

RESEARCH

Open Access



# Rewiring *Escherichia coli* to transform formate into methyl groups

Michael K. F. Mohr<sup>1</sup>, Ari Satanowski<sup>2,3</sup>, Steffen N. Lindner<sup>4</sup>, Tobias J. Erb<sup>2,5</sup> and Jennifer N. Andexer<sup>1\*</sup>

## Abstract

**Background** Biotechnological applications are steadily growing and have become an important tool to reinvent the synthesis of chemicals and pharmaceuticals for lower dependence on fossil resources. In order to sustain this progression, new feedstocks for biotechnological hosts have to be explored. One-carbon (C<sub>1</sub>-)compounds, including formate, derived from CO<sub>2</sub> or organic waste are accessible in large quantities with renewable energy, making them promising candidates. Previous studies showed that introducing the formate assimilation machinery from *Methylobacterium extorquens* into *Escherichia coli* allows assimilation of formate through the C<sub>1</sub>-tetrahydrofolate (C<sub>1</sub>-H<sub>4</sub>F) metabolism. Applying this route for formate assimilation, we here investigated utilisation of formate for the synthesis of value-added building blocks in *E. coli* using S-adenosylmethionine (SAM)-dependent methyltransferases (MT).

**Results** We first used a two-vector system to link formate assimilation and SAM-dependent methylation with three different MTs in *E. coli* BL21. By feeding isotopically labelled formate, methylated products with 51–81% <sup>13</sup>C-labelling could be obtained without substantial changes in conversion rates. Focussing on improvement of product formation with one MT, we analysed the engineered C<sub>1</sub>-auxotrophic *E. coli* strain C<sub>1</sub>S. Screening of different formate concentrations allowed doubling of the conversion rate in comparison to the not formate-supplemented BL21 strain with a share of more than 70% formate-derived methyl groups.

**Conclusions** Within this study transformation of formate into methyl groups is demonstrated in *E. coli*. Our findings support that feeding formate can improve the availability of usable C<sub>1</sub>-compounds and, as a result, increase whole-cell methylation with engineered *E. coli*. Using this as a starting point, the introduction of additional auxiliary enzymes and ideas to make the system more energy-efficient are discussed for future applications.

**Keywords** Formic acid, Methylation, SAM, C<sub>1</sub>-building block, Methyltransferase, Biotransformation

\*Correspondence:

Jennifer N. Andexer

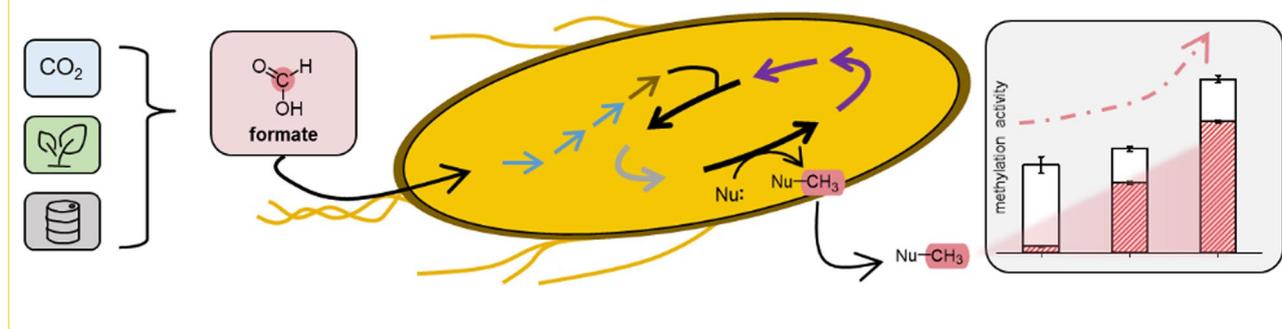
jennifer.andexer@pharmazie.uni-freiburg.de

