were assembled in a library of 266 combinations of genes (14 Mdh sequences * 19 Das sequences) and transformed using a robotic platform (Supplementary Figure S3). To prevent the induction of the formaldehyde detoxification pathway (formaldehyde to CO₂) (Figure 1), we used an *E. coli* strain Δ *frmA*, deleted for the formaldehyde dehydrogenase i.e. the first gene of this pathway. Enzyme expression were measured at two reasonable *E. coli* growth temperatures i.e. 30°C and 37°C. The higher expression obtained at 30°C overnight (Supplementary Figure S4) led us to use this temperature for all subsequent experiments.

To analyze the performance of the 266 different enzyme combinations, methanol incorporation was measured for each pathway variant using dynamic ¹³C-labeling experiments as shown in Supplementary Figure S3. We used the ¹³C-labeling 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 ¹³C-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. gernerii* 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 is also worth noting that although *B. stearrowthermophilus* Mdh and *Mycobacterium* Das have favorable *in vitro* and *in vivo* activities (Ro et al., 1997; Whitaker et al., 2017), and were well-expressed in *E. coli* (Supplementary Figure S5), they led to very low ^{13}C -enrichment (< 3%) in most of the tested combinations (Figure 2A).

Finally, the highest methanol incorporation was achieved when either *A. gernerii* Mdh or *Burkholderia sp. TSV86* Mdh was expressed in combination with *P. angusta* Das (opt). With these combinations, the ^{13}C -enrichments of PEP were respectively 13 and 12 times higher than with the query pathway (*B. methanolicus* PB1 Mdh2 / *P. angusta* Das). In particular, the fractions of PEP with one ^{13}C atom were 31% and 32% and reached 17% and 13.7% for two ^{13}C

atoms, respectively (Figure 2B). The incorporation of more than one labeled carbon into PEP demonstrates that the recycling of Xu5P is functional in both combinations.

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

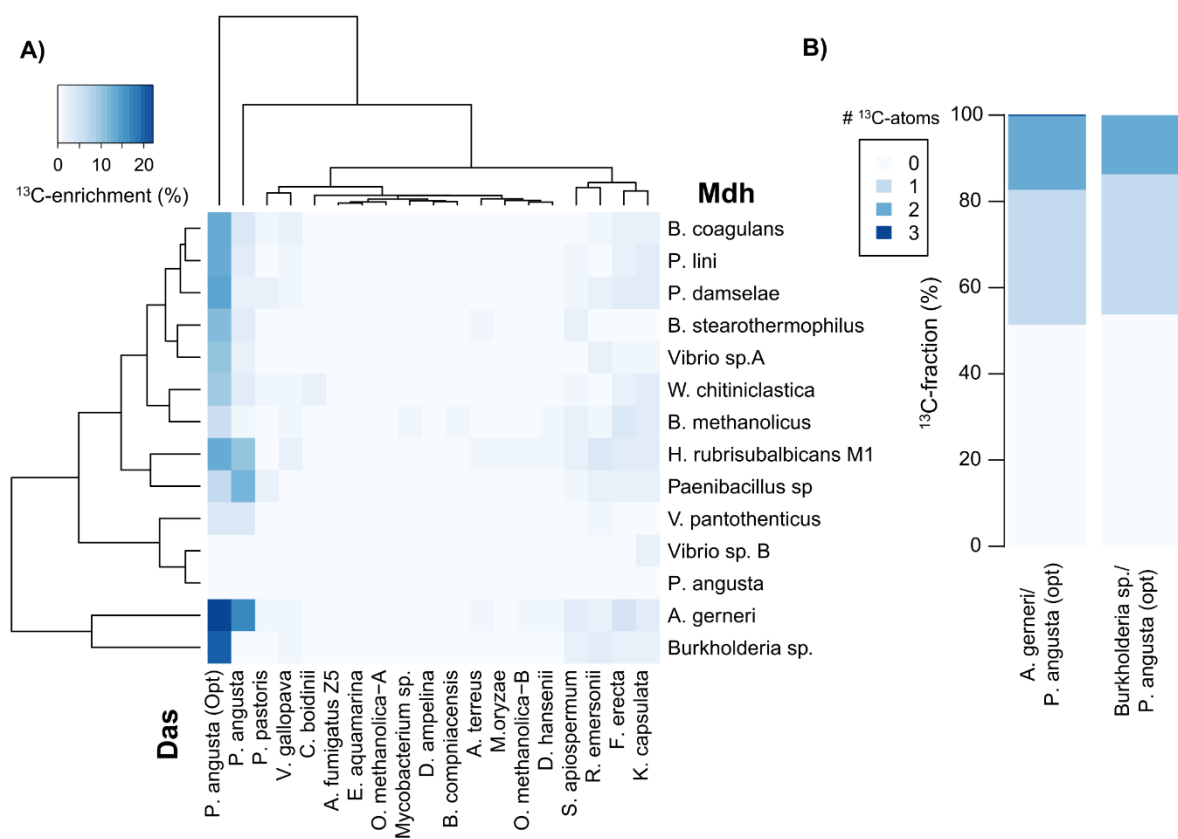


Figure 2: Screening of the different combinations of the hybrid methylotrophic pathway. (A) Heatmap showing the ^{13}C -enrichment of phosphoenolpyruvate (PEP) in *E. coli* ΔfrmA expressing different combinations of Mdh and Das homologues. The ^{13}C -enrichment of PEP was measured at the steady state 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 90 min in *E. coli* ΔfrmA expressing *P. angusta* Das (opt) with either *A. generii* Mdh or *Burkholderia* Mdh.

Characterizing the cellular behavior of the synthetic methylotroph

In order to uncover the specific make-up of the new synthetic methylotrophic *E. coli* strain with regard to methanol utilization, we performed a physiological and transcriptomic analysis of the strain expressing *A. gernerii* Mdh and *P. angusta* Das (opt) grown on xylose with and without additional methanol (Table 1, Figure 3). For this analysis, both genes were cloned into the pSEVA424 vector as a single operon to avoid the metabolic burdening of the cells with double-antibiotic selection (Silva et 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 $\Delta frmA$ strain (Table 1) suggests the presence of a promiscuous alcohol dehydrogenase that replaces FrmA in the reaction from formaldehyde to S-formylglutathione, which is then converted to formate by FrmB. Up-regulation of *dhaL* and *dhaM*, which encode the dihydroxyacetone kinase (DAK) pathway, was also observed in presence of methanol, (Figure 3B). Because the expression of the *dhaKLM* operon is induced by DHA (Bächler et al., 2005), this confirms the presence of DHA in the cells and thus the co-assimilation of methanol with xylose. However, the genes encoding alternative DHA assimilation routes (i.e. the glycerol (*gldA*, *glpK* and *glpD*) and the FSA (*fsaA* and *fsaB*) pathways) were not transcriptionally activated or even down-regulated (Figure 3B). These results are consistent with the conclusion of Peiro et al. that DHA is mainly assimilated via the dihydroxyacetone kinase (DAK) (Peiro et al., 2019). However, they appear to contradict those of the flux balance analysis that predict that Fsa may be involved in the regeneration of the C1 acceptor. Finally, the gene encoding the transketolase *tktA* was up-regulated on methanol (Figure 3B). This enzyme catalyzes the formation of Xu5P, which plays a key role in the cyclic operation of our synthetic pathway. However, Xu5P is also the entry point of xylose in the metabolism and, interestingly, the presence of methanol improved the expression of *xyIE* involved in its transport through the cellular membrane (Figure 3B), particularly during the exponential growth phase. This result corroborates the higher specific xylose uptake rate observed when the synthetic *E. coli* strain was grown in media supplemented with methanol (Table 1).

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

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* $\Delta frmA_pSEVA424\text{-Mdh}\text{-Das(opt)}$ 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.

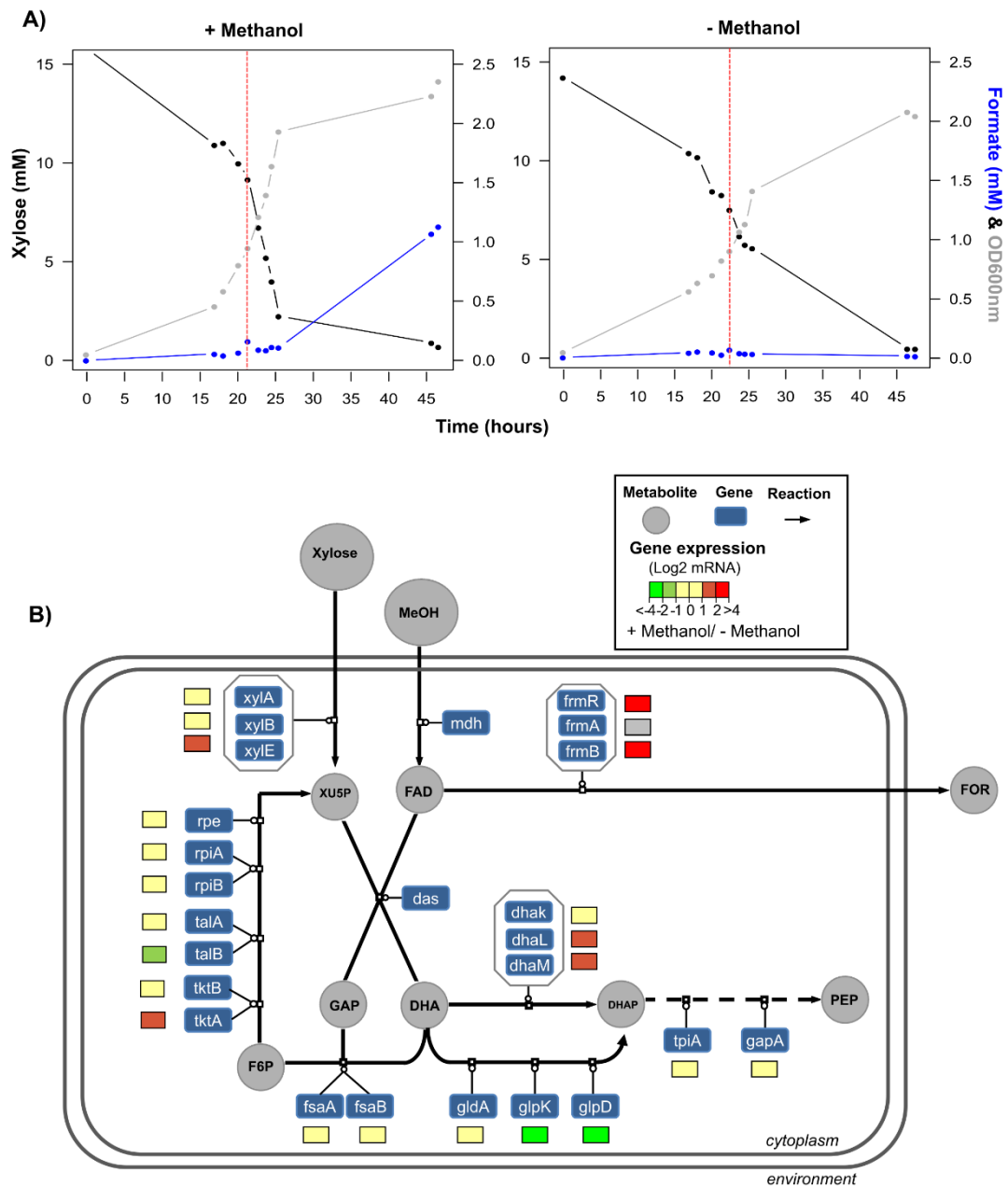


Figure 3. Response of the new synthetic methylotroph to methanol. (A) Time course analysis of xylose (black), formate (blue) and biomass (gray). Red dotted lines indicates the sampling point (T1 & T2) for transcriptomic analysis. Cells were grown in a minimal synthetic medium with 15mM of xylose with or without 150 mM of methanol at 30°C temperature, 250 rpm shaking. One exemplary 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 OD₆₀₀ = 1 (T1) and when xylose was completely consumed by the cells at OD₆₀₀ = 2 (T2).

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).

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 $\Delta ptsA::kan$ 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 $\Delta ptsA::kan$ 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 $\Delta frmA$ 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 -enrichment 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 -enrichment 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).

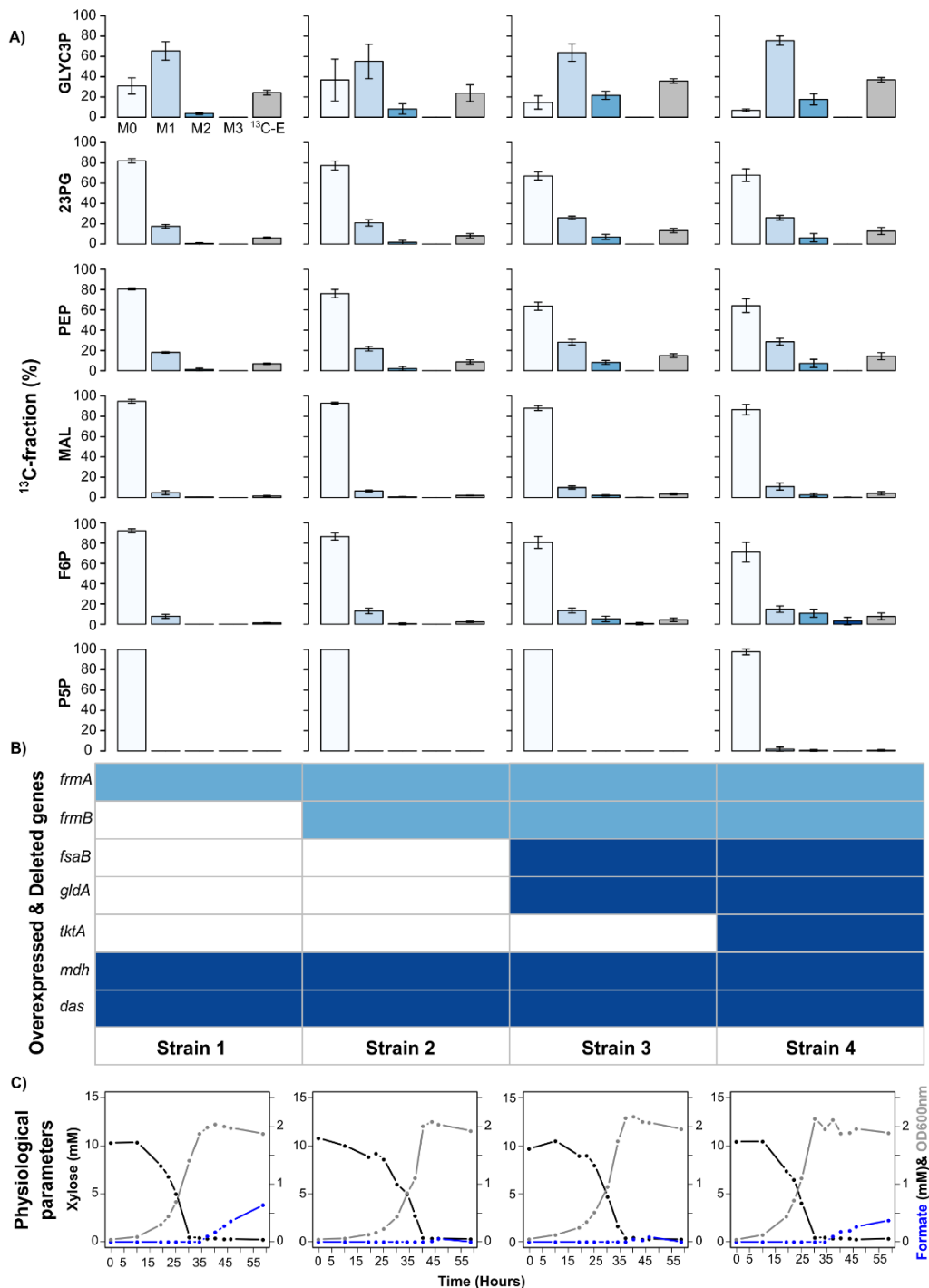


Figure 4: Phenotypic characterization of rationally designed *E. coli* strains. A) Labeling patterns of the intracellular metabolites 2 & 3 phosphoglycerate (23PG), fructose-6-phosphate (F6P), glycerol-3-phosphate (GLYC3P), pool of pentoses-5-phosphate (P5P), phosphoenolpyruvate (PEP) and malate (Mal) within the different strains after 90 min of culture in M9 medium with 655 mM ^{13}C -methanol. Mean of 3 replicates. 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 exemplary experiment for each strain is shown (n= 2).

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. gernerii* 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. gernerii* 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. stearothersophilus* 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* strain expressing a linear methanol assimilation pathway (Kim

et al., 2020) raises questions about the relevance of establishing methylotrophy in *E. coli* using cyclic pathways. Arguments in favor of pursuing the quest for growth on pure methanol using cyclic pathways are (i) the independence of such pathways from other carbon sources, and (ii) a recent study reporting an *E. coli* strain expressing an autotrophic cycle capable of producing all its biomass carbon from CO₂ (Gleizer et al., 2019).

References

- Ahmad, M., Hirz, M., Pichler, H., and Schwab, H. (2014). Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. *Appl Microbiol Biotechnol* 98, 5301-5317.
- Bächler, C., Schneider, P., Bähler, P., Lustig, A., and Erni, B. (2005). *Escherichia coli* dihydroxyacetone kinase controls gene expression by binding to transcription factor *DhaR*. *The EMBO Journal* 24, 283-293.
- Becker, J., and 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, 3328-3350.
- Bennett, R.K., Gonzalez, J.E., Whitaker, W.B., Antoniewicz, M.R., and 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., *et al.* (2013). Novel Methanol Dehydrogenase Enzymes from *Bacillus*.
- Chen, C.T., Chen, F.Y., Bogorad, I.W., Wu, T.Y., Zhang, R., Lee, A.S., and Liao, J.C. (2018). Synthetic methanol auxotrophy of *Escherichia coli* for methanol-dependent growth and production. *Metabolic engineering* 49, 257-266.
- Cherepanov, P.P., and 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, 2603-2622.
- Chistoserdova, L., Kalyuzhnaya, M.G., and Lidstrom, M.E. (2009). The expanding world of methylotrophic metabolism. *Annu Rev Microbiol* 63, 477-499.
- Chung, B.K., Selvarasu, S., Andrea, C., Ryu, J., Lee, H., Ahn, J., Lee, H., and 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., and 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, 6640-6645.
- 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., and 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., and 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., and 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., and 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., *et al.* (2019). Conversion of *Escherichia coli* to Generate All Biomass Carbon from CO₂. *Cell* 179, 1255-1263 e1212.

Gonzalez, J.E., Bennett, R.K., Papoutsakis, E.T., and 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., and Rogers, E.J. (2013). Transformation of chemically competent *E. coli*. *Methods Enzymol* 529, 329-336.

He, H., Edlich-Muth, C., Lindner, S.N., and 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., and J.C., P. (2018). Synthetic Methylo-trophy: Past, Present, and Future. In *Methane Biocatalysis: Paving the Way to Sustainability*, Kalyuzhnaya M., and X. XH., eds. (Springer, Cham).

<http://enerkem.com/fr/>.

<http://www.methanol.org/>.

Huang, Y., Niu, B., Gao, Y., Fu, L., and 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., *et al.* (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., and 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., and Brautaset, T. (2013). Methylo-trophic *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., and 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., and Tannenbaum, S.E. (1968). Single-Cell protein. *Economic Botany* 22, 42-50.

Meyer, F., Keller, P., Hartl, J., Groninger, O.G., Kiefer, P., and Vorholt, J.A. (2018). Methanol-essential growth of *Escherichia coli*. *Nat Commun* 9, 1508.

Millard, P., Delepine, B., Guionnet, M., Heuillet, M., Bellvert, F., and 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., *et al.* (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., and 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., and 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., and 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 (2009). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).

Richter, G., Volk, R., Krieger, C., Lahm, H.W., Rothlisberger, U., and 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*, 4050-4056.

Ro, Y.T., Eom, C.Y., Song, T., Cho, J.W., and Kim, Y.M. (1997). Dihydroxyacetone synthase from a methanol-utilizing carboxydobacterium, *Acinetobacter* sp. strain JC1 DSM 3803. *J Bacteriol* *179*, 6041-6047.

Rohlhil, J., Sandoval, N.R., and Papoutsakis, E.T. (2017). Sort-Seq Approach to Engineering a Formaldehyde-Inducible Promoter for Dynamically Regulated *Escherichia coli* Growth on Methanol. *Acs Synth Biol* *6*, 1584-1595.

Roth, T.B., Woolston, B.M., Stephanopoulos, G., and Liu, D.R. (2019). Phage-Assisted Evolution of *Bacillus methanolicus* Methanol Dehydrogenase 2. *Acs Synth Biol* *8*, 796-806.

Russmayer, H., Buchetics, M., Gruber, C., Valli, M., Grillitsch, K., Modarres, G., Guerrasio, R., Klavins, K., Neubauer, S., Drexler, H., *et al.* (2015). Systems-level organization of yeast methylotrophic lifestyle. *BMC biology* *13*, 80.

Sangar, V., Blankenberg, D.J., Altman, N., and Lesk, A.M. (2007). Quantitative sequence-function relationships in proteins based on gene ontology. *BMC Bioinformatics* *8*, 294.