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Graphical abstract



## Background

The majority of industrial processes depends on fossil resources for the synthesis of compounds such as natural products, fuels, polymers, or fine chemicals including pharmaceuticals [1, 2]. With depleting fossil resources and the emerging effects of climate change, defossilisation is an important strategy to create a greener and more sustainable chemical industry [1]. Development of biochemical applications utilising alternative and renewable feedstocks to support or replace existing chemical processes is key to support this transformation [3,4]. One promising approach is the application of CO<sub>2</sub>-derived feedstocks in chemobio-hybrid processes, *i.e.* initial (electro)chemical reduction of the greenhouse gas to produce reduced one-carbon (C<sub>1</sub>-) compounds such as formate and their subsequent valorisation through biotechnological efforts [5–7]. Additionally, enzymatic strategies for the less energy-intensive reduction of CO<sub>2</sub> are investigated which will allow utilisation of CO<sub>2</sub>-derived formate for one-step whole-cell biocatalysis without prior electrochemical reduction in future applications [8, 9]. Formate has the benefit of stability, biodegradability, ease of handling and low toxicity, which means that it poses a low risk towards humans and the environment and is in accordance with many principles of green chemistry [7, 10–12]. For this reason, the development of biotechnological strategies based on the C<sub>1</sub>-compound formate is an emerging field [7, 10, 13].

In nature, one important strategy to transfer C<sub>1</sub>-compounds onto various molecules is *S*-adenosylmethionine (SAM)-dependent methylation catalysed by methyltransferases (MTs, Fig. 1a) [14–16]. SAM-dependent methylation is involved in important processes such as DNA or histone methylation for epigenetic regulation

of gene expression [17, 18]. Various natural products with biological activity are methylated and methylation of small molecular drugs at specific positions has been shown to increase their potency by up to three orders of magnitude, called the “magic methyl effect” [19]. This makes selective methylation an important step in lead structure optimisation during development of new pharmaceuticals (Fig. 1b). Chemical introduction of methyl groups is in many cases challenging to direct and commonly mediated under energy-consuming conditions by the use of toxic and environmentally harmful methylation agents such as methyl iodide or diazomethane [20]. In comparison to chemical methylation, MTs introduce the methyl group stereo-, regio- and chemoselectively, only synthesising one specific product; by including radical SAM MTs, chemically inaccessible sp<sup>3</sup>-hybridised carbons could be targeted as well [14, 21]. While chemical methylation is not the major driving force for the use of fossil resources, it is a hazardous technique necessary for countless syntheses, and would benefit from substitution by a strategy which is safer, more environmental-friendly and, hence, more in accordance with the principles of green chemistry; an example for such a substitution strategy is enzymatic, SAM-dependent methylation [12, 15, 16, 22, 23].

Intracellular methylation activity is often a major limiting factor when natural compounds are produced in biotechnological hosts, emphasising the need to develop strategies to improve conversion rates to synthesise methylated products [22–25]. Optimisation of intracellular SAM availability and acceleration of the cofactor’s regeneration can improve the formation of methylated products using whole-cell methylation [22–26]. One

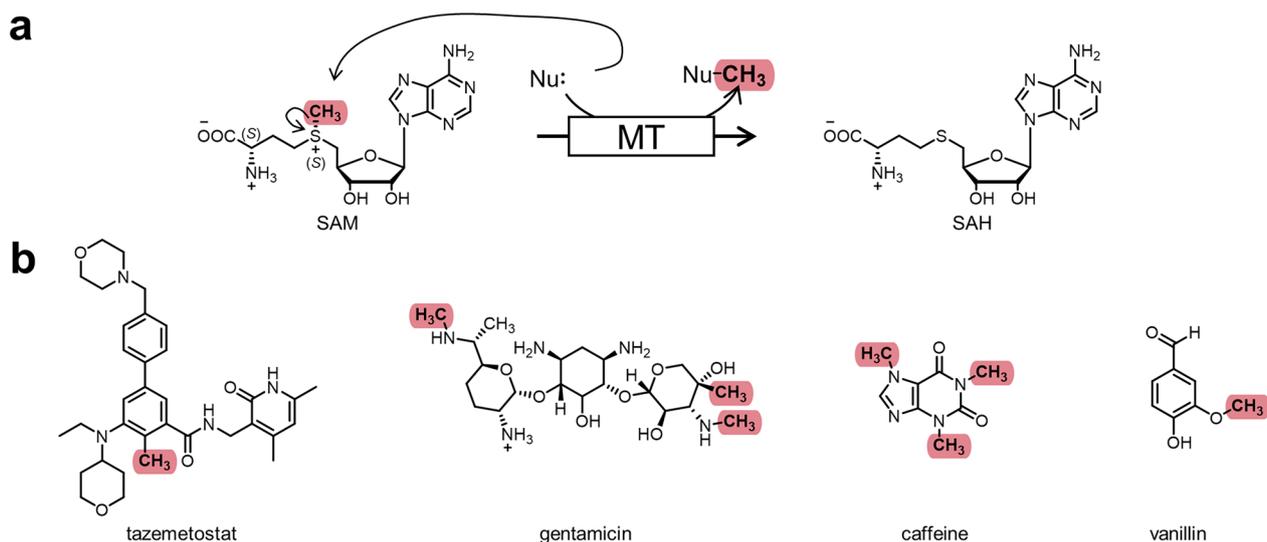
strategy which proved successful is increasing concentrations of the SAM precursor methionine, in particular its regeneration from homocysteine [23, 26–28]. In *E. coli*, homocysteine is methylated to methionine through transfer of a methyl group from methyl-tetrahydrofolate (methyl-H<sub>4</sub>F) catalysed by one of the methionine synthases *EcMETE* and *EcMETH* (Fig. 2) [29]. Majorly, the C<sub>1</sub>-H<sub>4</sub>F pool is fuelled from serine and glycine, catalysed by serine hydroxymethyltransferase (*EcGLYA*) and the glycine cleavage system (*EcGCS*), respectively [30, 31]. The C<sub>1</sub>-H<sub>4</sub>F pool supplies C<sub>1</sub>-carbon units for the synthesis of nucleobases and amino acids such as methionine or histidine, as well as for formylation of methionine, making it essential for protein production and cell growth [6, 32, 33]. Okano et al. reported that increasing the availability of C<sub>1</sub>-compounds to improve methyl-H<sub>4</sub>F production increases the formation of methylated products in *E. coli*. [28] We surmised that the C<sub>1</sub>-compound formate could be an ideal supplier of C<sub>1</sub>-building blocks to fuel the C<sub>1</sub>-H<sub>4</sub>F metabolism of biotechnological hosts and to drive whole-cell methylation forward.

In previous publications, strategies to assimilate formate into the metabolism of *E. coli* have been established [13, 34]. Introduction of the formate-H<sub>4</sub>F ligase from *Methylorubrum extorquens* (*MexFTFL*) into *E. coli* allows synthesis of formyl-H<sub>4</sub>F from formate and H<sub>4</sub>F in an adenosine-5'-triphosphate (ATP)-dependent reaction (Fig. 2a) [13, 35, 36]. Further, introduction of formyl-H<sub>4</sub>F cyclohydrolase (*MexFCH*) and methylene-H<sub>4</sub>F