Schrader, J., Schilling, M., Holtmann, D., Sell, D., Filho, M.V., Marx, A., and Vorholt, J.A. (2009). Methanol-based industrial biotechnology: current status and future perspectives of methylotrophic bacteria. *Trends in biotechnology* *27*, 107-115.

Shleev, S.V., Shumakovich, G.P., Nikitina, O.V., Morozova, O.V., Pavlishko, H.M., Gayda, G.Z., and Gonchar, M.V. (2006). Purification and characterization of alcohol oxidase from a genetically constructed over-producing strain of the methylotrophic yeast *Hansenula polymorpha*. *Biochemistry (Mosc)* *71*, 245-250.

Silva-Rocha, R., Martinez-Garcia, E., Calles, B., Chavarria, M., Arce-Rodriguez, A., de Las Heras, A., Paez-Espino, A.D., Durante-Rodriguez, G., Kim, J., Nikel, P.I., *et al.* (2013). The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic acids research* *41*, D666-675.

Silva, F., Queiroz, J.A., and Domingues, F.C. (2012). Evaluating metabolic stress and plasmid stability in plasmid DNA production by *Escherichia coli*. *Biotechnol Adv* *30*, 691-708.

Simakov, D. (2017). *Renewable Synthetic Fuels and Chemicals from Carbon Dioxide* (Springer International Publishing).

Suzek, B.E., Wang, Y., Huang, H., McGarvey, P.B., Wu, C.H., and the UniProt, C. (2014). UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics* *31*, 926-932.

Team, R. (2015). RStudio: Integrated Development for R. (RStudio, Inc., Boston, MA).

Vieira, G., Carnicer, M., Portais, J.C., and Heux, S. (2014). FindPath: a Matlab solution for *in silico* design of synthetic metabolic pathways. *Bioinformatics* *30*, 2986-2988.

Wang, Y., Fan, L., Tuyishime, P., Zheng, P., and Sun, J. (2020). Synthetic Methylophony: A Practical Solution for Methanol-Based Biomanufacturing. *Trends in biotechnology*.

Whitaker, W.B., Jones, J.A., Bennett, R.K., Gonzalez, J.E., Vernacchio, V.R., Collins, S.M., Palmer, M.A., Schmidt, S., Antoniewicz, M.R., Koffas, M.A., *et al.* (2017). Engineering the biological conversion of methanol to specialty chemicals in *Escherichia coli*. *Metabolic engineering* 39, 49-59.

Whitaker, W.B., Sandoval, N.R., Bennett, R.K., Fast, A.G., and Papoutsakis, E.T. (2015). Synthetic methylophony: engineering the production of biofuels and chemicals based on the biology of aerobic methanol utilization. *Current opinion in biotechnology* 33, 165-175.

Windass, J.D., Worsley, M.J., Pioli, E.M., Pioli, D., Barth, P.T., Atherton, K.T., Dart, E.C., Byrom, D., Powell, K., and Senior, P.J. (1980). Improved conversion of methanol to single-cell protein by *Methylophilus methylophony*. *Nature* 287, 396-401.

Woolston, B.M., King, J.R., Reiter, M., Van Hove, B., and Stephanopoulos, G. (2018). Improving formaldehyde consumption drives methanol assimilation in engineered *E. coli*. *Nat Commun* 9, 2387.

Wu, T.Y., Chen, C.T., Liu, J.T., Bogorad, I.W., Damoiseaux, R., and Liao, J.C. (2016). Characterization and evolution of an activator-independent methanol dehydrogenase from *Cupriavidus necator* N-1. *Appl Microbiol Biotechnol* 100, 4969-4983.

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

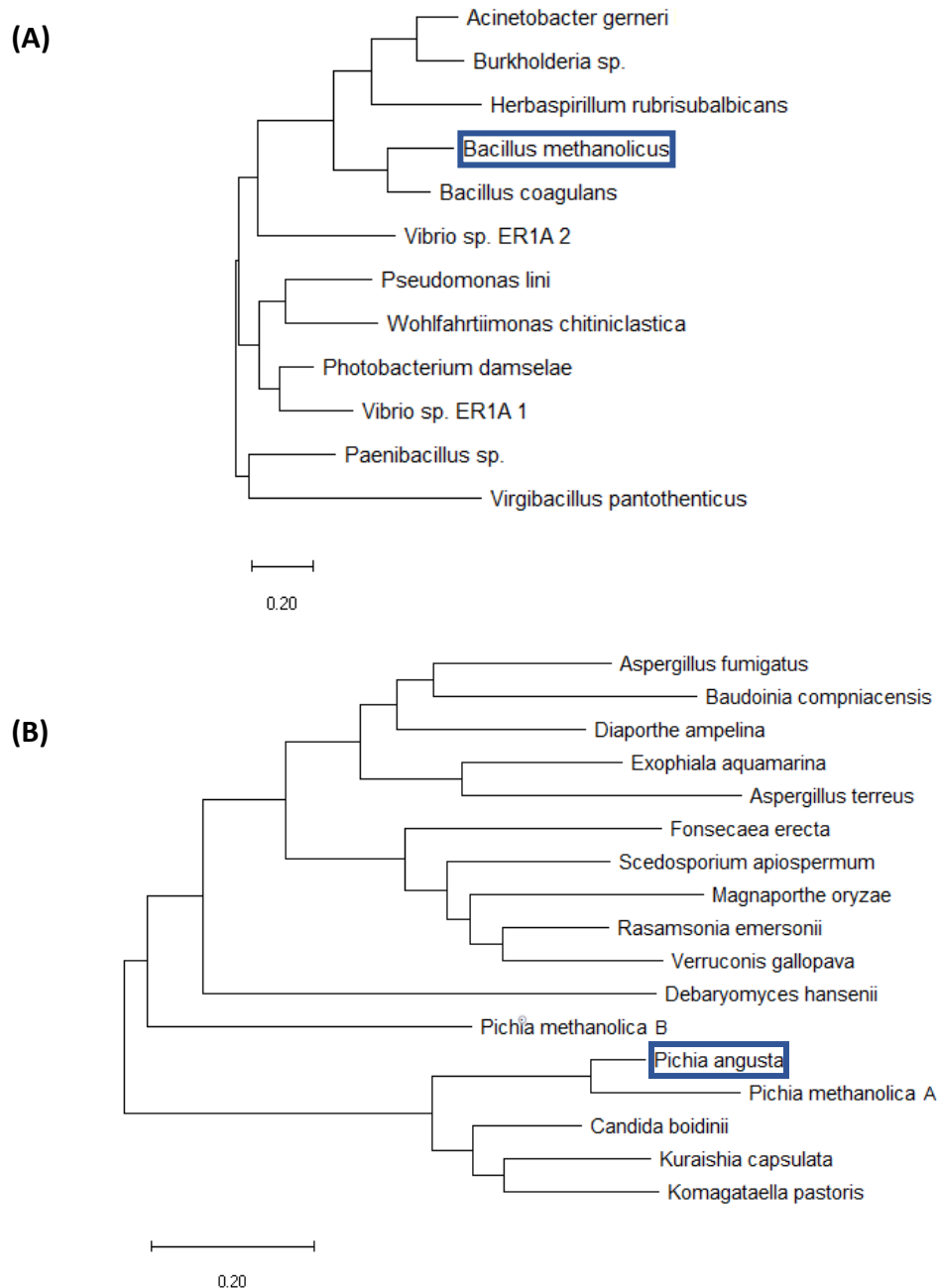


Figure S2: Unrooted phylogenetic trees of selected Mdh (A) and Das (B) homologues. The evolutionary history was inferred by using the Maximum Likelihood method and Le_Gascuel_2008 model [1]. The unrooted tree with the highest log likelihood (-6124.84) is shown. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Evolutionary analyses were conducted in MEGA X [2]. The query sequences are highlighted in blue. The sequences of *A. geroeri* Mdh and *Burkholderia* Mdh share 77.2% identity.

1. Le S.Q. and Gascuel O. (2008). An Improved General Amino Acid Replacement Matrix. *Mol Biol Evol* 25(7):1307-1320.

2. Kumar S., Stecher G., Li M., Knyaz C., and Tamura K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution* 35:1547-1549.

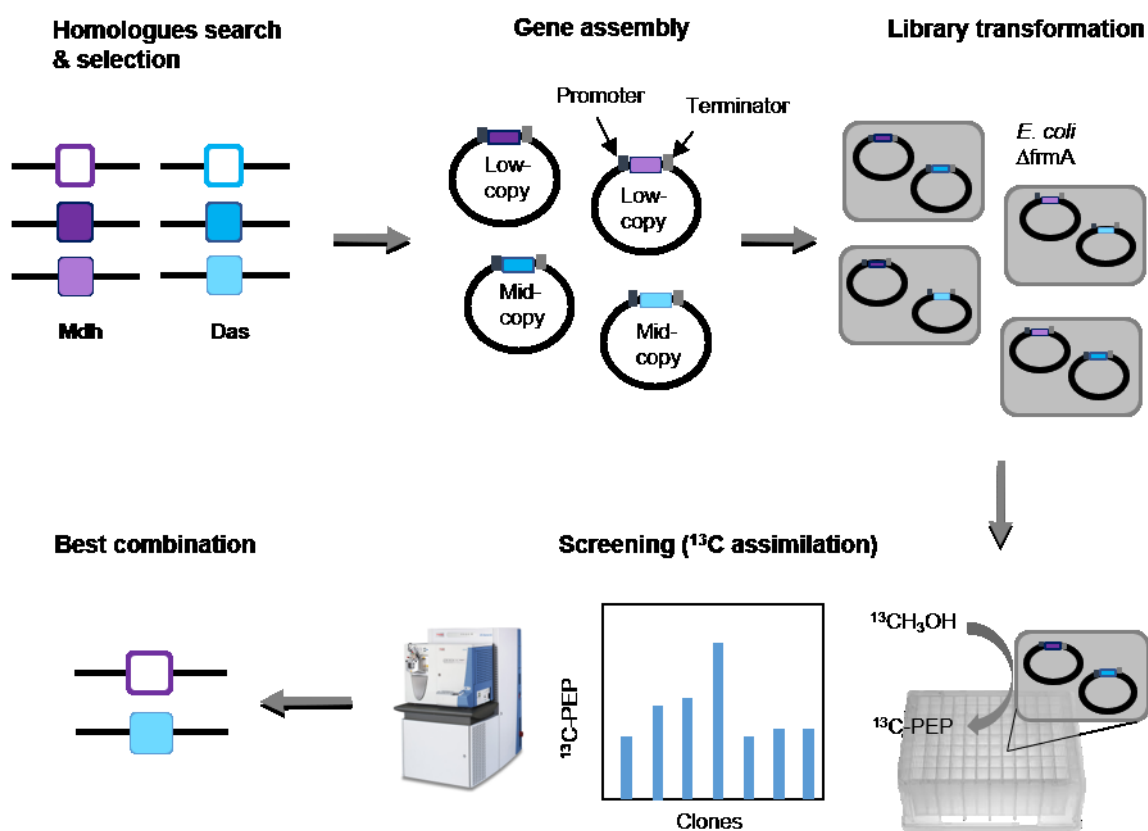


Figure S3: Overall scheme of the combinatorial assembly and screening of the synthetic pathway. Genes encoding Mdh and Das homologues are synthesized and cloned into two different plasmids between a promoter and terminator. A library is created by mixing the two arrays of plasmids and co-transforming each two-gene combination into the selected strain. The transformed strains are individually cultured in presence of ^{13}C -methanol and screened by measuring ^{13}C incorporation levels into PEP, a direct metabolite of the synthetic pathway. The best combination is identified based on the highest ^{13}C -incorporation level.

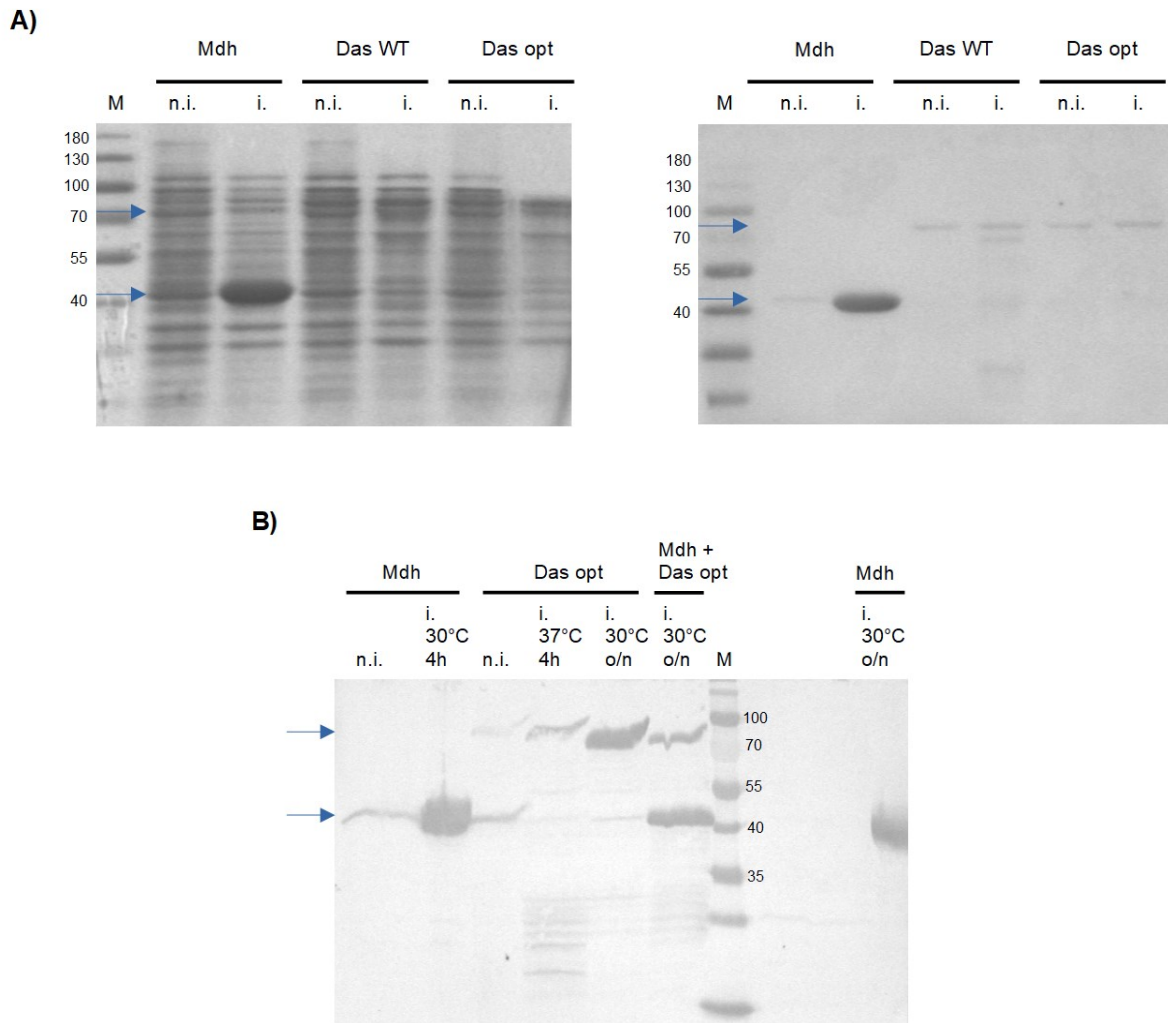
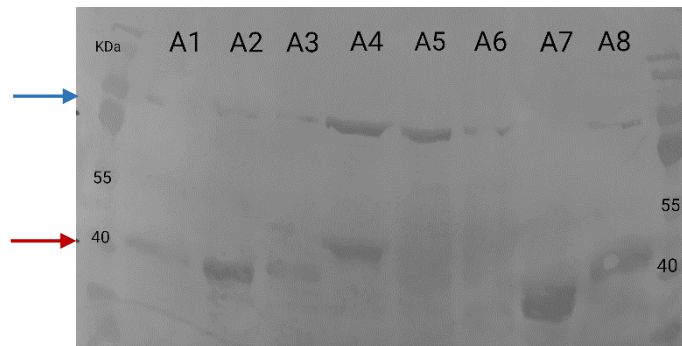


Figure S4: Expression analysis of *B. methanolicus* Mdh and *P. angusta* Das in different conditions. (A) Cell lysates corresponding to $OD_{600} = 1.2$ were separated on 12% SDS-PAGE gel and visualized by Coomassie staining (left) or Western blotted (right) using mouse anti-6×His antibody (1/10000) and anti-mouse AP-conjugated antibody (1/10000). Expression at 37°C for 4h after addition of 1 mM IPTG. M = molecular marker, n.i. = not induced, i. = induced. WT = wild-type, opt = codon-optimized. The expressed proteins are indicated by the blue arrows. Mdh = 40.8 kDa, Das = 78.7 kDa. (B) Western blot analysis of the expression at different incubation times and temperatures after IPTG induction (o/n = overnight).

(A)



(B)

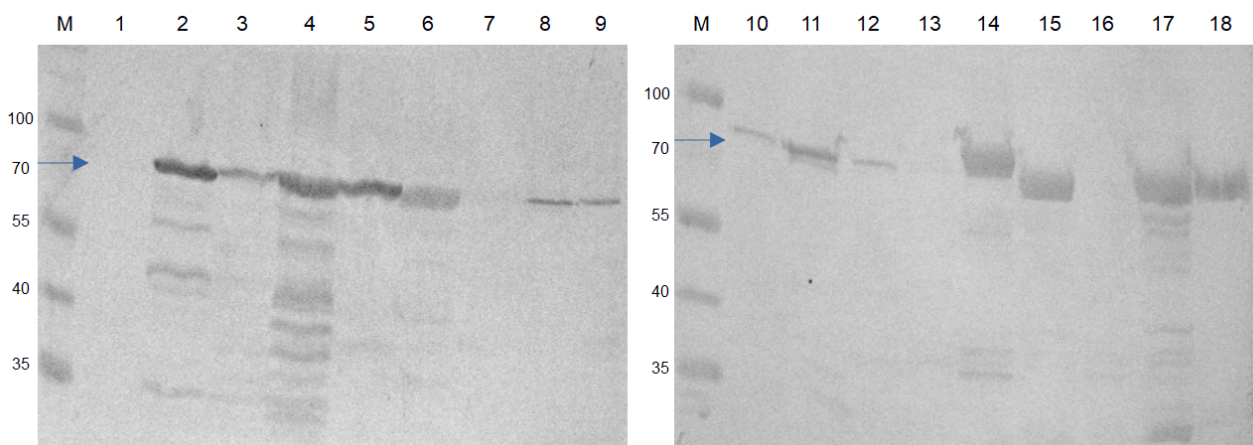


Figure S5: Western Blot analysis of expression of (A) Mdh and (B) Das homologues. Cell lysates corresponding to OD = 1.25 were loaded. Induction with 1 mM IPTG was carried at 30°C overnight. The expressed Das are indicated by a blue arrow whereas the expressed Mdh are indicated by a red arrow. A) Mdh expression. A1: *Bacillus methanolicus*, A2: *Pseudomonas lini*, A3: *Photobacterium damsela*, A4: *Bacillus coagulans*, A5: *Vibrio* sp. ER1A (1), A6: *Vibrio* sp. ER1A (2), A7: *Bacillus stearothermophilus*, A8: *Acinetobacter gernerii* DSM 14967. The co-expressed Das is from *P. angusta* (opt). B) Das expression. 1: *Aspergillus fumigatus*, 2: *Rasamsonia emersonii*, 3: *Exophiala aquamarina*, 4: *Diaporthe ampelina*, 5: *Aspergillus terreus*, 6: *Verruconis gallopava*, 7: *Kuraishia capsulata*, 8: *Fonseca erecta*, 9: *Baudoinia compniacensis*, 10: *Magnaporthe oryzae*, 11: *Scedosporium apiospermium*, 12: *Pichia methanolica* B (D7UPI3), 13: *Debaryomyces hansenii*, 14: *Mycobacterium* sp., 15: *Pichia angusta* (opt), 16: Dha reductase *Pichia pfancensis*, 17: *Pichia angusta* wild-type, 18: *Pichia angusta* wild-type not-induced. The codon-optimized version of Das (15) is more stable (no proteolytic degradation) compared to Das wild-type (17).