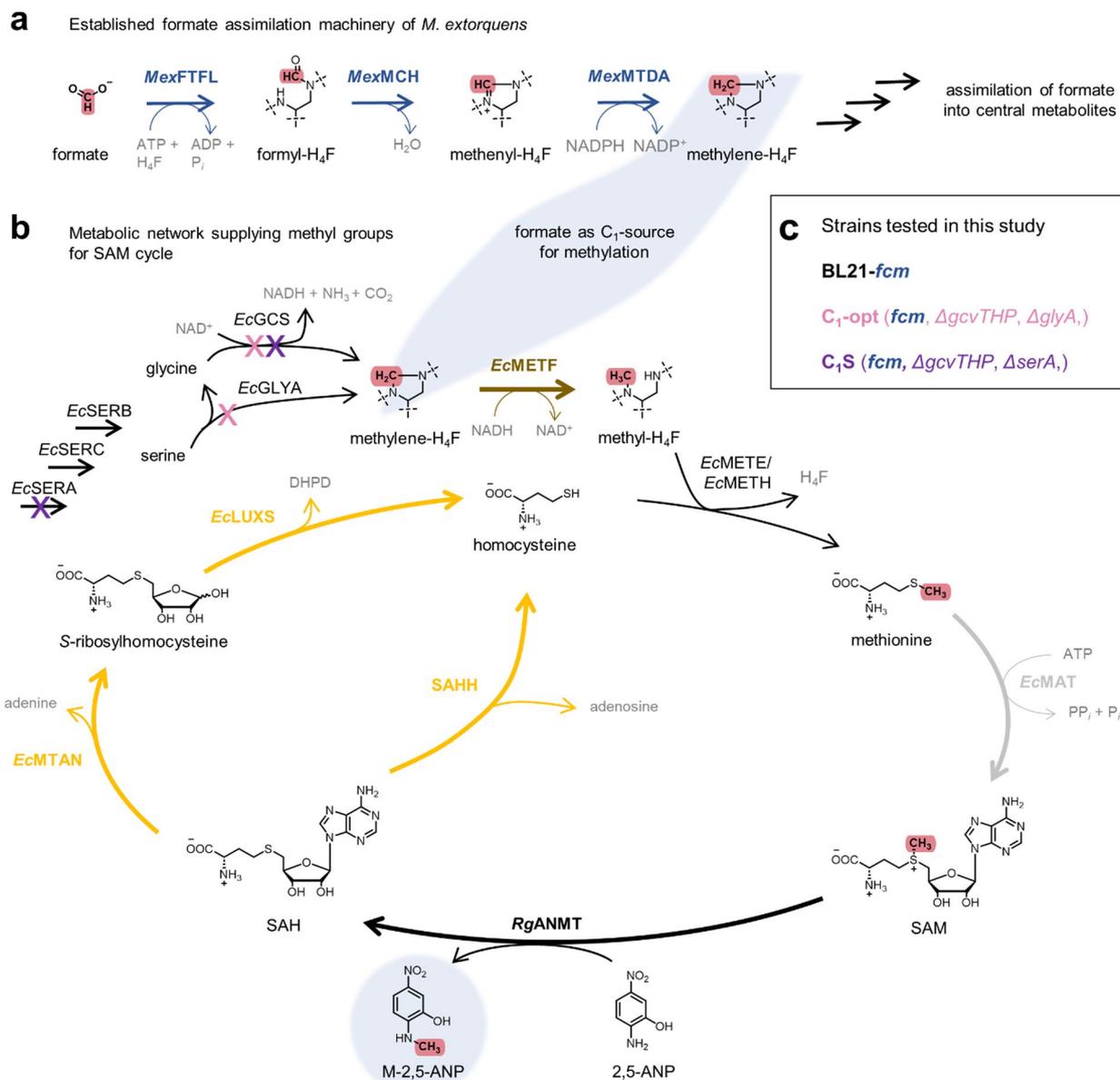
dehydrogenase A (*MexMTDA*), both from *M. extorquens*, cyclise and reduce formyl-H<sub>4</sub>F to methylene-H<sub>4</sub>F, the central metabolite of the C<sub>1</sub>-H<sub>4</sub>F metabolism, which can be used to further introduce formate's carbon into central metabolites [35–37]. This formate assimilation strategy composed of *MexFTFL*, *MexFCH* and *MexMTDA* (together referred to as FCM) proved powerful and was used for full assembly of all carbon scaffolds from formate and CO<sub>2</sub> in synthetic formatotrophic *E. coli* (Fig. 2a) [13, 34, 38].

In order to harness this established formate assimilation strategy for SAM-dependent methylation, the generated methylene-H<sub>4</sub>F has to be reduced once more by the NADH-dependent methylene-H<sub>4</sub>F reductase (*EcMETE*, Fig. 2) to methyl-H<sub>4</sub>F [39]. *EcMETE* or *EcMETH* transfer the methyl group onto homocysteine to produce the proteinogenic amino acid methionine, the precursor of SAM [29]. Previously reported introduction of FCM in engineered *E. coli* strains allowed all methyl groups of methionine to derive from formate when <sup>13</sup>C-labelled formate was fed, emphasising the potential to use formate for the synthesis of methyl groups [40]. Growth of the synthetic formatotrophs solely on formate is approximately 10-times slower than on established feedstocks [13, 38], making a strategy with growth from established feedstocks and supply of formate to push methylation of selected building blocks an interesting option.