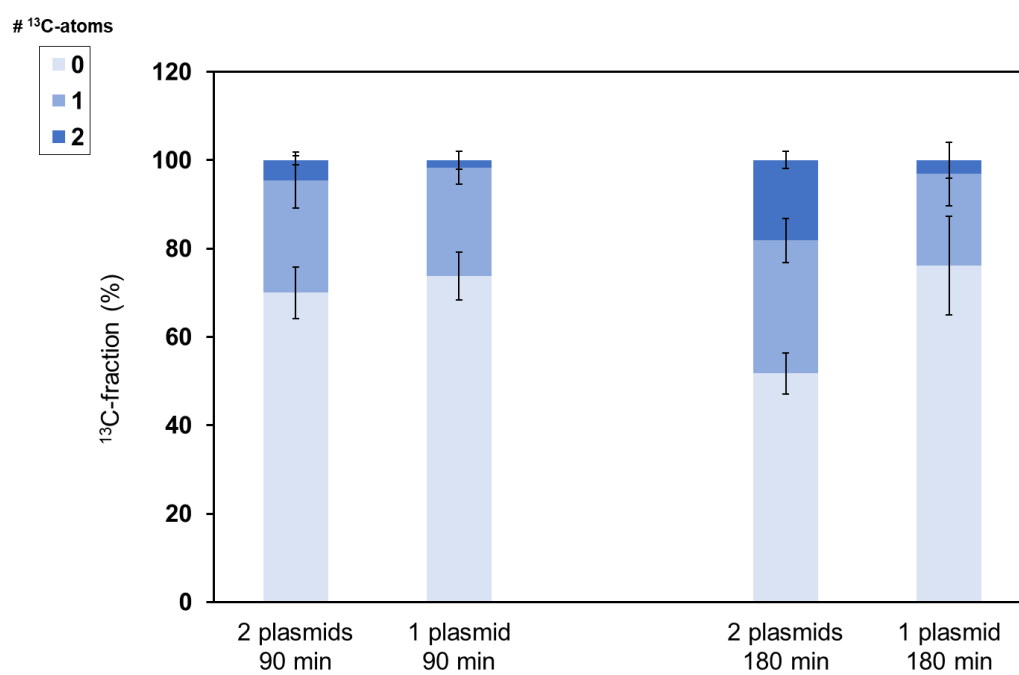


Figure S6: ^{13}C -Methanol assimilation in the methylotrophic *E. coli* $\Delta frmA$ expressing the synthetic pathway from one or two vectors. Labeling pattern of PEP after 90 min and 180 min of cultivation in M9 medium containing 655 mM ^{13}C -methanol. Data presented as mean \pm standard error of the mean (n = 3).

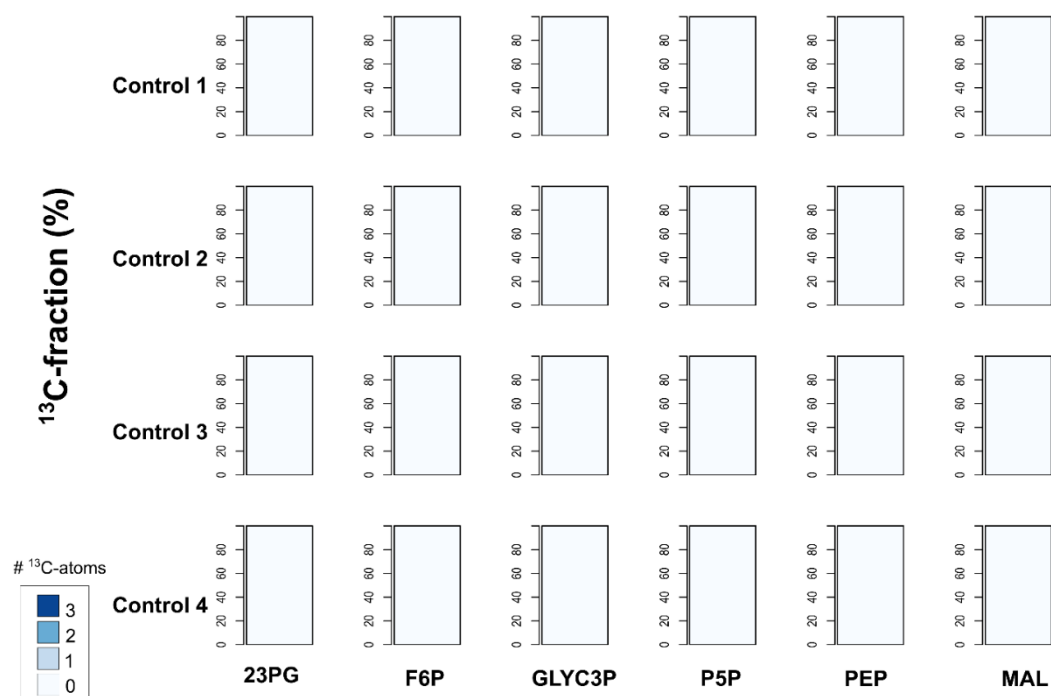


Figure S7: ^{13}C -Methanol assimilation into central metabolism intermediates in the control strains of the genealogy of methylotrophic *E. coli*. Labeling pattern of the intracellular metabolites 2 & 3 phosphoglycerate (23PG), fructose-6-phosphohate (F6P), glycerol-3-phosphate (GLYC3P), pool of pentoses-5-phosphate (P5P), phosphoenolpyruvate (PEP) and malate (Mal) within the different control strains containing an empty plasmid after 90 min of cultivation in M9 medium containing 655 mM ^{13}C -methanol.

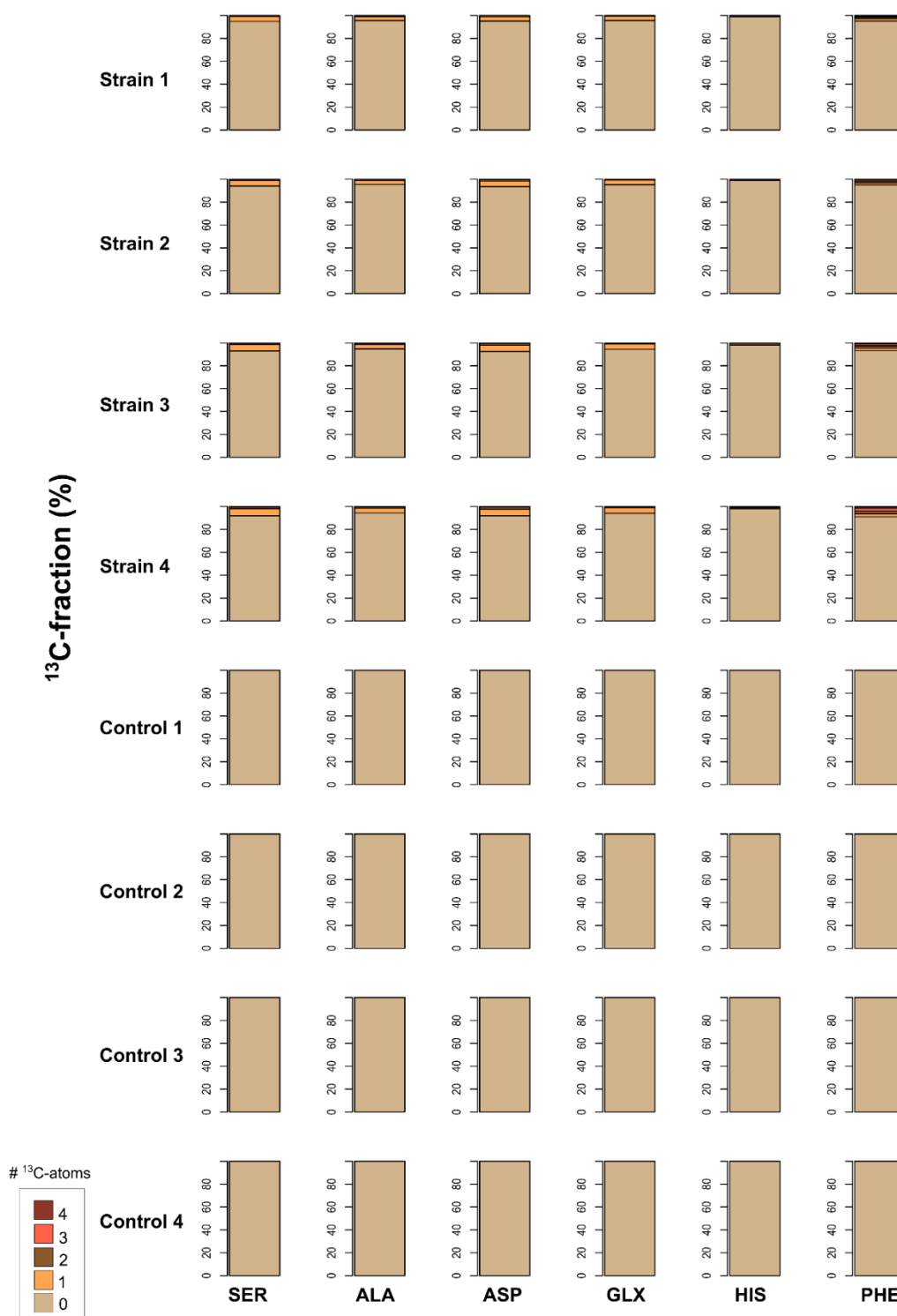


Figure S8: ^{13}C -Methanol assimilation into proteinogenic amino acids in the genealogy of methylotrophic *E. coli*. Labeling pattern of the proteinogenic amino acids serine (SER), alanine (ALA), asparagine (ASP), pool of glutamate and glutamine (GLX), histidine (HIS) and phenylalanine (PHE) within the different strains (see Figure 4B for details) and their respective controls containing an empty plasmid after 48h of cultivation in M9 medium containing 655 mM ^{13}C -methanol.

Plasmid, bacterial strain or primer	Description (genotype and/or relevant characteristics) or sequence of oligonucleotide primer	Reference or source
<i>E. coli</i> bacterial strains		
BW25113	rrnB3 ΔlacZ4787 hsdR514 Δ(araBAD)567 Δ(rhaBAD)568 rph-1	(Baba et al., 2006)
BW25113 ΔfrmA	ΔfrmA::frr	This study
BW25113 ΔptsA	ΔptsA::neo	This study
BW25113ΔfrmRAB	ΔfrmRAB::cat	This study
Plasmids		
pSEVA134	Medium-copy-number, lacIq/Ptrc promoter with lacO, pBBR1 ori, AmpR.	(Muller et al., 2015)
pSEVA424	Low-copy-number, lacIq/Ptrc promoter with lacO, pBBR1 ori, KmR.	(Silva-Rocha et al., 2013)
pSEVA424-Mdh	Derivative of pSEVA-424 containing <i>A. gernerii mdh</i> .	This study
pSEVA134-Das	Derivative of pSEVA-134 containing codon-optimized <i>P. angusta das</i>	This study
pSEVA424-Mdh-Das operon	Operon composed of <i>A. gernerii mdh</i> + codon-optimized <i>P. angusta das</i>	This study
pSEVA134-tkA	Derivative of pSEVA134 overexpressing <i>tkA</i>	This study
Primers		
P91	cactccgcttctggataatgtttttgcgccgac	This study
P92	ggtgatgCATATGTTTTCTCCTCCTAGGgtgtgGGTACC GGATCCTCACGACGCGGCCACTTACATTGCAGCTTC	This study
P83	cacacCCTAGGAGGAGGAAAAACATATGC	This study
P93	CGTAATTATTGGGGACCCCTGGATTCTCACC	This study
P102	aattcacaccctaggAGGAGGAAAAACATATGC	This study
P103	GGAGTCCAAGACTAGTTTACAGTTTGTTAACTTTATCA TGG	This study
P129	CGTGCGCTCTTATCCTGTAAACAGAAAACCGTACCGG GAGAGAGCAATGgtgtaggctggagctgcttc	This study
P130	GTCCAGATACAGTTCCATGATGTTTCCTTTTACAGTTC CAGTTCATGTTGcatatgaatatcctccttag	This study
P145	TAGAATACCCCTATAGTATATTGCATGCAGATGATG AGGTGCGAAatggtgtaggctggagctgcttc	This study
P146	GTAGGCCGGATAAGGCGTTCACGCCGCATCCGGCAG TCGTGCTATTacacatatgaatatcctccttag	This study

Table S1: Strains and plasmids used in this study

	<i>P. angusta</i> (Opt)	<i>P. angusta</i>	<i>P. pastoris</i>	<i>V. gallopava</i>	<i>C. boidinii</i>	<i>A. fumigatus</i> Z5	<i>E. aquamarina</i>	<i>O. methanolica</i> -A	<i>Mycobacterium</i> sp.
<i>Burkholderia</i> sp.	19,99	0,45	0,19	1	0,32	0	0,1	0,15	0
<i>A. gernerii</i>	22,02	15,96	1,74	1,67	0	0	0	0	0,3
<i>P. angusta</i>	0	0,01	0	0	0	0	0	0	0
<i>Vibrio</i> sp. B	0	0	0	0	0	0	0	0	0
<i>V. pantothenticus</i>	4,35	3,63	0,5	0,22	0	0	0	0	0
<i>Paenibacillus</i> sp	7	11,57	2,46	0,62	0,17	0	0	0,22	0,28
<i>H. rubrisubalbicans</i> M1	13	10,44	0,37	1,93	0,35	0	0	0	0,54
<i>B. methanolicus</i>	6	1,63	0,63	1,04	0,52	0,13	0,23	0,67	0,9
<i>W. chitinielastica</i>	9,47	3,06	1,4	1,07	1,99	0	0,35	0,2	0,45
<i>Vibrio</i> sp.A	9,83	1,93	0,32	0,42	0,14	0	0,09	0,05	0,29
<i>B. stearothermophilus</i>	11,43	2,71	0,69	0,47	0,43	0	0	0	0
<i>P. damselae</i>	13,34	2,1	1,81	1,41	0,49	0	0	0	0,29
<i>P. lini</i>	12,98	2,69	0,54	1,24	0	0	0	0	0,36
<i>B. coagulans</i>	12,62	3,53	1,43	1,91	0,08	0	0,1	0,2	0,41

	<i>Mycobacterium</i> sp.	<i>D. ampelina</i>	<i>B. compniacensis</i>	<i>A. terreus</i>	<i>M.oryzae</i>	<i>O. methanolica</i> -B	<i>D. hansenii</i>	<i>S. aptospermum</i>	<i>R. emersonii</i>	<i>F. erecta</i>	<i>K. capsulata</i>
<i>Burkholderia</i> sp.	0	0,36	0,28	0,57	0	0,31	0,2	1,93	3,5	1,78	1,98
<i>A. gernerii</i>	0,3	0,69	0,52	0,95	0,69	0,91	1,02	3,19	2,05	4,52	2,78
<i>P. angusta</i>	0	0	0	0	0	0	0	0	0	0	0
<i>Vibrio</i> sp. B	0	0	0	0	0	0,04	0	0	0	0	2,18
<i>V. pantothenticus</i>	0	0	0,01	0,11	0	0	0	0,17	1,03	0,64	0,66
<i>Paenibacillus</i> sp	0,28	0	0,02	0,42	0,13	0,23	0,23	0,93	1,79	2,16	2,62
<i>H. rubrisubalbicans</i> M1	0,54	0,43	0,27	0,93	1,13	1,5	1,21	2,19	3,76	3,22	3,15
<i>B. methanolicus</i>	0,9	0,58	0,99	0,25	0,8	0,71	0,97	2,44	0,93	3,64	2,99
<i>W. chitinielastica</i>	0,45	0,17	0,28	0,66	0,14	0,45	0,61	1,59	0,87	2,44	3,11
<i>Vibrio</i> sp.A	0,29	0,07	0	0	0,01	0,11	0,11	0,72	1,84	1,74	1,38
<i>B. stearothermophilus</i>	0	0,03	0,45	0,91	0,03	0,44	0,61	1,97	0,82	0,18	0,26
<i>P. damselae</i>	0,29	0,4	0,42	0,27	0,43	0,74	0,53	1,07	2,42	2,77	3,3
<i>P. lini</i>	0,36	0	0,34	0,19	0	0,38	0,42	1,39	0,25	1,92	2,98
<i>B. coagulans</i>	0,41	0,14	0,16	0,3	0,2	0,16	0,41	0,81	1,12	1,99	1,8

Table S2: ¹³C- enrichment of PEP in % using the combinatorial library. Labelling levels were measured 90 minutes after ¹³C pulse.

Mdh	Das	Mean Isotopic enrichment of PEP (%)
A. gernerii	Aspergillus fumigatus (OPT)	0.72+/-0.32
	Candida boidinii (OPT)	0.87 +/-0.28
	Debaryomyces hansenii (OPT)	0.95 +/- 0.20
	Kuraisha capsulata (OPT)	1.51 +/-0.09
	Pichia methanolica A (OPT)	0.50 +/-0.05
	Rasamsonia emersonii (OPT)	4.69+/-0.56

Table S3: ^{13}C -enrichment of PEP obtained using different Das enzymes having codon-optimized sequences in combination with Mdh from *A. gernerii*. Average and standard deviation of triplicates are given

Pathway	Cultivation conditions	Host	Metabolite	Time	¹³ C-enrichement (in %)	Reference
RuMP	1M methanol	<i>E. coli</i>	PEP	10 min	26,1	Muller et al., 2015
RuMP	60mM methanol	<i>E. coli</i>	PEP	48h	19,4	Whitaker et al., 2017
RuMP	60mM methanol + 1g/l yeast extract	<i>E. coli</i>	3 phosphoglycerate	48h	53,4	Whitaker et al., 2017
RuMP	60mM methanol + 0.5 g/l yeast extract	<i>E. coli</i>	PEP	48h	50	Bennett et al., 2018
RuMP + PPP	60mM methanol + 0.5 g/l yeast extract	<i>E. coli</i>	PEP	48h	59,2	Bennett et al., 2018
RuMP	188mM methanol + 30g/L glucose / anaerobic	<i>E. coli</i>	Glycerol 3 phosphate	48h	2	Zhang et al., 2018
RuMP	50mM methanol + 55mM glucose	<i>E. coli</i>	Glycerol 3 phosphate	25h	8	Wang et al., 2019
RuMP / auxotroph	500mM methanol + 5mM gluconate	<i>E. coli</i>	PEP	75h	21	Meyer et al., 2018
RuMP / auxotroph	250mM methanol + 50 mM xylose	<i>E. coli</i>	Butanol	4h	30	Chen et al., 2018
FLS pathway	250mM methanol + 1g/L yeast extract	<i>E. coli</i>	Phenylalanine	48h	8,7	Wang et al., 2017
SERINE based	200mM methanol + 30mM xylose	<i>E. coli</i>	Pyruvate	3h	9	Yu et Liao, 2019
SACA	250mM methanol + 1g/L yeast extract + 2g/L tryptone	<i>E. coli</i>	PEP	16h	18	Lu et al., 2019

Table S5: Overview of the ¹³C-enrichement obtained in different synthetic methylotrophic strains from previous studies.

Chapter 4

Encapsulation of a synthetic methylotrophic pathway in a bacterial microcompartment

Manuscript in preparation

Abstract:

Spatial organization in metabolic engineering offers great possibilities to improve the catalytic efficiency of a pathway by locally increasing the concentration of both enzymes and substrates. Empty bacterial microcompartments (eBMCs) are scaffolds that can be engineered to encapsulate heterologous enzymes in close proximity with their substrates. In this study, we constructed methylotrophic BMCs (mBMCs) by encapsulating a methanol dehydrogenase (Mdh) and a dihydroxyacetone synthase (Das) into an eBMC to enhance methanol conversion to dihydroxyacetone (DHA) and glyceraldehyde-3-phosphate (Ga3P). Both proteins were encapsulated within eBMCs using an encapsulation peptide, resulting in the formation of mBMCs observed by TEM. We showed that the addition of encapsulation peptides not only impacted enzymes activities but also promoted their aggregation. ¹H-NMR analysis revealed that methanol conversion to DHA and Ga3P was improved in purified mBMCs compared to non-encapsulated enzymes but was not optimal because unspecific reactions were taking place. Overall, mBMCs represent a promising solution to increase methanol conversion *in vivo* and thus might help to establish synthetic methylotrophy in *E. coli*.

Introduction

Methanol is becoming an attractive substrate for biorefineries as a new economy can be generated from its utilization (Olah, 2013; Schrader et al., 2009). Methylotrophs are organisms that can naturally use reduced one-carbon (C1) compounds as their sole carbon and energy source. However, the use of natural methylotrophs in the biotech industry is quite limited at the moment by our knowledge on their physiology and by the few genetic tools available to engineer them (Schrader et al., 2009; Delépine et al., 2020). Over the last years, to overcome this issue, an alternative has been to design synthetic methylotrophs by introducing methylotrophic pathways into non-native methylotrophic production hosts such as *Escherichia coli*.

Various attempts were done by introducing native methylotrophic pathways that relied either (i) on circular processes as the serine cycle (Yu and Liao, 2018), the ribulose monophosphate (RuMP) pathway (Chen et al., 2020), the xylulose monophosphate (XuMP) pathway (Dai et al., 2017) or a mix between the RuMP and XuMP pathways (De Simone et al., 2020) or (ii) on linear pathways such as the reductive glycine pathway (Kim et al., 2020) or the formolase pathway (Wang et al., 2017). Despite significant developments in the field, growth on methanol remains a challenge, which was recently won in *E. coli* by the implementation of a cyclic pathway allowing growth on pure methanol (Chen et al., 2020) or a linear pathway allowing growth on methanol and CO₂ (Kim et al., 2020). Different approaches as adaptive laboratory evolution (ALE) or rational engineering were used to optimize synthetic methylotrophs.

Spatial optimization is another approach that has not been extensively investigated yet. This approach can be of particular interest for the establishment of synthetic methylotrophy to improve methanol conversion as demonstrated in recent studies using multi-enzymes complexes of methylotrophic enzymes (Price et al., 2016; Fan et al., 2018). Another strategy commonly found in nature is to compartmentalize metabolic pathways. For example, in methylotrophic yeasts, the XuMP pathway is localized in the peroxisomes (Russmayer et al., 2015). This strategy was not only adopted by eukaryotic cells but also by some prokaryotic cells. By compartmentalizing metabolic pathways, cells protect themselves against the toxicity of reactive aldehyde intermediates (Held et al., 2013; Sampson and Bobik, 2008) while improving the pathway efficiency by encapsulating enzymes in close space with their

substrates (Chessher et al., 2015; Jakobson et al., 2017). Modelling studies reported that compartmentalization might increase up to 6-fold the catalytic activity compare to non-compartmentalized reactions (Conrado et al., 2007).

Unlike lipid-based eukaryotic organelles, bacterial microcompartments (BMCs) are semipermeable proteinaceous organelles encasing enzymes involved in specific metabolic process. A study revealed that microcompartments are found in 23 bacterial phyla (Axen et al., 2014). BMCs are large complexes of 100-600 MDa and have a diameter of 100-200 nm (Cheng et al., 2008; BOBIK et al., 1999). They are divided in two broad categories, carboxysomes involved in anabolism and metabolosomes involved in catabolism. Metabolosomes are involved in the degradation of substrates by first converting them into aldehyde and then detoxifying the aldehyde-intermediate into alcohol and acid derivatives. The best characterized metabolosomes are the 1,2-propanediol utilization (Pdu) BMC and the ethanolamine utilization (Eut) BMC, propanaldehyde and acetaldehyde being respectively the aldehyde intermediates of the two pathways (Chowdhury et al., 2014). Previous studies have demonstrated that it is possible to form empty BMCs (eBMCs) in *E. coli* by expressing seven of the eight shell proteins of the Pdu BMC from *Citrobacter freundii* (Parsons et al., 2010; Mayer et al., 2016). It was also demonstrated that non-native enzymes could be addressed in the eBMCs by fusing heterologous proteins with the N-terminal 18 amino acids of PduD (D18) or PduP (P18), named encapsulation peptides (Fan and Bobik, 2011; Fan et al., 2010). Several examples in the literature already show the variety of refunctionalization of eBMCs that have been achieved (Lawrence et al., 2014; Lee et al., 2016; Liang et al., 2017; Yung et al., 2017). The first proof of concept was the establishment of an “ethanol bioreactor” in *E. coli* by addressing a pyruvate decarboxylase and an alcohol dehydrogenase fused with encapsulation peptides (Lawrence et al., 2014). Co-localization of both enzymes in the eBMC resulted in the increase of ethanol production compared to non-encapsulated enzymes. Later, to produce 1,2-propanediol, a more complex pathway was encapsulated into eBMCs by addressing four enzymes (Lee et al., 2016).

We thus hypothesised that the compartmentalization of a methylotrophic pathway in an eBMC could help improving methanol conversion. Indeed, methanol oxidation to formaldehyde was shown to be a limiting step for methanol assimilation. This reaction is ensured by a methanol dehydrogenase (Mdh). However, Mdhs exhibit higher efficiency for

formaldehyde reduction than for methanol oxidation (Krog et al., 2013). Therefore, Mdh encapsulation in the BMC should favour methanol oxidation by locally increasing both substrate and enzyme concentrations. In addition, co-localization of Mdh with an enzyme using formaldehyde as a substrate, should be a driving force to enhance methanol conversion. Formaldehyde is a key intermediate of methylotrophic pathways but is also a toxic metabolite for the cells. Therefore, encapsulation of formaldehyde in the BMC would not only enhance the substrate channelling towards the assimilation process but would also protect cells from its toxicity by limiting formaldehyde presence in the cytoplasm. Moreover, by specifically targeting enzymes in BMCs, competitive enzymatic reactions should be avoided which should improve the overall efficiency of the methylotrophic pathway.

In this study, we constructed a methylotrophic BMC (mBMC) by encapsulating a synthetic methylotrophic pathway into a Pdu based eBMC to enhance methanol conversion. We choose the hybrid methylotrophic pathway composed by a NAD-dependent Mdh from *Acinetobacter gnereri* and a dihydroxyacetone synthase (Das) from *Pichia angusta* (De Simone et al., 2020). To construct the mBMC, Mdh and Das were fused with an encapsulation peptide to be addressed into eBMCs. The formation of the mBMCs in *E. coli* was observed by TEM. The *in vitro* activity of the purified mBMCs was analysed by $^1\text{H-NMR}$. The compartmentalization of the methylotrophic pathway offers a new solution to improve *in vivo* methanol assimilation in *E. coli*.

Material and methods

Bacterial strains and plasmids construction

Molecular biology was carried out either in *E. coli* JM109 or Top10 strains and protein production was done in *E. coli* BL21*DE3 strain (Supplementary Table 1). Plasmids were constructed in order to add the N-terminal His-tag with either D18 or P18 encapsulation peptide. Primers are listed in Supplementary Table 2. *Mdh* and *das* genes were amplified with flanking *NdeI* and *SpeI* restrictions sites and were ligated into pET14b, pET14b-D18 and pET14b-P18 vectors using *NdeI* and *SpeI* restriction sites. To assemble tagged and untagged Mdh and Das on the same plasmid, plasmids pCPE-1 and pCPE-2 (Supplementary Table 3) were constructed by a “Link and Lock” approach based on the compatible sticky ends formed by digestion with *XbaI* and *Clal* (McGoldrick et al., 2005).

Overexpression, purification of recombinant protein and SDS-PAGE

BL21*DE3 competent cells were transformed with either the pET14b-Mdh, pET14b-D18-Mdh, pET14b-P18-Mdh, pET14b-Das, pET14b-D18-Das or pET14b-P18-Das to produce the untagged or tagged Mdh and Das. A colony was cultured in 5mL of LB overnight at 37°C with shaking. 3mL of the starter culture were used to inoculate 300mL of LB supplemented with ampicillin (100mg/L). For D18-Das and P18-Das, 2XYT media was used. Cultures were grown at 37°C with shaking until OD₆₀₀ 0.6-0.8 was reached, protein production was induced by the addition of 0.4mM of isopropyl- β -thiogalactopyranoside (IPTG). The cultures were then incubated overnight at 19°C with shaking. Cells were harvested by centrifugation at 4000 rpm for 15 minutes at 4°C; pellet was resuspended in Lysis Buffer (0.1 M PBS pH 7.4, 300mM NaCl 10mM imidazole) along with 100 μ M phenylmethylsulfonyl fluoride (PMSF). Cells were lysed by sonication and cell debris was removed by centrifugation at 45,000g for 30 minutes at 4°C. Recombinant protein was purified from the soluble fraction using a Co-NTA column by immobilized metal ion affinity chromatography (IMAC) and eluted using Lysis Buffer supplemented with 300mM imidazole. Imidazole was removed using PD-10 desalting columns (GE Healthcare) stabilized with 0.1 M PBS pH 7.4 using manufacturer’s instructions. Mdh concentration was measured using a nanodrop, molecular weight and extinction coefficient settings were calculated using Expasy Prot Param tool (expasy.org/tools/protparam.html) (Gasteiger, 2003). Das concentration was measured using Pierce BCA protein assay kit (Sigma) using bovine serum albumin (BSA) as a standard. Purified proteins were analysed on 12% Mini-

PROTEAN TGX gel (Bio-rad, Hercules, CA, USA) and gel was stained using InstantBlue Protein Stain (Expedeon).

Activity assays

1. Methanol dehydrogenase (Mdh)

The activity of Mdh for the oxidation of methanol to formaldehyde was measured by following the initial rate of reduction of NAD⁺ to NADH at 340 nm. Activity assays were carried out at 30 °C in 0.4 mL reactions containing 0.1 M PBS pH 7.4, 1.5 M methanol, 3 mM NAD⁺, 20 mM MgCl₂ and 50 μM of Mdh.

2. Dihydroxyacetone synthase (Das)

The activity of Das for the conversion of formaldehyde and xylulose-5-phosphate to dihydroxyacetone and glyceraldehyde-3-phosphate was measured in a coupled reaction with glyceraldehyde-3-phosphate dehydrogenase (gldA) by following the oxidation of NADH to NAD⁺ at 340 nm. Activity assays were carried out 30 °C in 0.18 mL containing 0.1 M PBS pH 7.4, 2.5 mM formaldehyde, 2.5 mM xylulose-5-phosphate, 2.5 mM NADH, 0.1 mM thiamine pyrophosphate (TPP), 20 mM MgCl₂, 1 μM of gldA and 20 μM of Das.

TEM analysis

BL21*DE3 competent cells were co-transformed with pLysS-PduABJKNU and either the pET14b-Mdh, pET14b-D18-Mdh, pET14b-P18-Mdh, pET14b-Das, pET14b-D18-Das, pET14b-P18-Das or pCPE-1 to express the eBMC and the untagged or tagged Mdh and Das. A colony was cultured in 5 mL of LB overnight at 37°C with shaking. 0.5 mL of the starter culture were used to inoculate 50 mL of LB supplemented with chloramphenicol (37 mg/L) and ampicillin (100 mg/L). Cultures were grown at 37°C with shaking until OD₆₀₀ 0.6-0.8 was reached, protein production was induced by the addition of 0.4 mM of isopropyl-β-thiogalactopyranoside (IPTG). The cultures were then incubated overnight at 19°C with shaking. Cells were harvested by centrifugation at 3000 g for 10 minutes. The cell pellet was resuspended in 2 ml 2.5% Glutaraldehyde in 100 mM sodium cacodylate pH 7.2 and fixed for 2 hours with gentle rotating. Cells were pelleted by centrifugation at 6000 g for 2 minutes and were washed twice with 100 mM sodium cacodylate. Cells were post-fixed with 1% osmium tetroxide in 100 mM sodium cacodylate for 2 hours and subsequently washed twice with dH₂O. Cells were dehydrated by incubation in an ethanol gradient, 50% EtOH for 10 minutes, 70% EtOH

overnight, 90% EtOH for 10 minutes followed by three washes in 100% EtOH. Cells were then washed twice with propylene oxide for 15 minutes. Cell pellets were embedded by resuspension in 1 ml of a 1:1 mix of propylene oxide and Agar LV Resin and incubated for 30 minutes with rotation. Cell pellet was infiltrated twice in 100% Agar LV resin. The cell pellet was resuspended in fresh resin and transferred to a 1ml BEEM embedding capsule, centrifuged for 5 minutes at 3000 x g to concentrate the cells to the tip of the mould and incubated for 20 hours at 60 °C to polymerise.

Samples were ultra-thin sectioned on an RMC MT-XL ultramicrotome with a diamond knife (diatome 45°) sections (60-70 nm thick) were collected on uncoated 300-mesh copper grids. Grids were stained in 4.5% uranyl acetate in 1% acetic acid solution followed by two washes in dH₂O. Grids were then stained with Reynolds lead citrate for 3 minutes followed by a wash in ddH₂O. Electron Microscopy was performed using a JEOL-1230 transmission electronic microscope equipped with a Gatan multiscan digital camera at an accelerating voltage of 80 kV.

Purification of BMCs

BL21*DE3 competent cells were co-transformed with pLysS-PduABJKNU and pCPE-1 or pCPE-2. A colony was cultured in 5mL of LB overnight at 37°C with shaking. 0.5mL of the starter culture were used to inoculate 50mL of LB supplemented with chloramphenicol (37 mg/L) and ampicillin (100 mg/L). Cultures were grown at 37°C with shaking until OD₆₀₀ 0.6-0.8 was reached, protein production was induced by the addition of 0.4mM of isopropyl-β-thiogalactopyranoside (IPTG). The cultures were then incubated overnight at 19°C with shaking. Cells were harvested by centrifugation at 4,000 rpm for 10 minutes at 4 °C. Pellet was resuspended in 30 mL of Yeast Protein Extraction Reagent Plus (YPER Plus, Thermo Scientific) supplemented with 500 units of Benzonase Nuclease (Sigma) and 100μM PMSF and incubated at room temperature for 3 hours with gentle shaking. Cell lysate was pelleted by centrifugation at 11,300 g for 8 minutes at 4 °C. Pellet was resuspended in 2 mL of 20 mM Tris HCl pH 8.0, 20mM NaCl. Supernatant was centrifuged for 5 minutes at 11,000 g at 4 °C. The supernatant was collected, NaCl concentration was increased to 80 mM with 3 M NaCl, and was then centrifuged at 4°C for 15 minutes at 11,000 g. The pellet was resuspended in 1 mL of 20 mM Tris HCl pH 8.0 and was centrifuged in an Amicon Ultra-4 MWCO 100 kDa for 3

minutes at 2,000 g. Final fraction was quantified by Pierce BCA protein assay kit (Sigma) using BSA as a standard prior to do the enzymatic assay.

¹H-NMR analysis

All NMR analysis were performed on an Avance III 800-MHz spectrometer (Bruker, Rheinstetten, Germany) equipped with a 5-mm QPCI cryogenic probe head except for the Mdh and Das combination experiment that was performed on an Avance III 500-MHz (Bruker, Rheinstetten, Germany) equipped with cryogenic probe head .

1. Das ¹³C-labelling experiment

Conditions assay were set as previously described using ¹³C-formaldehyde and without addition of gldA in the reaction mix. After 2 hours, spectra were acquired by quantitative ¹H one dimensional NMR at 280K with water presaturation using zgpr30 sequence. 64 scans were accumulated after 4 dummy scans. Using Fourier transform, the time domain function (FID) was converted to the frequency domain function (spectrum). The spectra phase was manually adjusted, baseline was automatically corrected and spectrum was aligned using the internal standard 3-methylsilylpropionic-2,2,3,3-d4 acid sodium salt (TSP-d4) signal with the Bruker software TopSpin (v4.0.5).

2. Mdh and Das combination

Enzymatic assays were carried out at 30°C in 1mL containing 0.1 M PBS pH 7.4, 500 mM methanol, 2.5 mM xylulose-5-phosphate, 2.5 mM NAD⁺, 0.1 mM TPP, 5 mM MgSO₄, 40 μM of Mdh and 10 μM of Das. Spectra were acquired at 286 K by quantitative ¹H one dimensional NMR with water presaturation using zgpr30 sequence. 16 scans were accumulated after 4 dummy scans. Using Fourier transform, the time domain function (FID) was converted to the frequency domain function (spectrum). The spectra phase and baseline were corrected manually and spectrum was aligned using the internal standard 3-methylsilylpropionic-2,2,3,3-d4 acid sodium salt (TSP-d4) signal with the Bruker software TopSpin (v4.0.5).

3. Methylophilic BMCs

Enzymatic assays were carried out at 30°C with shaking, in 500 μL containing 20 mM TrisHCl pH 8.0, 500 mM methanol, 2.5 mM xylulose-5-phosphate, 2.5 mM NAD⁺, 0.1 mM TPP, 1 mM MgSO₄, 1-2.5 mg/mL of purified BMC. A 3 mm NMR tube was filled with 200 μL of the reaction mix and placed in a 5 mm NMR tube filled with 300 μL of TSP-d4. Spectra were acquired at 303 K by quantitative ¹H one dimensional NMR with water presaturation using

zgpr30 sequence. 16 scans were accumulated after 4 dummy scans. Using Fourier transform, the time domain function (FID) was converted to the frequency domain function (spectrum). The spectra phase and baseline were corrected manually and spectrum was aligned using the internal standard TSP-d4 signal with the Bruker software TopSpin (v4.0.5).

Results and discussion

Expression, purification and enzymatic assay of Mdh from *A. gernerii* and Das from *P. angusta*

Mdh from *A. gernerii* and Das from *P. angusta* have never been studied, we thus measured the enzymatic activities of the purified Mdh and Das. The *mdh* and *das* encoding genes were cloned and expressed in the His-tag plasmid pET14b in *E. coli*, then proteins were purified. SDS-PAGE analysis showed that the purified Mdh and Das migrated at their expected size, of approximately 43 kDa and 80.9 kDa respectively (Supplementary Figure S1). Mdh activity was followed by monitoring NADH appearance (Figure 1A) and Das was coupled with a glycerol dehydrogenase (GldA) to monitor NADH disappearance (Figure 1B). Dihydroxyacetone (DHA) is produced by Das and can be reduced to glycerol by GldA using NADH as a cofactor. GldA reduction activity was tested before monitoring Das activity (data not shown). Both enzymes were found to be active in the tested condition (Figure 1C). In our “*in vivo*”-like conditions (i.e. pH 7.4 at 30°C), Mdh specific activity to convert methanol into formaldehyde was low (0.03 $\mu\text{mol}/\text{min}/\text{mg}$) compared to the specific activities of others Mdhs (Krog et al., 2013). *A. gernerii* is a mesophilic organism whose optimal growth temperature is 30°C. Therefore, the chosen temperature for the enzymatic assay should not affect much Mdh activity. However, it was shown that alcohol oxidation by Mdhs was more favourable in alkaline conditions and that the catalytic activity was higher for long chain alcohol (Krog et al., 2013). Moreover, formaldehyde reduction was also shown to be more favourable than methanol oxidation. The specific activity of the Das from *P. angusta* was of 0.585 $\mu\text{mol}/\text{min}/\text{mg}$ and was closed to the one found for the partially purified Das from *Candida boidinii* in similar conditions (0.766 $\mu\text{mol}/\text{min}/\text{mg}$) (Waites and Quayle, 1983).

C)

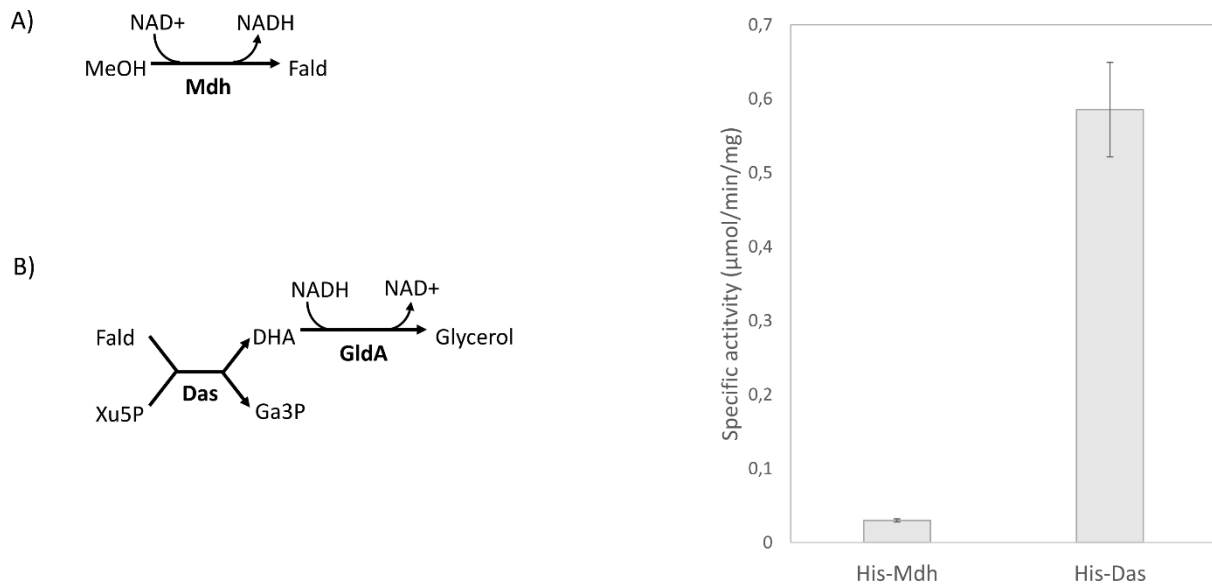


Figure 1: Enzymatic reactions of Mdh and Das. (A) methanol oxidation by Mdh and (B) condensation of formaldehyde with xylulose-5-phosphate by Das coupled with the dihydroxyacetone reduction by glycerol dehydrogenase. (C) Specific activities of Mdh from *A. gernerii* and the codon-optimised Das from *P. angusta*. Values averaged from three replicates, bars show SD.

Dihydroxyacetone synthase (Das), glycerol dehydrogenase (gldA), methanol dehydrogenase (Mdh). Dihydroxyacetone (DHA), formaldehyde (Fald), glyceraldehyde-3-phosphate (GA3P), methanol (MeOH), xylulose-5-phosphate (Xu5P).

Das is also named formaldehyde transketolase. Therefore, to give clear evidence of Das transketolase activity with Xu5P and formaldehyde, we used ^{13}C -formaldehyde combined with ^{12}C -Xu5P and analysed the ^{13}C -transition in the reaction products by ^1H -NMR. As expected, DHA and Ga3P were produced. In aqueous solution, Ga3P and DHA exist respectively under two forms. DHA is found at 78% under its native form and at 22% under its hydrated form (i.e. propane-1,2,2,3-tetrol) (Peiro et al., 2019). Similarly, in aqueous solution, an equilibrium between the diol and aldehyde forms of Ga3P exists, respectively represented at 97% and 3% (Trentham et al., 1969). ^{13}C satellites were present only around DHA peak at 4.42 ppm and its diol derivate at 3.58 ppm (Figure 2). This confirms that DHA is originating from formaldehyde and the glycolic aldehyde group of Xu5P. DHA being a symmetrical molecule of three carbons, we cannot distinguish which atom is labelled between the C1 and C3. The reaction was

complete as the expected amount of DHA and Ga3P were found in an equimolar ratio (Supplementary Figure S2). This assay is the first example showing Das transketolase activity *in vitro* with its native substrates (i.e. Xu5P and formaldehyde). Our data are consistent with a previous study demonstrating the formation ^{13}C -DHA from ^{13}C -methanol and hydroxypyruvate using of a combination of a Das with an alcohol oxidase from *C. boidinii* (Yanase et al., 1995).