To activate the methyl group, native methionine adenosyltransferases (*EcMAT*, Fig. 2b) convert ATP and



**Fig. 1** SAM-dependent methylation. **a** Reaction mechanism of S-adenosylmethionine (SAM)-dependent transmethylation of a nucleophilic substrate by methyltransferases (MTs) creating the byproduct S-adenosylhomocysteine (SAH). **b** Pharmaceuticals and natural compounds carrying “magic methyl” groups, which significantly influence potency. The relevant methyl in tazemetostat was chemically introduced during lead optimisation, C-methylation of gentamicin is mediated by a radical SAM MT, other methyl groups are introduced by conventional MTs



**Fig. 2** Metabolic network to convert formate into transferable methyl groups. **a** Established formate assimilation machinery consisting of *MexFTFL*, *MexMCH* and *MexMTDA* from *M. extorquens* (together abbreviated as “FCM”) as it is implemented for the generation of synthetic formatotrophic *E. coli*. **b** Metabolic pathway for the native generation of  $\text{C}_1$ - $\text{H}_4\text{F}$  metabolites from serine by *EcGLYA* and glycine by *EcGCS* for the supply of methyl groups to the SAM regeneration cycle. Heterologous introduction of an MT and supply of the MT substrate, here *RgANMT* with 2,5-ANP, enable formation of the methylated product and the byproduct SAH, which can be degraded to homocysteine either by *EcMTAN* and *EcLUXS*, or by heterologously introduced SAHH. The connection between formate assimilation and  $\text{C}_1$ - $\text{H}_4\text{F}$  supply for the SAM cycle via methylene- $\text{H}_4\text{F}$  is highlighted in light blue; genes deleted in this study are labelled in the corresponding colour of the strains shown in **(c)**. 2,5-ANP - 2,5-aminonitrophenol, ATP - adenosine-5'-triphosphate, ADP - adenosine-5'-diphosphate,  $\text{C}_1$  - one carbon, DHPD - 4,5-dihydroxy-2,3-pentanedione,  $\text{H}_4\text{F}$  - tetrahydrofolate, NAD(P)H - nicotinamide adenine dinucleotide (phosphate), M-2,5-ANP - N-methyl-2,5-aminonitrophenol, SAM - S-adenosylmethionine, SAH - S-adenosylhomocysteine, *EcGCS* - glycine cleavage system from *E. coli*, *EcGLYA* - serine hydroxymethyltransferase from *E. coli*, *EcLUXS* - S-ribosylhomocysteine lyase from *E. coli*, *EcMAT* - methionine adenosyltransferase from *E. coli*, *EcMETE/EcMETH* - methionine synthase from *E. coli*, *EcMETF* - methylene- $\text{H}_4\text{F}$  reductase from *E. coli*, *EcMTAN* - S-methyl-5'-thioadenosine/SAH nucleosidase from *E. coli*, *EcSERA* - 3-phospho-glycerate dehydrogenase from *E. coli*, *EcSERB* - phosphoserine phosphatase from *E. coli*, *EcSERC* - phosphoserine aminotransferase from *E. coli*, *MexFTFL* - formate- $\text{H}_4\text{F}$  ligase from *M. extorquens*, *MexMCH* - formyl- $\text{H}_4\text{F}$  cyclohydrolase from *M. extorquens*, *MexMTDA* - methylene- $\text{H}_4\text{F}$  dehydrogenase A from *M. extorquens*, *RgANMT* - anthranilate N-MT from *Ruta graveolens*

methionine to SAM as a first step of the SAM regeneration cycle [41]. The cofactor is subsequently utilised by MTs for methylation of their substrate, thereby, introducing the C<sub>1</sub>-compound into the product. Byproduct of the methylation reaction is *S*-adenosylhomocysteine (SAH), a known inhibitor of MTs [14]. In native *E. coli* metabolism, SAH is degraded in an irreversible reaction to adenine and *S*-ribosylhomocysteine by methyl thioadenosine/SAH nucleosidase (*EcMTAN*) [42]. *S*-ribosylhomocysteine is cleaved by *S*-ribosylhomocysteine lyase (*EcLUXS*) to homocysteine and the quorum sensing precursor 4,5-dihydroxypentane-2,3-dione (DHPD) [43]. Another option to degrade SAH are SAH hydrolases (SAHH, Fig. 2b) that cleave SAH to adenosine and homocysteine [44]. Transfer of a methyl group from methyl-H<sub>4</sub>F regenerates methionine from homocysteine allowing repetition of methylation with another C<sub>1</sub>-building block.