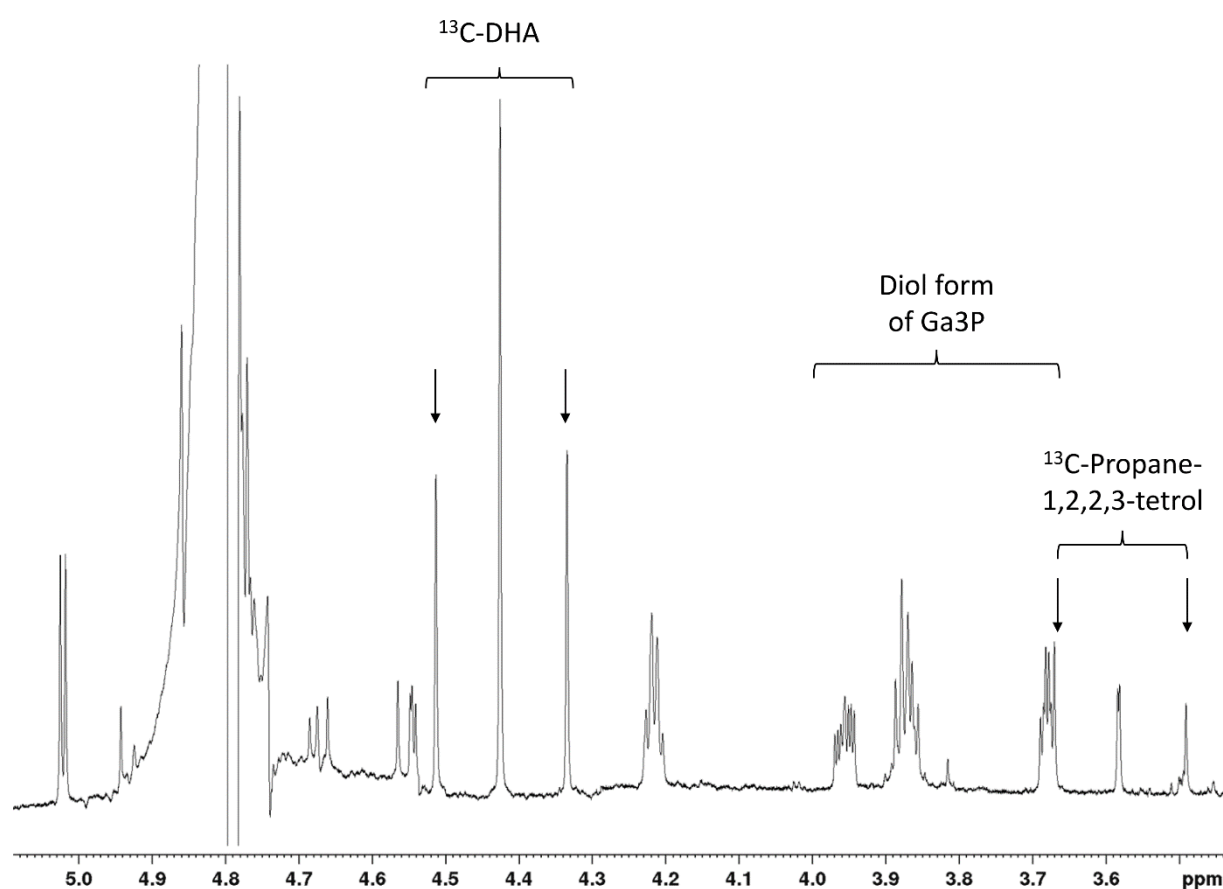


Figure 2: Products from ^{13}C -formaldehyde and xylulose-5-phosphate rearrangement by Das analyzed by ^1H -NMR. ^{13}C satellites are indicated by arrows.

Dihydroxyacetone (DHA), Glyceraldehyde-3-phosphate (Ga3P).

Effect of encapsulation peptides on the activity of Mdh and Das

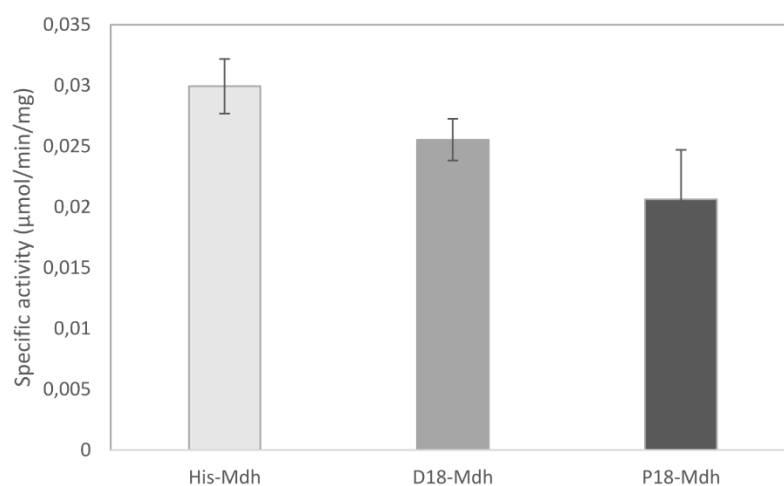
It was previously shown that the addition of the D18- or P18- encapsulation peptides was required to target heterologous enzymes into recombinantly produced eBMCs. (Lawrence et al., 2014; Lee et al., 2016). However, the addition of the encapsulation peptides was shown to

affect the enzyme activity (Lee et al., 2016; Juodeikis et al., 2020). It is therefore determinant to choose carefully the encapsulation peptide for the enzyme to be encapsulated in the BMCs. First, Mdh and Das were cloned separately with D18 and P18 N-terminal encapsulation peptide followed by a His-tag. Once expressed, proteins were purified by IMAC and the enzymatic activity was compared to the protein harbouring only the His-tag, now referred as “untagged” enzyme.

Encapsulation peptides were found to have a highly variable effect on the specific activities of the enzymes (Figure 3). During methanol oxidation, the addition of the D18 encapsulation peptide resulted in a decrease of 15% of the Mdh specific activity in comparison to the untagged control. Tagging Mdh with the P18 encapsulation peptide reduced the activity by 31% (Figure 3A). However, fusing Das with the D18 encapsulation peptide had a dramatic effect by reducing Das activity by 73%. Surprisingly, the P18 encapsulation peptide increased Das activity by 178% (Figure 3B).

Until now, the addition of encapsulation peptides was shown to either decrease or to have few impact on the enzymatic activity (Lee et al., 2016). It is therefore consistent with what we observed with the tagged-Mdh and the D18-Das. However, P18-Das is new example among to exhibiting an improved activity by the addition of an encapsulation peptide (Juodeikis et al., 2020). Even if the D18 and P18 encapsulation peptides are predicted to be structurally similar, the two tags do not have the same influence on the enzymes activity. Predict the effect of an encapsulation peptide on the enzymatic activity is not possible at the moment as shown by (Juodeikis et al., 2020). From this experiment, the D18 encapsulation peptide seemed to be a good candidate to fuse with Mdh while Das should be fused with the P18.

A)



B)

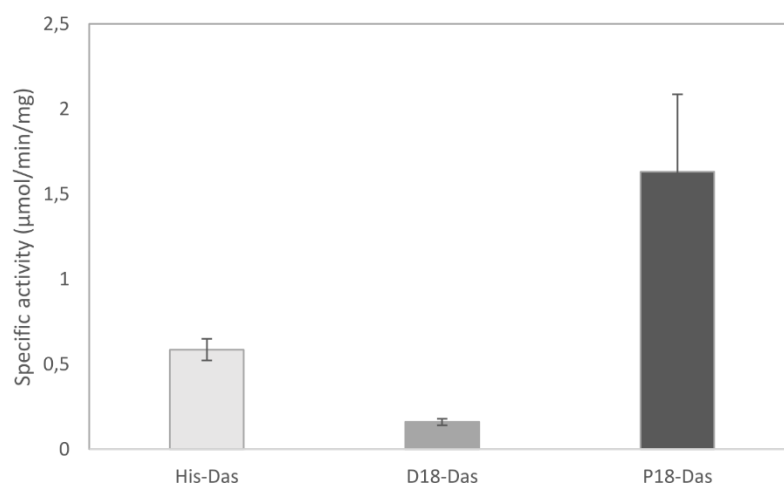


Figure 3: The effect of the encapsulation peptides on the specific activities of the (A) Mdh from *A. gernerii* and of (B) the Das from the codon-optimised *P. angusta*. Values averaged from three replicates, bars show SD.

Demonstration of the activity of Mdh and Das combination

Once the activity of the single enzymes determined, we decided to verify the *in vitro* operation of Mdh and Das combination. Based on the specific activities determined previously, we tested the production of DHA using the combination of the best performing enzymes (i.e. His-Mdh and P18-Das) in the same conditions. We followed methanol and Xu5P conversion by $^1\text{H-NMR}$. After 7 hours, a singlet at 4.4 ppm corresponding to DHA chemical shift was produced and DHA identification was confirmed thanks to a spike with a standard of DHA (data not shown). The reaction was not complete since only $625\mu\text{M}$ of DHA from 2.5mM

of Xu5P was produced. However, the reaction was specific since only DHA was produced and no other peaks were found on the NMR spectra. When Mdh and Das were tested alone (Figure 1), the reaction was complete after 50 minutes and 10 minutes respectively. Here, the reaction is not complete but three times less methanol was used and a ratio of 4:1 of Mdh and Das was used to test the combination. In native and in synthetic methylotrophy, the impact of this ratio on the methanol assimilation has not been explored yet. However, few studies investigated the *in vitro* combination of Mdh with two other methylotrophic enzymes by constructing multienzyme complexes (Price et al., 2016; Fan et al., 2018). Price et al. showed that by increasing the molar ratio of the formaldehyde-condensing enzyme, methanol conversion was improved (Price et al., 2016). Further exploration could help to determine if a complete reaction of methanol and Xu5P could be achieved by modulating the Mdh and Das ratio *in vitro*.

Effect of encapsulation peptides on Mdh and Das aggregation

The addition of encapsulation peptides is known to promote protein aggregation which allows recruiting the tagged proteins into BMCs (Kerfeld and Erbilgin, 2015). We thus decided to study the effect of D18 and P18 encapsulation peptides on the level of aggregation of Mdh and Das.

BL21*DE3 competent cells expressing the tagged Mdh and Das were visualized by TEM (Figure 4). Protein aggregates were observed in cells expressing the tagged Mdh and Das (Figure 4, middle and lower panel) regardless of the nature of the encapsulation peptides (i.e. D18 or P18). As expected, these aggregates were not observed when the untagged Mdh was produced. Surprisingly, the untagged Das was found to form aggregates (Figure 4, higher panel). During the purification process, even if some untagged Das was found in insoluble fraction, the enzyme was mainly found in the soluble fraction. In contrast, D18- and P18-Das were mainly found in the insoluble fraction during the purification process (data not shown). This shows that the addition of the encapsulation peptides still modified the proteins solubility. Natively, Das is found in the peroxisome of yeasts. In *C. boidinii*, when the *pmp47* was deleted, a gene encoding for a putative peroxisomal transporter, Das was only found in the cytosol and formed aggregates (Sakai et al., 1996; van der Klei et al., 2006). In our conditions, this suggests that the His-tag addition to Das did not promote its aggregation in

cells but it is rather its localization in the cytoplasm. Indeed, peroxisome pH was found to be alkaline in some eukaryotic cells (Dansen et al., 2000; Shen et al., 2013). Protein solubility can be affected by the pH as it modifies the charge state of proteins. At a pH equals to the pI of the protein, the protein is the least soluble. By using ExPASy Prot Param tool, we determined that the theoretical pI of the untagged Das is 6.70. Therefore, as cytosol pH is 7.4, the aggregation of untagged Das might be promoted. Another explanation would be that untagged Das aggregation would come from its native peroxisome targeting signal (PTS) PTS1. PTS1 is a C-terminus sequence allowing proteins to cross the peroxisome membrane (van der Klei et al., 2006). However, no role of PTS1 in protein aggregation has been described yet. Overall, our data showed that both D18 and P18 encapsulation peptides promoted aggregation of both Mdh and Das. However, between the two encapsulation peptides, no difference on the protein aggregation-level was found.

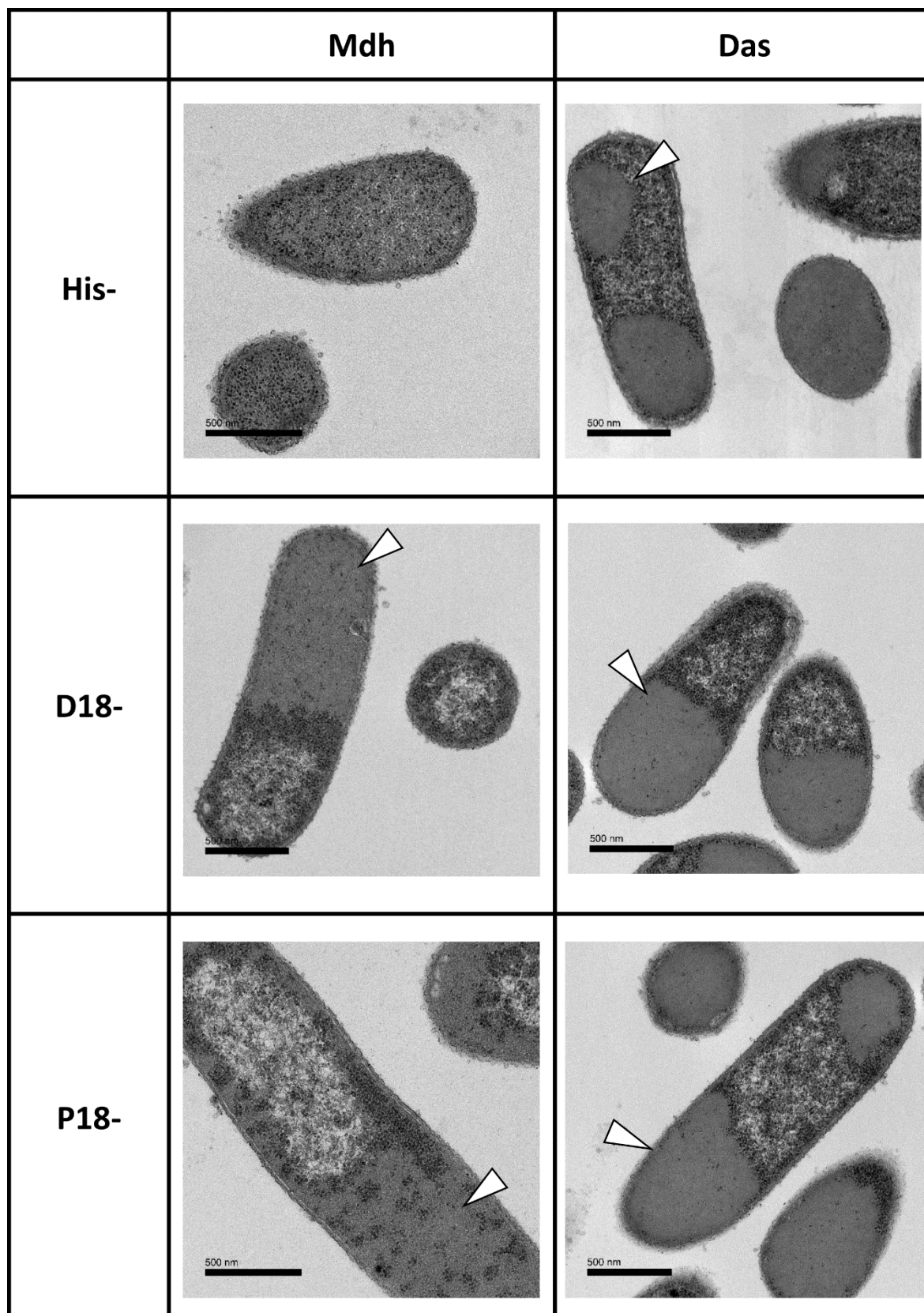


Figure 4: TEM analysis of *E. coli* strains producing tagged (D18 and P18) and untagged (His) Mdh and Das. Arrows indicate protein aggregate. Scale bar shows 0.5 μm .

Effect of encapsulation peptides on Mdh and Das recruitment into eBMCs

Several studies demonstrated that heterologous proteins could be addressed into eBMCs by adding encapsulation peptides (Lawrence et al., 2014; Lee et al., 2016; Yung et al., 2017; Juodeikis et al., 2020). However, the question of a possible difference of protein recruitment into eBMCs due to the nature of the encapsulation peptides had been addressed only recently (Juodeikis et al., 2020). We thus decided to determine the influence of the encapsulation peptides on the level of encapsulation of Mdh and Das into eBMCs.

BL21*DE3 competent cells were transformed with genes encoding for each version of the tagged-enzyme and for the eBMC. Gene expression was verified by SDS-PAGE (Supplementary Figure S3). Tagged enzymes and BMCs were co-produced (Supplementary Figure S3A) and as expected the production level of Mdh and Das decreased due to the eBMCs production compared to the production of Mdh or Das alone (Supplementary Figure S3B). Next, strains expressing the eBMC with each version of the tagged-enzyme were analysed by TEM (Figures 5A). We defined BMCs as structures where angular facets could be seen containing different levels of electron density depending on their filling (Lee et al., 2018). The electron density found within the eBMCs represents *E. coli* endogenous proteins unspecifically encapsulated within the eBMCs. However, when tagged Mdh and Das were overexpressed together with the eBMC, higher density was found in some of the BMCs. Based on these observations, for each condition, one counting of at least 600 purified BMCs was performed on randomized samples. When a high electron density was found within the BMCs, these were considered as filled (Figure 5B). Compared to the untagged versions, both D18 and P18 encapsulation peptides allow more Mdh and Das to be addressed in BMCs (65-68% for tagged Mdh and 75-80% for tagged Das). However, we could not conclude about a possible difference of the encapsulation rate between the D18 and P18 encapsulation peptides. Although the method used to determine the electron density differences is subjective because depending on the operator's eye, we can nevertheless conclude that the addition of encapsulation peptides to Mdh and Das tended to allow more proteins to be addressed in BMCs. Juodeikis et al. used two different methods to investigate the influence of encapsulation peptides on the proteins recruitment into eBMCs (Juodeikis et al., 2020). One was based on confocal microscopy fluorescence and immuno-TEM, they showed that both D18 and P18 encapsulation allowed Citrine to be specifically encapsulated into eBMCs in contrast to when the Citrine was fused

with the L20 encapsulation peptide. The second method was by using metallothioneins, proteins that are known to bind various metal ions easing their visualization by TEM and therefore the comparative analysis. Their statistical analysis showed that with both D18 and L20 encapsulation peptide, at least 95 % proteins were addressed into eBMCs when only 50 % did when fused with the P18 encapsulation peptide. It is most likely that the nature of both the encapsulation peptide and the protein influence the protein recruitment into eBMCs. However, the role of encapsulation peptide is determinant in the protein recruitment into eBMCs as they not only promote protein aggregation but were also shown to interact with some protein shells (Fan et al., 2012; Lawrence et al., 2014). Interestingly, the level of protein aggregation was also shown to influence the size of recombinant BMCs (Mayer et al., 2016; Juodeikis et al., 2020). Based on this data and the ones obtained with the enzymatic assay, we decided to choose the D18 encapsulation peptide for Mdh and the P18 for Das.