In order to monitor transformation of formate into the methyl groups of MT-derived products, a model MT reaction is needed, such as the anthranilate *N*-MT from *Ruta graveolens* (*RgANMT*, Fig. 2b). *RgANMT* transfers a methyl group in a SAM-dependent reaction onto the amino group of the unnatural substrate 2,5-aminonitrophenol (2,5-ANP) producing *N*-methyl-2,5-ANP (M-2,5-ANP, Fig. 2b) [26, 45]. The unnatural substrate and product undergo minimal degradation in *E. coli* and can freely permeate in and out of the bacterial cell, making this system a well-suited model for investigations on improving whole-cell methylation. Conversion rates of the model MT-substrate system were shown to be strongly dependent on the intracellular availability of methionine, resulting in a great potential to use it to explore the influence of formate on whole-cell methylation [26].

Recent research towards the application of formate mainly focussed on the incorporation of formate-derived carbon into central metabolites, either to allow growth or the production of small organic compounds [13, 46]. Reports on improving whole-cell methylation, on the contrary, focus on improving efficiency of the SAM regeneration cycle and intracellular availability of methionine, mostly omitting the supply of C<sub>1</sub>-units from

the C<sub>1</sub>-H<sub>4</sub>F metabolism [23, 25, 26, 47]. In this study we combine these two separated fields of research to investigate formate as a supplementary source of C<sub>1</sub>-units to increase regeneration of methionine and, therefore, whole-cell methylation with *E. coli*. We first investigated the sole whole-cell biotransformation of formate into transferred methyl groups using the described formate assimilation machinery FCM in combination with different MTs. We then continued using engineered *E. coli* strains (Fig. 2c) to allow higher formation of a methylated product by the supplementation of formate.

## Methods

### Chemicals

All chemicals were purchased in the highest available purity from certified vendors: <sup>13</sup>C-formate (>99% <sup>13</sup>C-enriched) was obtained as sodium formate from Sigma Aldrich, unlabelled sodium formate was obtained from Carl Roth chemicals; both were according to our knowledge not synthesised by electrochemical reduction. 2,5-Aminonitrophenol (2,5-ANP), 2,4-methoxy nitroaniline (2,4-MNA), 2,7-dihydroxynaphthalene (DHN), perchloric acid (70%), and formic acid (LC-MS grade) were purchased from Sigma Aldrich; NaCl, NH<sub>4</sub>Cl, MgSO<sub>4</sub>, KH<sub>2</sub>PO<sub>4</sub>, K<sub>2</sub>HPO<sub>4</sub>, CaCl<sub>2</sub>, agarose, LB-agar, LB-medium, D-(+)-glucose, D-(+)-xylose, glycerol, *iso*-propyl-β-thiogalactopyranosid (IPTG), kanamycin, chloramphenicol, and glycine were purchased from Carl Roth; 2×Phusion DNA polymerase master mix, acetonitrile (HPLC grade or LC-MS grade) were purchased from Thermo Fisher Scientific. Ethidium bromide was purchased from Merck KGaA, In-Fusion 5×master mix was purchased from Takara Bio Inc.; high fidelity restriction enzymes, CutSmart buffer, T4-DNA Ligase, 10×T4-DNA ligase buffer, and 1 kb Plus DNA ladder were purchased from New England Biolabs; spectinomycin was purchased from VWR International.

### Strains and plasmids

See Table 1.