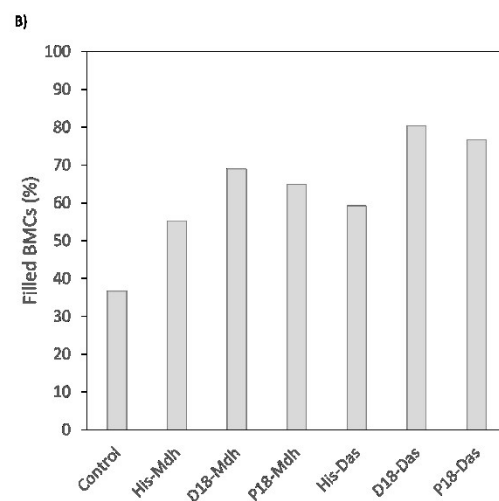
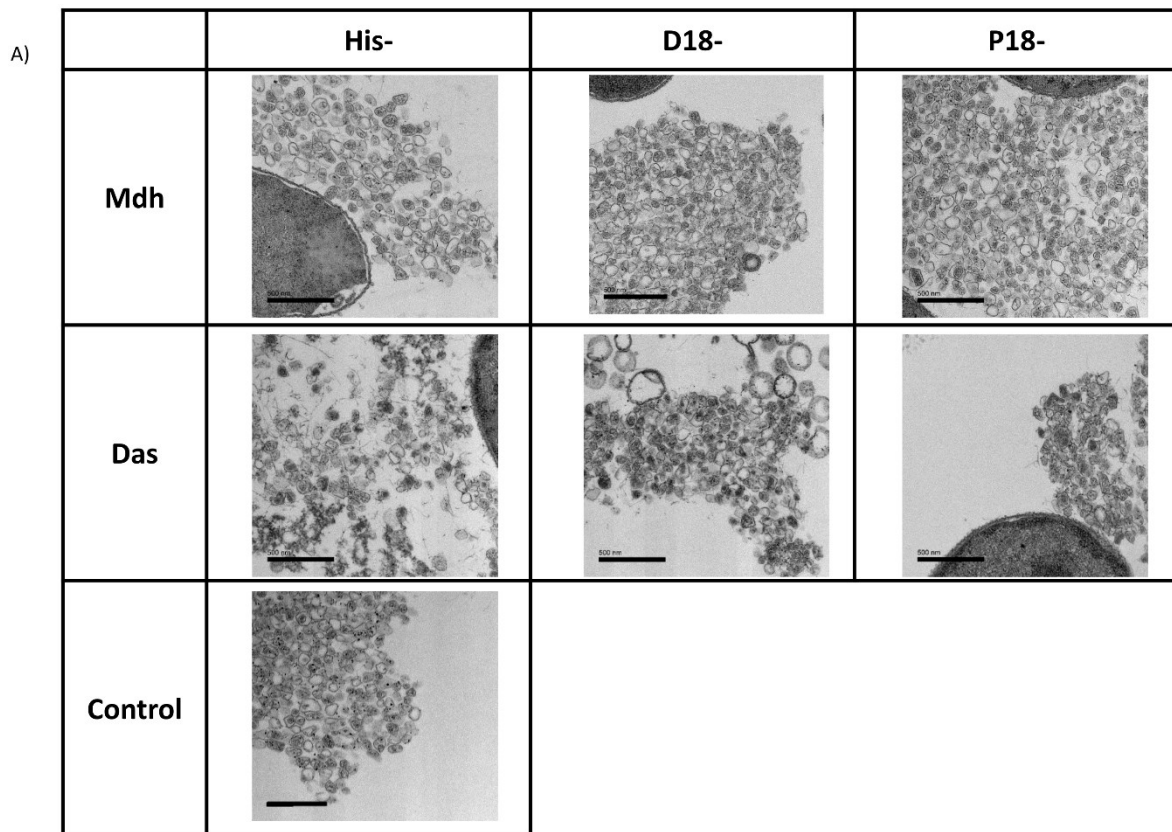


Figure 5: (A) TEM analysis of *E. coli* strains producing tagged (D18 and P18) and untagged (His) Mdh in the presence of BMCs (PduA-U). Control condition is the production of eBMC alone (and with the empty plasmid). Scale bar shows 0.5 μm . (B) Percentage of purified BMCs coproduced with the tagged or untagged Mdh and Das. Successful encapsulation was indicated by electron density within the lumen of the BMC. One count of a single biological replicate was done ($n \geq 600$).

***In vitro* characterization of methylotrophic BMCs by $^1\text{H-NMR}$**

In order to assess the functionality of the encapsulated D18-Mdh and P18-Das into eBMCs, the methylotrophic BMCs (mBMCs) were purified and analysed by $^1\text{H-NMR}$. First, mBMCs were isolated as previously described by (Lawrence et al., 2014). When co-expressed with the eBMC, D18-Mdh and P18-Das were present in the final BMC fraction (Supplementary Figure S4A). As the BMC purification protocol relies on differential salt precipitation steps and thus, on differential solubility behaviour, we wanted to make sure that the Mdh and Das were indeed encapsulated within the mBMC and not co-precipitating with the final BMC fraction. Untagged Mdh and Das were also co-expressed with the eBMC (Supplementary Figure S4B). Both untagged Mdh and Das were found in the final fraction but to a smaller extent compared to tagged Mdh and Das. In addition, untagged Das was found in a larger proportion than the untagged Mdh. Because the untagged Das form aggregates in the cytosol, this can promote the formation of BMCs around it. This is consistent with the previous data showing that more BMCs were filled when untagged Mdh or Das were expressed with the eBMC compared to the eBMC alone. When untagged Mdh and Das were co-expressed without the eBMC, proteins were not found in the final fraction. Thus, untagged Mdh and Das presence in the final fraction with the eBMC does not come from a coprecipitation but most probably comes from an unspecific encapsulation (Supplementary Figure S4C).

Methanol and Xu5P were used to test the functionality of the mBMCs (Figure 6A). After 16 hours, $^1\text{H-NMR}$ analysis revealed that 80 μM of DHA were produced by mBMCs which corresponds to 4.2 % of the consumed Xu5P whereas no DHA was detected in the eBMC condition (Figure 6B). These data demonstrate that the DHA production was specific and more specifically linked to D18-Mdh and P18-Das presence in the mBMC. Surprisingly, in the absence of methanol or Xu5P, we observed a small production of DHA (i.e. 10 μM without Xu5P and 20 μM without methanol) and NADH. DHA production in these conditions is may be due to a promiscuous activity of Mdh, which would partly explain the NADH production. Because we used Amicon centrifugal filters during the final step of the BMC purification, glycerol was found in all the final BMC fractions (data not shown). As described by the manufacturer, traces of glycerol are found on the Amicon membranes. Mdh was shown to exhibit higher catalytic activity for long chain alcohol and in alkaline conditions (Krog et al., 2013). As alkaline conditions were used for the mBMC analysis, it is therefore possible that

Mdh exhibits a promiscuous activity converting glycerol to DHA. This hypothesis can be supported by the fact (i) that no DHA was produced in the eBMC condition, linking DHA production by mBMCs without methanol or Xu5P to the presence of the protein content of the mBMCs (i.e. Mdh and Das) (ii) and that two times more DHA was produced in the absence of methanol than in the absence of Xu5P, which suggests that there might be a competition between methanol and glycerol as substrate for Mdh. In addition, Mdh was shown to convert glycerol into glyceraldehyde (Wolf et al., 1982). Peaks overlapping and low concentrations made the identification of glyceraldehyde impossible on the $^1\text{H-NMR}$ spectra. To prove that in our conditions Mdh can convert glycerol into glyceraldehyde, a spike with a glyceraldehyde standard will be needed to determine glyceraldehyde presence. If this hypothesis turns out to be true, this would suggest that another enzyme, most probably an isomerase, is present in the mBMC to convert glyceraldehyde into DHA.

In addition to DHA, two main molecules were produced by the mBMCs. According to their chemical shift, pyruvate (one singlet at 2.36 ppm) and its hydrated form, hydroxypyruvate (one singlet at 1.49 ppm), were identified (Nemeria et al., 2005). Several singlets in the methyl area were also found but further investigation is required to identify them. We also noticed that 2.17 mM of Xu5P were consumed by the mBMCs in absence of methanol that pyruvate and its hydrated form were produced. In this condition, 270 μM of pyruvate were produced whereas 2.3 mM of pyruvate were produced by the mBMCs in the presence of methanol and Xu5P. This suggests that the pyruvate production may come from a promiscuous activity from Das. However, this is not consistent with the fact that the highest amount of Xu5P was consumed (2.35 mM) in the eBMC in presence of Xu5P and methanol, resulting in a production of 330 μM pyruvate and its hydrated form (Figure 6B). In addition, several peaks in the methyl area, as the ones observed in the mBMC condition, were also found. Unfortunately, in any of our conditions, we haven't been able to complete the carbon balance from the Xu5P consumption and a possible glycerol consumption. Finally, no pyruvate nor hydroxypyruvate were produced during the demonstration of the operation of the purified Mdh and Das together. This demonstrates that the Xu5P consumption observed in the mBMC and eBMC does not come from a promiscuous activity of Das but most probably come from one or more enzymes unspecifically addressed within the mBMCs. Pdu BMC formation has not been elucidated yet, but some evidences let assume that Pdu BMC formation is following the same

type of assembly than β -carboxysomes. In β -carboxysomes, the protein core first assemble *via* proteins interacting by their protein domain or most commonly *via* aggregation of proteins carrying an encapsulation peptide and then the shell is forming around the proteins aggregate (Kerfeld and Erbilgin, 2015). Hence, one possibility is that the BMC shell is forming around endogenous proteins from *E. coli* that are forming aggregates thanks to native encapsulation peptides. BL21 like other *E. coli* strains (e.g. BW25513 or MG1655) contain the *eut* operon in their genome. By using the encapsulation peptide of EutE or EutC, proteins found within the Eut BMC, heterologous proteins could be addressed in the Pdu BMC (Jakobson et al., 2015). Unlike BW25513 and MG1655, the *eut* operon is intact in BL21 but none of the Eut proteins exhibits an enzymatic activity that would explain the pyruvate production from Xu5P. Moreover, the *eut* operon is only expressed in the presence of ethanolamine (Roof and Roth, 1992). Therefore, it is very unlikely that this possibility takes place in our conditions.

Regarding the nature of the unspecific reaction(s) taking place in our conditions, if we consider that the C5 Xu5P is converted into the C3 pyruvate, the missing molecule is a C2-molecule that would be in the same quantity range than the pyruvate. However, we did not observe any peak meeting with this criterion on the NMR spectra obtained with the eBMCs or the mBMCs. Pdu shell proteins have not been described to have any enzymatic activity so far, hence this enzymatic activity should result from an enzyme present within the BMC. The only enzyme described to use Xu5P as a substrate and to produce one C3-molecule and one C2-molecule is a TPP-dependent phosphoketolase using an inorganic phosphate and producing Ga3P and acetyl-phosphate (Henard et al., 2017) but such enzyme exhibiting has not been described in *E. coli* yet. We could also imagine that Xu5P could react with a C1 molecule. The 1-deoxy-xylulose-5-phosphate synthase is a protein present in *E. coli* that was described *in vitro* to produce 1-deoxy-Xu5P and CO₂ from hydroxypyruvate and Ga3P or Xu5P and CO₂ from pyruvate and Ga3P but we did not find the same amounts of Ga3P and pyruvate or hydroxypyruvate (Schürmann et al., 2002; Brammer et al., 2011). In addition, it is also possible that Xu5P can be condensed with C2 molecule. Indeed, DHA was showed to be unstable in solution and to non-enzymatically be degraded in other products as formate, glycolate or acetate (Peiro et al., 2019). Therefore, DHA stability in these conditions should also be investigated to determine if these non-enzymatic reactions are taking place. Glycerol involvement in these unspecific metabolic reactions should also be investigated.

Taken together, our data showed for the first time the *in vitro* functionality of a synthetic methylotrophic BMC but also highlighted that unspecific activity impaired the fully operation of the methylotrophic activities. These data question the enzymatic specificity of the BMC. However, one should keep in mind that here, the eBMC production is under a strong promoter (i.e. T7 RNA polymerase promoter) which enables the production of hundreds of BMCs in the cells (Figure 6C). Therefore, the probability to have unspecific enzymes encapsulated within the mBMCs is very high compared to native conditions.

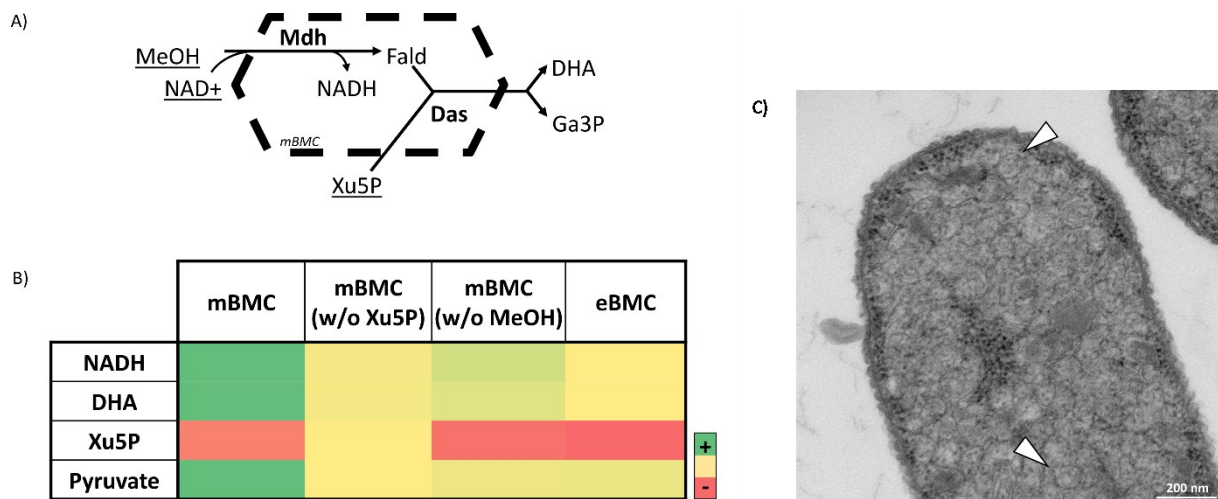


Figure 6: mBMC *in vitro* analysis. (A) Schematic representation of the methylotrophic BMC (mBMC). Substrates are underlined and enzyme are in bold. (B) Heatmap is showing metabolites consumption (red) or production (green) or neither consumed or produced (yellow) 16 hours after the beginning of the enzymatic assay. Values are averaged from the concentration difference after 16 hours of two duplicates. Levels of production or consumption were compared by metabolites and not by conditions. (C) TEM analysis of *E. coli* strains producing mBMCs. Arrows show mBMCs. Scale bar indicates 200 nm.

Dihydroxyacetone synthase (Das), methanol dehydrogenase (Mdh). Dihydroxyacetone (DHA), formaldehyde (Fald), glyceraldehyde-3-phosphate (Ga3P), methanol (MeOH), xylulose-5-phosphate (Xu5P).

Effect of the encapsulation on the methylotrophic pathway efficiency

To evaluate whether that methanol conversion to DHA was improved due to the encapsulation of Mdh and Das into mBMCs, we tested if similar amounts of DHA could be produced using the same amounts of non-encapsulated Mdh and Das. We quantified D18-Mdh and P18-Das amount in the mBMC by SDS-PAGE. A sample of the mBMC was loaded in a SDS-PAGE gel with known amounts of untagged Mdh and Das. To determine the quantity of D18-Mdh and P18-Das within the mBMC, a standard curve was plotted to correlate the quantity of protein with the intensity of a band. Untagged Mdh and Das were then tested in the same conditions (i.e at 30°C, in TrisHCl 20 mM, pH = 8), with the same amount as the one measured in the mBMC. After 16 hours, even if a small amount of Xu5P was consumed, no DHA nor pyruvate were produced. However new singlets in the methyl area of the spectrum appeared. But these were minor considering the amount of Xu5P consumed. As these peaks were not produced in the combination of Mdh and Das tested at pH 7.4, this suggests that a promiscuous activity of Mdh and/or Das may take place in alkaline conditions. But most importantly, in the same conditions and with the same amount of Mdh and Das than in the mBMCs, DHA production was not possible with the non-encapsulated enzymes (Figure 7). This data supports that the idea that mBMCs enhance the methanol conversion to DHA by locally increasing the Mdh and Das and substrate concentrations. Enhancement of the substrate channelling by co-localization of a heterologous pathway in a BMC was previously predicted and showed *in vivo* (Lawrence et al. 2014, Jakobson et al., 2017), here we demonstrated it for the first time *in vitro*, using a methylotrophic pathway.

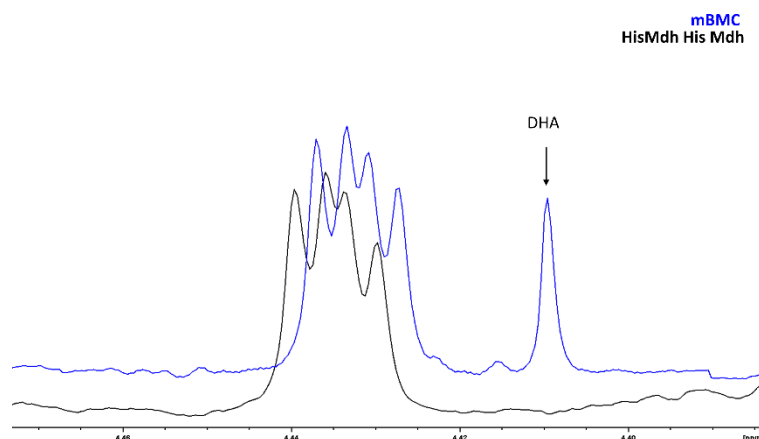


Figure 7: $^1\text{H-NMR}$ spectra showing the DHA production by the mBMCs (in blue) or the non-encapsulated Mdh and Das (in black) after 16 hours of reaction. Spectra are shown with a horizontal offset of 0.015 ppm and a vertical step of 10 %.

Dihydroxyacetone (DHA)

Conclusion

In this study, we successfully engineered mBMCs allowing a specific DHA production from methanol and Xu5P. We showed that co-localization of Mdh and Das within the eBMC enhanced the substrate channelling compared to non-encapsulated enzymes. $^1\text{H-NMR}$ is an approach without *a priori* that turned out to be determinant for the analysis of the *in vitro* mBMCs activity. It highlighted that methanol conversion to DHA by the mBMCs was limited by side reactions most probably coming from unspecific enzymes encapsulated within the mBMCs. Further investigations will enable the identification of these side reactions as well as the involved enzymes. The next step will be now to follow the fate of ^{13}C -methanol into central metabolism in strains expressing the mBMCs to determine if the presence of mBMCs can enhance the methanol assimilation *in vivo*. Further engineering on the mBMC could also be considered by adding a third enzyme that would ensure the recycling of NADH to balance the redox ratio within the mBMC allowing to sustain its activity while making it independent from the cell from a redox point of view. Finally, the construction of a genuine methylotrophic BMC may require the encapsulation of a complete pentose phosphate pathway allowing the regeneration of the C1 acceptor (i.e. Xu5P). This design would mimic the organization of the methylotrophic pathway in yeasts and eventually ensure a fully methylotrophic lifestyle to *E. coli*.

References

- Axen, S.D., Erbilgin, O., and Kerfeld, C.A. (2014). A Taxonomy of Bacterial Microcompartment Loci Constructed by a Novel Scoring Method. *PLoS Comput Biol* *10*, e1003898.
- BOBIK, T.A., HAVEMANN, G.D., BUSCH, R.J., WILLIAMS, D.S., and ALDRICH, H.C. (1999). The Propanediol Utilization (pdu) Operon of *Salmonella enterica* Serovar Typhimurium LT2 Includes Genes Necessary for Formation of Polyhedral Organelles Involved in Coenzyme B₁₂-Dependent 1,2-Propanediol Degradation[†]. *J. BACTERIOL.* *181*, 9.
- Brammer, L.A., Smith, J.M., Wade, H., and Meyers, C.F. (2011). 1-Deoxy-d-xylulose 5-Phosphate Synthase Catalyzes a Novel Random Sequential Mechanism. *J. Biol. Chem.* *286*, 36522–36531.
- Chen, F.Y.-H., Jung, H.-W., Tsuei, C.-Y., and Liao, J.C. (2020). Converting *Escherichia coli* to a Synthetic Methyloph Growing Solely on Methanol. *Cell* *182*, 933-946.e14.
- Cheng, S., Liu, Y., Crowley, C.S., Yeates, T.O., and Bobik, T.A. (2008). Bacterial microcompartments: their properties and paradoxes. *BioEssays* *30*, 1084–1095.
- Chessher, A., Breitling, R., and Takano, E. (2015). Bacterial Microcompartments: Biomaterials for Synthetic Biology-Based Compartmentalization Strategies. *ACS Biomaterials Science & Engineering* *1*, 345–351.
- Chowdhury, C., Sinha, S., Chun, S., Yeates, T.O., and Bobik, T.A. (2014). Diverse Bacterial Microcompartment Organelles. *Microbiol Mol Biol Rev* *78*, 438–468.
- Conrado, R.J., Mansell, T.J., Varner, J.D., and DeLisa, M.P. (2007). Stochastic reaction–diffusion simulation of enzyme compartmentalization reveals improved catalytic efficiency for a synthetic metabolic pathway. *Metabolic Engineering* *9*, 355–363.
- Dai, Z., Gu, H., Zhang, S., Xin, F., Zhang, W., Dong, W., Ma, J., Jia, H., and Jiang, M. (2017). Metabolic construction strategies for direct methanol utilization in *Saccharomyces cerevisiae*. *Bioresource Technology* *245*, 1407–1412.
- Dansen, T.B., Wirtz, K.W.A., Wanders, R.J.A., and Pap, E.H.W. (2000). Peroxisomes in human fibroblasts have a basic pH. *Nat Cell Biol* *2*, 51–53.
- De Simone, A., Vicente, C.M., Peiro, C., Gales, L., Bellvert, F., Enjalbert, B., and Heux, S. (2020). Mixing and matching methylotrophic enzymes to design a novel methanol utilization pathway in *E. coli*. *Metabolic Engineering* *61*, 315–325.
- Delépine, B., López, M.G., Carnicer, M., Vicente, C.M., Wendisch, V.F., and Heux, S. (2020). Charting the Metabolic Landscape of the Facultative Methyloph *Bacillus methanolicus*. *5*, 16.
- Fan, C., and Bobik, T.A. (2011). The N-Terminal Region of the Medium Subunit (PduD) Packages Adenosylcobalamin-Dependent Diol Dehydratase (PduCDE) into the Pdu Microcompartment. *Journal of Bacteriology* *193*, 5623–5628.
- Fan, C., Cheng, S., Liu, Y., Escobar, C.M., Crowley, C.S., Jefferson, R.E., Yeates, T.O., and Bobik, T.A. (2010). Short N-terminal sequences package proteins into bacterial microcompartments. *Proceedings of the National Academy of Sciences* *107*, 7509–7514.
- Fan, C., Cheng, S., Sinha, S., and Bobik, T.A. (2012). Interactions between the termini of lumen enzymes and shell proteins mediate enzyme encapsulation into bacterial microcompartments. *Proceedings of the National Academy of Sciences* *109*, 14995–15000.
- Fan, L., Wang, Y., Tuyishime, P., Gao, N., Li, Q., Zheng, P., Sun, J., and Ma, Y. (2018). Engineering Artificial Fusion Proteins for Enhanced Methanol Bioconversion. *ChemBioChem* *19*, 2465–2471.
- Gasteiger, E. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research* *31*, 3784–3788.
- Held, M., Quin, M.B., and Schmidt-Dannert, C. (2013). Eut Bacterial Microcompartments: Insights into Their Function, Structure, and Bioengineering Applications. *J Mol Microbiol Biotechnol* *23*, 308–320.

Henard, C.A., Smith, H.K., and Guarnieri, M.T. (2017). Phosphoketolase overexpression increases biomass and lipid yield from methane in an obligate methanotrophic biocatalyst. *Metabolic Engineering* 41, 152–158.

Jakobson, C.M., Kim, E.Y., Slininger, M.F., Chien, A., and Tullman-Ercek, D. (2015). Localization of Proteins to the 1,2-Propanediol Utilization Microcompartment by Non-native Signal Sequences Is Mediated by a Common Hydrophobic Motif. *Journal of Biological Chemistry* 290, 24519–24533.

Jakobson, C.M., Tullman-Ercek, D., Slininger, M.F., and Mangan, N.M. (2017). A systems-level model reveals that 1, 2-Propanediol utilization microcompartments enhance pathway flux through intermediate sequestration. *PLoS Computational Biology* 13, e1005525.

Juodeikis, R., Lee, M.J., Mayer, M., Mantell, J., Brown, I.R., Verkade, P., Woolfson, D.N., Prentice, M.B., Frank, S., and Warren, M.J. Effect of metabolosome encapsulation peptides on enzyme activity, coaggregation, incorporation, and bacterial microcompartment formation. *MicrobiologyOpen* n/a, e1010.

Kerfeld, C.A., and Erbilgin, O. (2015). Bacterial microcompartments and the modular construction of microbial metabolism. *Trends in Microbiology* 23, 22–34.

Kim, S., Lindner, S.N., Aslan, S., Yishai, O., Wenk, S., Schann, K., and Bar-Even, A. (2020). Growth of *E. coli* on formate and methanol via the reductive glycine pathway. *Nat Chem Biol*.

van der Klei, I.J., Yurimoto, H., Sakai, Y., and Veenhuis, M. (2006). The significance of peroxisomes in methanol metabolism in methylotrophic yeast. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1763, 1453–1462.

Krog, A., Heggeset, T.M., Müller, J.E., Kupper, C.E., Schneider, O., Vorholt, J.A., Ellingsen, T.E., and 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.

Lawrence, A.D., Frank, S., Newnham, S., Lee, M.J., Brown, I.R., Xue, W.-F., Rowe, M.L., Mulvihill, D.P., Prentice, M.B., Howard, M.J., et al. (2014). Solution Structure of a Bacterial Microcompartment Targeting Peptide and Its Application in the Construction of an Ethanol Bioreactor. *ACS Synthetic Biology* 3, 454–465.

Lee, M.J., Brown, I.R., Juodeikis, R., Frank, S., and Warren, M.J. (2016). Employing bacterial microcompartment technology to engineer a shell-free enzyme-aggregate for enhanced 1,2-propanediol production in *Escherichia coli*. *Metabolic Engineering* 36, 48–56.

Lee, M.J., Palmer, D.J., and Warren, M.J. (2018). Biotechnological Advances in Bacterial Microcompartment Technology. *Trends in Biotechnology*.

Liang, M., Frank, S., Lünsdorf, H., Warren, M.J., and Prentice, M.B. (2017). Bacterial microcompartment-directed polyphosphate kinase promotes stable polyphosphate accumulation in *E. coli*. *Biotechnol. J.* 12, 1600415.

Mayer, M.J., Juodeikis, R., Brown, I.R., Frank, S., Palmer, D.J., Deery, E., Beal, D.M., Xue, W.-F., and Warren, M.J. (2016). Effect of bio-engineering on size, shape, composition and rigidity of bacterial microcompartments. *Scientific Reports* 6.

McGoldrick, H.M., Roessner, C.A., Raux, E., Lawrence, A.D., McLean, K.J., Munro, A.W., Santabarbara, S., Rigby, S.E.J., Heathcote, P., Scott, A.I., et al. (2005). Identification and Characterization of a Novel Vitamin B₁₂ (Cobalamin) Biosynthetic Enzyme (CobZ) from *Rhodobacter capsulatus*, Containing Flavin, Heme, and Fe-S Cofactors. *Journal of Biological Chemistry* 280, 1086–1094.

Nemeria, N., Tittmann, K., Joseph, E., Zhou, L., Vazquez-Coll, M.B., Arjunan, P., Hübner, G., Furey, W., and Jordan, F. (2005). Glutamate 636 of the *Escherichia coli* Pyruvate Dehydrogenase-E1 Participates in Active Center Communication and Behaves as an Engineered Acetolactate Synthase with Unusual Stereoselectivity. *J. Biol. Chem.* 280, 21473–21482.

Olah, G.A. (2013). Towards Oil Independence Through Renewable Methanol Chemistry. *Angew. Chem. Int. Ed.* 52, 104–107.

Parsons, J.B., Dinesh, S.D., Deery, E., Leech, H.K., Brindley, A.A., Heldt, D., Frank, S., Smales, C.M., Lünsdorf, H., Rambach, A., et al. (2008). Biochemical and Structural Insights into Bacterial Organelle Form and Biogenesis. *Journal of Biological Chemistry* *283*, 14366–14375.

Parsons, J.B., Frank, S., Bhella, D., Liang, M., Prentice, M.B., Mulvihill, D.P., and Warren, M.J. (2010). Synthesis of Empty Bacterial Microcompartments, Directed Organelle Protein Incorporation, and Evidence of Filament-Associated Organelle Movement. *Molecular Cell* *38*, 305–315.

Peiro, C., Millard, P., de Simone, A., Cahoreau, E., Peyriga, L., Enjalbert, B., and 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., and Chen, W. (2016). Scaffoldless engineered enzyme assembly for enhanced methanol utilization. *Proceedings of the National Academy of Sciences* *113*, 12691–12696.

ROOFt, D.M., and Roth, J.R. (1992). Autogenous Regulation of Ethanolamine Utilization by a Transcriptional Activator of the eut Operon in *Salmonella typhimurium*. *174*, 10.

Ru\smayer, H., Buchetics, M., Gruber, C., Valli, M., Grillitsch, K., Modarres, G., Guerrasio, R., Klavins, K., Neubauer, S., and Drexler, H. (2015). Systems-level organization of yeast methylotrophic lifestyle. *BMC Biology* *13*, 80.

Sakai, Y., Saiganji, A., Yurimoto, H., Takabe, K., Saiki, H., and Kato, N. (1996). The absence of Pmp47, a putative yeast peroxisomal transporter, causes a defect in transport and folding of a specific matrix enzyme. *Journal of Cell Biology* *134*, 37–51.

Sampson, E.M., and Bobik, T.A. (2008). Microcompartments for B12-Dependent 1,2-Propanediol Degradation Provide Protection from DNA and Cellular Damage by a Reactive Metabolic Intermediate. *JB* *190*, 2966–2971.

Schrader, J., Schilling, M., Holtmann, D., Sell, D., Filho, M.V., Marx, A., and Vorholt, J.A. (2009). Methanol-based industrial biotechnology: current status and future perspectives of methylotrophic bacteria. *Trends in Biotechnology* *27*, 107–115.

Schürmann, M., Schürmann, M., and Sprenger, G.A. (2002). Fructose 6-phosphate aldolase and 1-deoxy-d-xylulose 5-phosphate synthase from *Escherichia coli* as tools in enzymatic synthesis of 1-deoxysugars. *Journal of Molecular Catalysis B: Enzymatic* *19–20*, 247–252.

Shen, J., Zeng, Y., Zhuang, X., Sun, L., Yao, X., Pimpl, P., and Jiang, L. (2013). Organelle pH in the Arabidopsis Endomembrane System. *Molecular Plant* *6*, 1419–1437.

Tang, C.T., Ruch, F.E., and Lin, C.C. (1979). Purification and properties of a nicotinamide adenine dinucleotide-linked dehydrogenase that serves an *Escherichia coli* mutant for glycerol catabolism. *Journal of Bacteriology* *140*, 182–187.

Trentham, D.R., McMurray, C.H., and Pogson, C.I. (1969). The active chemical state of d-glyceraldehyde 3-phosphate in its reactions with d-glyceraldehyde 3-phosphate dehydrogenase, aldolase and triose phosphate isomerase. *Biochemical Journal* *114*, 19–24.

Waites, M.J., and Quayle, J.R. (1983). Dihydroxyacetone Synthase: a Special Transketolase for Formaldehyde Fixation from the Methylotrophic Yeast *Candida boidinii* CBS 5777. *Microbiology* *129*, 935–944.

Wang, X., Wang, Y., Liu, J., Li, Q., Zhang, Z., Zheng, P., Lu, F., and Sun, J. (2017). Biological conversion of methanol by evolved *Escherichia coli* carrying a linear methanol assimilation pathway. *Bioresources and Bioprocessing* *4*.

Wolf, H. J. (1982). Process for preparing glyceraldehyde from glycerol with methanol dehydrogenase. US4353987A.

Yanase, H., Okuda, M., Kita, K., Sato, Y., Shibata, K., Sakai, Y., and Kato, N. (1995). Enzymatic preparation of [1, 3-¹³C] dihydroxyacetone phosphate from [¹³C] methanol and hydroxypyruvate using the methanol-assimilating system of methylotrophic yeasts. *Applied Microbiology and Biotechnology* 43, 228–234.

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

Yung, M.C., Bourguet, F.A., Carpenter, T.S., and Coleman, M.A. (2017). Re-directing bacterial microcompartment systems to enhance recombinant expression of lysis protein E from bacteriophage φX174 in *Escherichia coli*. *Microb Cell Fact* 16, 71.

Supplementary data

Strain	Genotype	Source
JM109	endA1, recA1, gyrA96, thi, hsdR17 (rk ⁻ , mk ⁺), relA1, supE44, Δ(lac-proAB), [F', traD36, proAB, laqIqZΔM15]	Promega
Top10	F- <i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) Φ80 <i>lacZ</i> ΔM15 Δ <i>lacX74 recA1 araD139</i> Δ(<i>araleu</i>)7697 <i>galU galK rpsL</i> (StrR) <i>endA1 nupG</i>	ThermoFisher
BL21* (DE3)	F- <i>ompT hsdS_B</i> (r _B ⁻ , m _B ⁻) <i>galdcmrne131</i> (DE3)	ThermoFisher

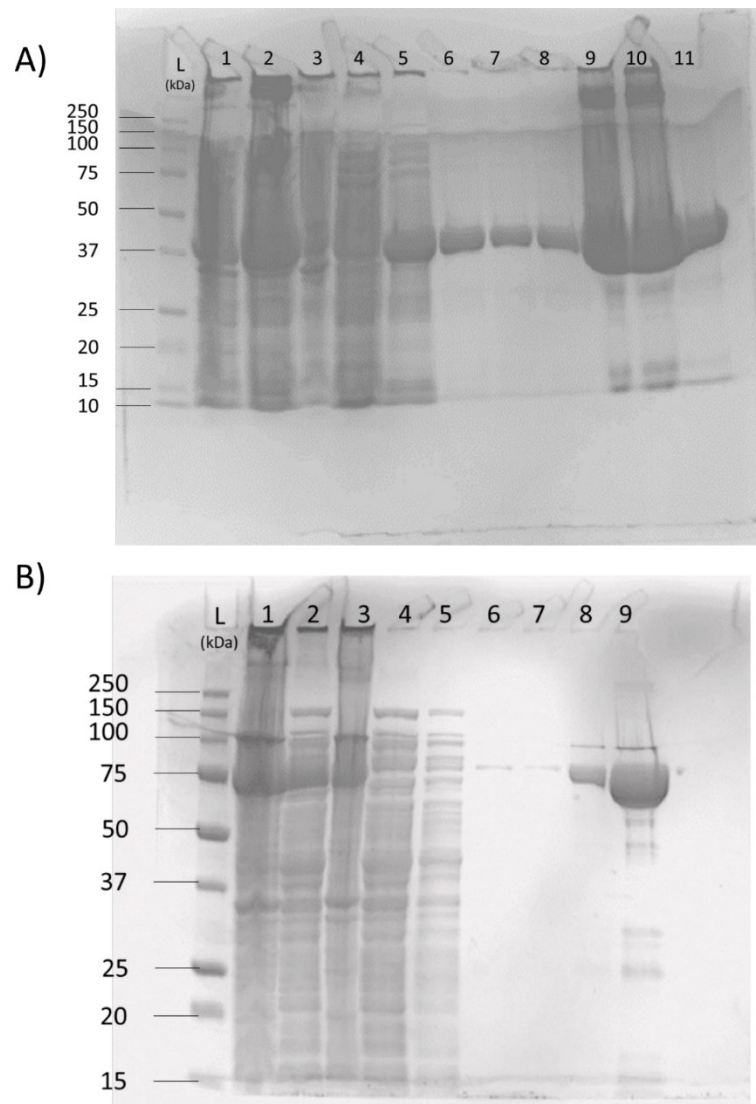
Supplementary Table 1: Strains used in this study

Name	Sequence 5' – 3'
Mdh_NdeI_FW	CACCATATGGCTTTTAAAAATCTTGCAGACC
Mdh_SpeI_RV	CACACTAGTTTACATTGCAGCTTCAAAAATCGC
Das_NdeI_FW	CACCATATGTCTATGCGCATTCCG
Das_SpeI_RV	CACACTAGTTTACAGTTTGTTAACTTTATCATGG

Supplementary Table 2: Oligonucleotides used in this study, restriction sites are underlined

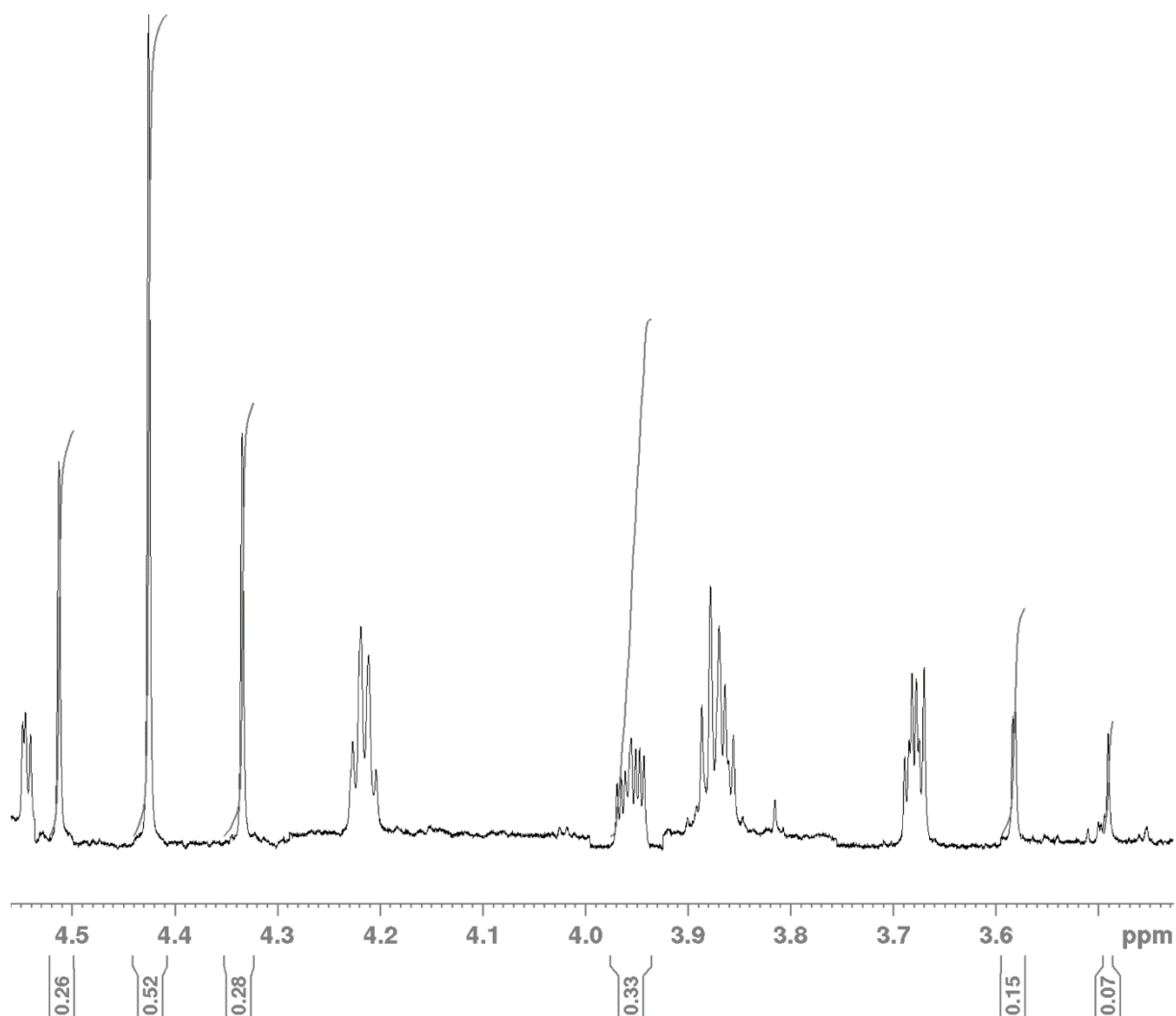
Plasmid name	Description	Source
pET14b	High copy number, overexpression vector containing a N-terminal His-tag with T7 promotor, Amp ^R , modified to include a <i>SpeI</i> site 5' of <i>BamHI</i>	Novagen
pET14b-D18	Overexpression vector derived from pET14b containing a N-terminal D18 encapsulating tag followed by a short amino acid linker (AMGSS) then a His-tag	(Lee et al., 2016)
pET14b-P18	Overexpression vector derived from pET14b containing a N-terminal P18 encapsulating tag followed by a short amino acid linker (PMGSS) then a His-tag	(Lee et al., 2016)
pLysS	Medium copy number, basal expression suppressor with T7 promotor, Cm ^R	Novagen
pLysS-PduABJKNU	pLysS containing Pdu shell genes required for the formation of empty BMCs	(Parsons et al., 2010)
pET14b-Mdh	PCR product of Mdh ligated into <i>NdeI/SpeI</i> sites of pET14b	This study
pET14b-Das	PCR product of Das ligated into <i>NdeI/SpeI</i> sites of pET14b	This study
pET14b-D18-Mdh	PCR product of Mdh ligated into <i>NdeI/SpeI</i> sites of pET14b-D18	This study
pET14b-D18-Das	PCR product of Das ligated into <i>NdeI/SpeI</i> sites of pET14b-D18	This study
pET14b-P18-Mdh	PCR product of Mdh ligated into <i>NdeI/SpeI</i> sites of pET14b-P18	This study
pET14b-P18-Das	PCR product of Das ligated into <i>NdeI/SpeI</i> sites of pET14b-P18	This study
pSEVA131-pTRC-gldA	Overexpression vector of <i>gldA</i> containing a N-terminal hexa histidine-tag	(Peiro et al., 2019)
pCPE-1	<i>XbaI/ClaI</i> fragment from pET14b-Das ligated into <i>XbaI/ClaI</i> sites of pET14b-Mdh	This study
pCPE-2	<i>XbaI/ClaI</i> fragment from pET14b-P18-Das ligated into <i>XbaI/ClaI</i> sites of pET14b-D18-Mdh	This study

Supplementary Table 3: Plasmids used in this study



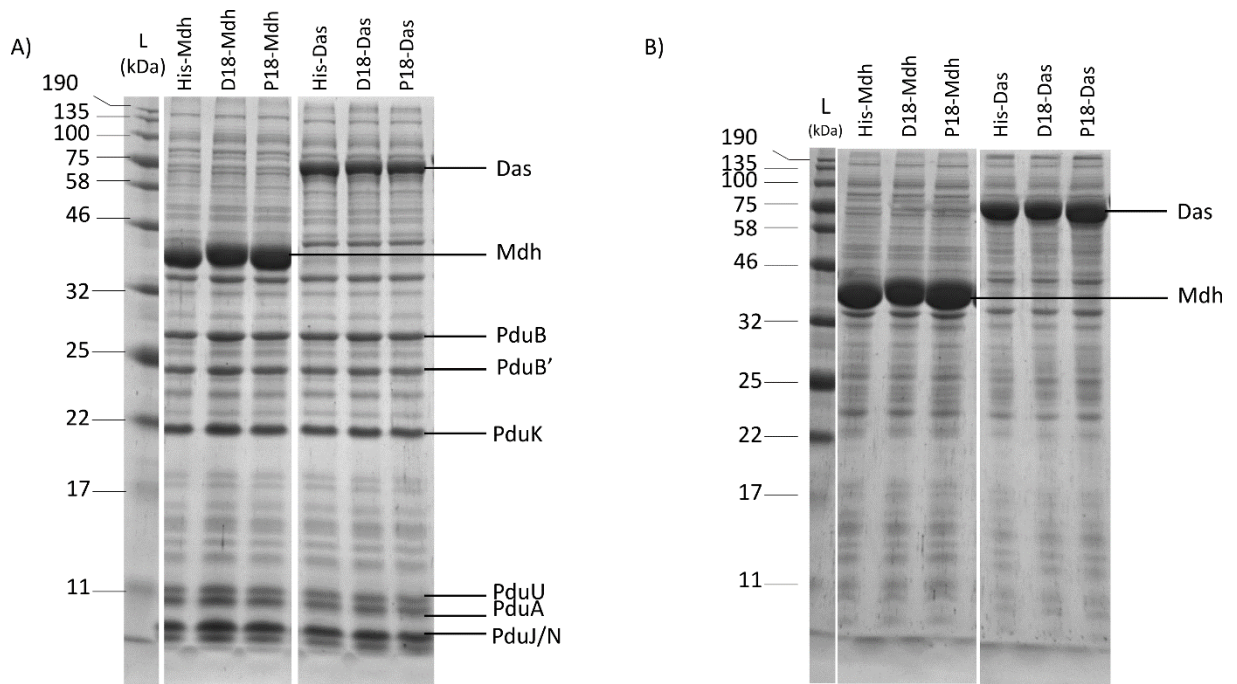
Supplementary Figure S1: SDS-PAGE of (A) Mdh from *A. gernerii* and (B) Das from *P. angusta* purified by IMAC.

In comparison to a molecular weight marker. Lane 1 – lysate, lane 2 – supernatant after centrifugation, lane 3 – pellet after centrifugation, lane 4 – supernatant flow through, lane 5 – wash buffer 1 flow through, lane 6 – wash buffer 2 flow through, lane 7 – wash buffer 3 flow through (100 mM imidazole), lane 8 – elution fraction 1 (100mM imidazole), lane 9 – elution fraction 2 (300mM imidazole), lane 10– elution fraction 3 (300mM imidazole), lane 11 – elution fraction 4 (400mM imidazole).

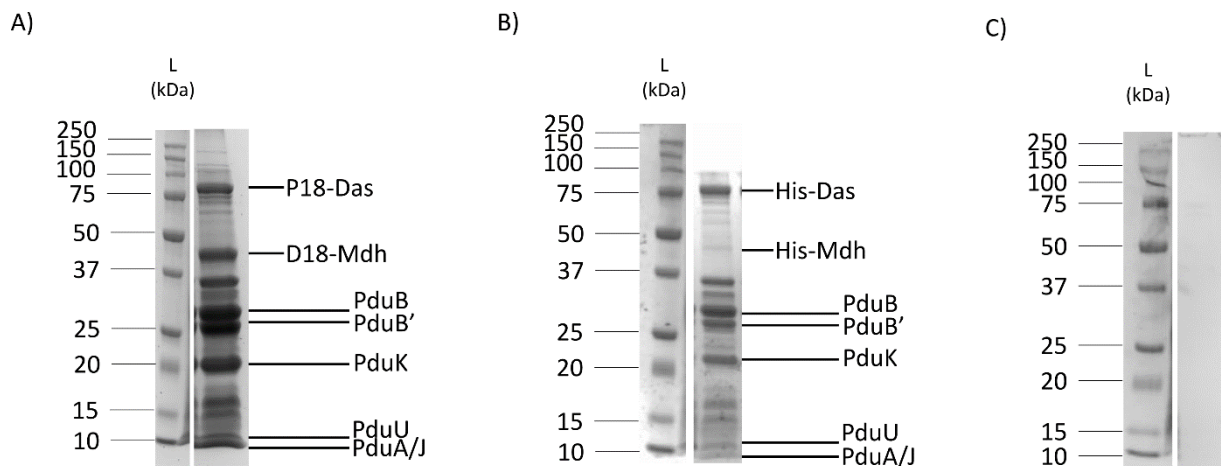


	DHA	Propanetetrol	Diol form of Ga3P
Integrated values	1.0064	0.2513	0.3328
Concentration (mM)	2.83		3.09

Supplementary Figure S2: DHA and Diol form of Ga3P quantification.



Supplementary Figure S3: SDS-PAGE of the untagged and tagged Mdh from *A. gernerii* and Das from *P. angusta* coproduced (A) with eBMCs or (B) without eBMCs (i.e. with empty plasmid). SDS samples were normalized by the OD.



Supplementary Figure S4: SDS-PAGE of the purified BMC fraction. eBMCs were co-expressed with (A) D18-Mdh and P18-Das or (B) His-Mdh and His-Das or (C) only His-Mdh and His-Das were co-expressed. Bands are less intense in panel B because protein induction was done at OD 0.4 whereas protein induction was done at OD 0.6-0.8 in panel A.

Conclusion and outlooks

This thesis aimed to functionally implement a new synthetic methylotrophic pathway composed of a bacterial Mdh and a yeast Das (i.e. MUT pathway) in the production host *E. coli*. To achieve this goal, an iterative approach was used to rationally and spatially optimize the MUT pathway in *E. coli*.

In **Chapter 2**, I investigated DHA metabolism in native *E. coli* strain using constraint-based modelling and system-level analysis approaches. The combination of these two approaches enabled to decipher enzymes involved in DHA assimilation while giving guidelines on the different constraints limiting *E. coli* to optimally grow on this carbon source. During our investigations, we found that DHA assimilation was improved by using $\Delta ptsA::kan$ mutant from the Keio collection. In this mutant, the polar effect of the kanamycin resistance cassette enabled the overexpression of both *gldA* and *fsaB* genes downstream, which encode for enzymes involved respectively in GLD and FSA pathways. This highlights, that prior to carry any phenotypical analysis using Keio mutants, it is crucial to remove the kanamycin resistance cassette to avoid unwanted phenotypes. During growth on DHA, our transcriptomic data showed that neither *fsaA* nor *fsaB* were upregulated. When we overexpressed both, only *fsaB* overexpression was associated with a higher DHA uptake rate showing its direct involvement in DHA assimilation. Moreover, when Lindner et al. engineered a synthetic pathway for glycerol assimilation overexpressing *gldA* and *fsaA* in a $\Delta gplK\Delta dhaK$ *E. coli* strain, growth on glycerol could not be rescued (Lindner et al., 2020). This result raises some interrogations regarding FsaA and FsaB involvement in *E. coli* metabolism. Indeed, FsaA and FsaB are isozymes sharing 70 % of similarity but because of its high specific activity for DHA, FsaA has been extensively used *in vitro* for synthetic chemistry applications (Sánchez-Moreno et al., 2012). However, according to our data, it seems that *in vivo* FsaB is more specific for DHA than FsaA. There is no data describing any particular role of FsaA in *E. coli* metabolism. *fsaA* is a single gene on *E. coli* genome, which complicates the task to link this enzyme to a particular metabolic process, its promoter as well as its regulation are not known yet. In contrast, *fsaB* is part of the same operon along with *gldA* and *ptsA*. *ptsA* encodes for a putative PTS system protein possibly involved in anaerobic fructose catabolism (Reizer et al., 1995). In the same

region of the genome, several genes encoding for enzymes involved in fermentation process are present. It is most likely that under anaerobic growth on fructose, *ptsA* would ensure fructose phosphorylation into F6P, then *FsaB* would convert F6P into DHA and Ga3P, finally *GldA* would produce glycerol from DHA. This hypothesis is supported by experiments made on *Zymomonas anaerobia* where DHA and glycerol were specific products of fructose fermentation (McGILL and Dawes, 1971). Elucidation of the role of *FsaA* and *FsaB* will not only complete our knowledge on *E. coli* but may also highlight how two enzymes sharing a high level of similarity at the protein-level, can have different roles *in vivo*.

Further engineering to improve DHA assimilation could be considered by engineering *FsaB*. By modifying one amino acid of the catalytic pocket, *FsaA* A129S exhibited improved catalytic efficiency for DHA than the native *FsaA* *in vitro* (Castillo et al., 2010). The residues composing *FsaA* and *FsaB* catalytic pockets are the same (Thorell et al., 2002). Hence, a similar modification of the catalytic pocket could be considered to improve the kinetic parameters of *FsaB* for DHA in the $\Delta ptsA::kan$ mutant.

We showed, in **Chapter 3**, that the introduction the MUT pathway only was not sufficient to allow cell growth on methanol. However, we demonstrated that methanol assimilation was improved by engineering a more efficient MUT pathway based on our combinatorial library, by completely blocking the dissimilation pathway and by improving the assimilation pathway. In parallel of this study, we decided to further improve the overall metabolic capacity towards methanol assimilation. We applied evolutionary engineering on a methylotrophic strain with a chromosomal integrated version of the improved MUT pathway (data not shown). We used an automated cultivation allowing the direct selection of evolved strains carrying novel functional metabolic pathways or enzymes with original properties (de Crécy-Lagard et al., 2001). During 83 days of evolution, in presence of 655 mM of methanol and a decreasing amount of xylose (from 15 to 5 mM), 10 strains were sequentially isolated and analyzed. The most evolved strain (i.e. EV83 strain) had a better fitness on the final conditions (i.e. 5 mM of xylose and 655 mM of methanol) and assimilated methanol faster than the parental strain. Despite improvements, the EV83 strain was still not able to grow on pure methanol. We believe that to continue the ALE experiment by including the genetic targets that were identified in **Chapter 3**, might lead to the selection of a strain able to grow on methanol, or at

least to the selection of a strain exhibiting a faster assimilation of methanol. This is supported by the study of Chen et al., six months of ALE experiment along with rational engineering were required to select a synthetic methylotrophic *E. coli* expressing the RuMP pathway able to grow on pure methanol (Chen et al., 2020). From this experiment, we will select new mutants with beneficial mutations that will have led to the overall improvement of the methanol assimilation. By combining 'omics' approaches and next generation sequencing, we will get insight into the set of mutations that was responsible of the final evolved strain phenotype.

From the transcriptomic data obtained in the **Chapter 3**, we observed that methanol influenced negatively the cellular machinery by downregulated genes involved in these processes. By constructing of a synthetic regulon (i.e. synthetic transcription factor) responding to methanol, the catabolic pathway could be coordinated with the overall cellular infrastructure which may allow to overcome these limitations. These data also showed that in the synthetic methylotrophic strain expressing the MUT pathway, the *dhaKLM* operon was upregulated in presence of methanol. This operon is activated in presence of DHA, which links the presence of DHA in the cells to the methanol assimilation. Hence, this metabolite could be used to follow methanol assimilation using a DHA biosensor. Biosensor engineering is gaining a lot of interest in metabolic engineering as it allows a dynamic response to the presence of a specific target (Brockman and Prather, 2015). A promoter inducible to DHA had already been described by (Bächler et al., 2005). The P_{dhaK} promoter was associated with *lacZ* to measure *dhaKLM* operon expression. In this experiment, 2 mM of DHA were used to induce the expression. One key element in the engineering of biosensor is its sensitivity. Therefore, by constructing a similar plasmid associating P_{dhaK} promoter with *gfp* gene, the sensitivity of P_{dhaK} promoter will be assessed first. The P_{dhaK} promoter could be used to express an essential gene which will then be used as selection pressure to boost methanol assimilation during ALE experiment (Raman et al., 2014; Liu et al., 2017).

In the last chapter of this thesis (**Chapter 4**), I built a methylotrophic BMC (mBMC) by encapsulating the MUT pathway in an eBMC. We showed that methanol conversion into DHA was higher in the mBMC compared to non-encapsulated Mdh and Das, demonstrating that substrate channelling was improved thanks to the local increase of Mdh and Das concentration in the mBMC. Those preliminary results represent a groundwork for the

utilization of BMCs as versatile micro-reactors for the encapsulation of synthetic metabolic pathways. However, our results on eBMCs showed that an unspecific activity resulting into pyruvate production was taking place. Other compounds were also produced but we were not able to identify them. This highlights the presence of unspecific enzymes within the mBMC competing with Das to use Xu5P as a substrate, which may impair Das formaldehyde fixation and by extend may limit methanol assimilation. Further investigations of the eBMC should be considered to determine the identity of this or these protein(s). As a starting point, a complete identification of the other compounds produced by the eBMC should be done. Indeed, by identifying the nature of these products, we will be able to determine which kind of enzymatic reactions is taking place within the eBMC while possibly completing the carbon balance from Xu5P consumption. By increasing the amount of Xu5P in the assay, the concentration of the products should increase which will ease the interpretation of the 2D-NMR experiments. A proteome analysis of the eBMC could also be considered to identify the enzymes unspecifically addressed in the eBMCs. The activity of eBMC using other substrates could be tested to determine if any other enzymatic activity can be found.

Then, the *in vivo* efficiency of the mBMC as well as its effect on *E. coli* metabolism will be investigated. Our collaborator demonstrated that the production of compounds could be achieved *in vivo* by the heterologous expression of BMCs associated with specific metabolic pathways. However, this was achieved using BL21*DE3 strain cultured in a rich medium (i.e. LB medium). Prior to tackle this new challenge, some questions are raising: (i) Which *E. coli* strain between BL21*DE3 and BW25113 strains should be used? Indeed, BL21*DE3 strain is known to reach high yields in the protein production, which is useful for *in vitro* analysis (e.g structural analysis, enzymatic assay). However, all the phenotypical analysis made so far in **Chapters 2 and 3** were done in *E. coli* BW25113 background. In addition, in **Chapter 4**, we showed that cellular crowding was occurring when mBMCs were expressed using T7 promotor in BL21*DE3 promoting unspecific encapsulation. In addition, it was shown that on LB, cellular growth in this strain was impacted from the too high expression level of the mBMCs (Lee et al., 2016). Consequently, a second question is raising: (ii) Should we use a less strong promoter than the T7 promoter for the *in vivo* experiment to avoid a metabolic burden? Taking into considerations all these aspects, we plan to use the BW25113 background expressing the eBMC on a low-copy plasmid and the tagged Mdh and Das on medium copy-plasmid. To

determine how the synthetic methylotrophic strain is responding to the presence of the mBMCs and if the mBMC improve the methanol assimilation in *E.coli* in depth analysis of the metabolic and physiological consequences will be done, including ^{13}C -methanol experiment to follow methanol assimilation at the intracellular level.

The engineering of an upgraded version of the mBMC (mBMC*) is on going. As a first step, the addition of GldA in the mBMC could be an asset to enhance methanol assimilation. Indeed, GldA presence in the mBMC* will not only improve DHA assimilation but will also recycle NADH to NAD^+ , ensuring the balance of NADH/NAD ratio within the mBMC*. In the native Pdu BMC or Eut BMC, enzymes ensuring the recycling of the cofactors are found within the BMCs, creating a “private pool” of cofactors in the BMCs. Moreover, the creation of a “NADH sink” by associating Mdh with the lactate dehydrogenase was shown to enhance methanol conversion (Price et al., 2016). The encapsulation peptide to address the GldA in the eBMC has already been determined (Lee et al., 2016), which is ease the task to add GldA in the mBMC*. Then, we could go one-step further by encapsulating Tkt as well as other enzymes of the PPP in the mBMC* to recycle Xu5P. In methylotrophic yeasts, the recycling of Xu5P was shown to rather take place in the peroxisome than in the cytoplasm (Russmayer et al., 2015). In addition, several steps pathway had already been successfully addressed in eBMCs (Lee et al., 2016). Natively, the Pdu BMC contains six enzymes to ensure the degradation of 1,2-PD. Therefore, the addition of the new enzymes within the mBMC* should not be a problem from a space point of view. The encapsulation peptide choice is determinant to address each new enzyme within the mBMC*. Consequently, the setup of a purification protocol as well as an enzymatic assay will be required for each of them. Then, mBMC* purification will allow to verify the presence of all the enzymes newly added within the mBMC*. This way, the upgraded version of the mBMC* will not only have a similar organization than the methylotrophic plasmid of *B. methanolicus* but will also mimic the yeast peroxisome by concomitantly ensuring efficient DHA assimilation, balancing NADH/NAD ratio and recycling of Xu5P. All together, these leads should favour the methanol growth of the synthetic methylotrophic *E. coli* strain.

With the recent advances, we witnessed that engineering synthetic methylotrophy in platform organism remains a challenge that can be won. There is still some room for the

implementation of novel methylotrophic pathways as the MUT pathway. Now, new exciting challenges are awaiting to turn synthetic methylotrophic strains into efficient compounds producers.

References

- Bächler, C., Schneider, P., Bähler, P., Lustig, A., and Erni, B. (2005). *Escherichia coli* dihydroxyacetone kinase controls gene expression by binding to transcription factor DhaR. *EMBO J* *24*, 283–293.
- Brockman, I.M., and Prather, K.L.J. (2015). Dynamic metabolic engineering: New strategies for developing responsive cell factories. *Biotechnol. J.* *12*.
- Castillo, J.A., Guérard-Hélaine, C., Gutiérrez, M., Garrabou, X., Sancelme, M., Schürmann, M., Inoue, T., Hélaine, V., Charmantray, F., Gefflaut, T., et al. (2010). A Mutant D-Fructose-6-Phosphate Aldolase (Ala129Ser) with Improved Affinity towards Dihydroxyacetone for the Synthesis of Polyhydroxylated Compounds. *Adv. Synth. Catal.* *352*, 1039–1046.
- Chen, F.Y.-H., Jung, H.-W., Tsuei, C.-Y., and Liao, J.C. (2020). Converting *Escherichia coli* to a Synthetic Methylophile Growing Solely on Methanol. *Cell* *182*, 933-946.e14.
- de Crécy-Lagard, V.A., Bellalou, J., Mutzel, R., and Marlière, P. (2001). Long term adaptation of a microbial population to a permanent metabolic constraint: overcoming thymineless death by experimental evolution of *Escherichia coli*. *BMC Biotechnology* *7*.
- Lee, M.J., Brown, I.R., Juodeikis, R., Frank, S., and Warren, M.J. (2016). Employing bacterial microcompartment technology to engineer a shell-free enzyme-aggregate for enhanced 1,2-propanediol production in *Escherichia coli*. *Metabolic Engineering* *36*, 48–56.
- Lindner, S.N., Aslan, S., Müller, A., Hoffart, E., Behrens, P., Edlich-Muth, C., Blombach, B., and Bar-Even, A. (2020). A synthetic glycerol assimilation pathway demonstrates biochemical constraints of cellular metabolism. *FEBS J* *287*, 160–172.
- Liu, S.-D., Wu, Y.-N., Wang, T.-M., Zhang, C., and Xing, X.-H. Maltose Utilization as a Novel Selection Strategy for Continuous Evolution of Microbes with Enhanced Metabolite Production. *ACS Synthetic Biology* *60*.
- McGILL, D.J., and Dawes, E.A. (1971). Glucose and Fructose Metabolism in *Zymomonas anaerobia*. *125*, 10.
- Price, J.V., Chen, L., Whitaker, W.B., Papoutsakis, E., and Chen, W. (2016). Scaffoldless engineered enzyme assembly for enhanced methanol utilization. *Proceedings of the National Academy of Sciences* *113*, 12691–12696.
- Raman, S., Rogers, J.K., Taylor, N.D., and Church, G.M. (2014). Evolution-guided optimization of biosynthetic pathways. *Proc Natl Acad Sci USA* *111*, 17803–17808.
- Reizer, J., Reizer, A., and Saier Jr, M.H. (1995). Novel phosphotransferase system genes revealed by bacterial genome analysis—a gene cluster encoding a unique Enzyme I and the proteins of a fructose-like permease system. *Microbiology* *141*, 961–971.
- Russmayer, H., Buchetics, M., Gruber, C., Valli, M., Grillitsch, K., Modarres, G., Guerrasio, R., Klavins, K., Neubauer, S., and Drexler, H. (2015). Systems-level organization of yeast methylophile lifestyle. *BMC Biology* *13*, 80.
- Sánchez-Moreno, I., Nauton, L., Théry, V., Pinet, A., Petit, J.-L., de Berardinis, V., Samland, A.K., Guérard-Hélaine, C., and Lemaire, M. (2012). FSAB: A new fructose-6-phosphate aldolase from *Escherichia coli*. Cloning, over-expression and comparative kinetic characterization with FSAA. *Journal of Molecular Catalysis B: Enzymatic* *84*, 9–14.
- Thorell, S., Schürmann, M., Sprenger, G.A., and Schneider, G. (2002). Crystal Structure of Decameric Fructose-6-Phosphate Aldolase from *Escherichia coli* Reveals Inter-subunit Helix Swapping as a Structural Basis for Assembly Differences in the Transaldolase Family. *Journal of Molecular Biology* *319*, 161–171.