**Table 1** Strains and plasmids used in this study (Gene sequences are given at the end of the SI)

Name	Genotype	Source
<i>E. coli</i> strains		
Stellar	<i>F</i> <sup>-</sup> , <i>endA1</i> , <i>supE44</i> , <i>thi-1</i> , <i>recA1</i> , <i>relA1</i> , <i>gyrA96</i> , <i>phoA</i> , $\Phi 80d$ <i>lacZ</i> $\Delta$ M15, $\Delta$ ( <i>lacZYA</i> — <i>argF</i> ) U169, $\Delta$ ( <i>mrr</i> — <i>hsdRMS</i> — <i>mcrBC</i> ), $\Delta$ <i>mcrA</i> , $\lambda$ –	Takara Bio Inc., Cat No. 636766
BL21(Gold)DE3	<i>F</i> – <i>ompT hsdS</i> ( <i>rB</i> – <i>mB</i> –) <i>gal dcm</i> $\lambda$ (DE3)	Novagen, Cat. No. 69450
ST18	<i>S17</i> $\lambda$ <i>pir</i> ( <i>pro thi hsdR</i> <sup>+</sup> <i>Tp</i> <sup>+</sup> <i>Sm</i> <sup>+</sup> ; chromosome::RP4-2 Tc::Mu-Kan::Tn7/ $\lambda$ <i>pir</i> ) $\Delta$ <i>hemA</i>	DSM 22074 [48,49]
SIJ488	<i>E. coli</i> K-12 MG1655 Tn7::para-exo-beta-gam; <i>prha</i> -FLP; <i>xylSpm-lscl</i>	[50]
BL21-AI (only used transfer <i>araB</i> ::T7RNAP-tetA)	<i>F</i> – <i>ompT hsdS</i> ( <i>rB</i> – <i>mB</i> –) <i>gal dcm</i> $\Delta$ <i>araB</i> ::T7RNAP-tetA	Invitrogen (Cat. No. 11540146)
BL21(Gold) DE3 $\Delta$ <i>metJ</i>	<i>F</i> – <i>ompT hsdS</i> ( <i>rB</i> – <i>mB</i> –) <i>gal dcm</i> $\lambda$ (DE3), $\Delta$ <i>metJ</i>	[22]
C <sub>1</sub> -Aux (C1A $\Delta$ 3)	SIJ488 $\Delta$ <i>gcvTHP</i> , $\Delta$ <i>glyA</i> , $\Delta$ <i>frmRAB</i>	[51]
C <sub>1</sub> -opt	C1-Aux $\Delta$ <i>pflAB</i> , $\Delta$ <i>araB</i> ::T7RNAP, <i>fcM</i>	This study
C <sub>1</sub> S	SIJ488 $\Delta$ <i>gcvTHP</i> , $\Delta$ <i>serA</i> , $\Delta$ <i>frmRAB</i> , $\Delta$ <i>ecaraB</i> ::T7RNAP, <i>fcM</i>	This study
C <sub>1</sub> -opt- <i>rganmt</i>	C <sub>1</sub> -opt harbouring pET22b:: <i>rganmt</i>	This study
C <sub>1</sub> S- <i>rganmt</i>	C <sub>1</sub> S harbouring pET28a:: <i>rganmt</i>	This study
C <sub>1</sub> S- <i>rganmt</i> + empty	C <sub>1</sub> S harbouring pET28a:: <i>rganmt</i> , pCDF-Duet-1	This study
C <sub>1</sub> S- <i>rganmt</i> + <i>ecmat</i>	C <sub>1</sub> S harbouring pET28a:: <i>rganmt</i> , pCDF:: <i>ecmat</i>	This study
C <sub>1</sub> S- <i>rganmt</i> + <i>ecmtan</i> + <i>ecluxS</i>	C <sub>1</sub> S harbouring pET28a:: <i>rganmt</i> , pCDF:: <i>ecmtan_ecluxS</i>	This study
C <sub>1</sub> S- <i>rganmt</i> + <i>ecmat</i> + <i>ecmtan</i> + <i>ecluxS</i>	C <sub>1</sub> S harbouring pET28a:: <i>rganmt</i> , pCDF:: <i>ecmat_T7_ecmtan_ecluxS</i>	This study
C <sub>1</sub> S- <i>rganmt</i> + <i>ecmetF</i>	C <sub>1</sub> S harbouring pET28a:: <i>rganmt</i> , pCDF:: <i>ecmetF</i>	This study
BL21- <i>fcM</i> - <i>rganmt</i>	BL21(GOLD)DE3 harbouring pFCM and pET28a:: <i>rganmt</i>	This study
BL21- <i>fcM</i> - <i>rganmt</i> - <i>ecmetF</i>	BL21(GOLD)DE3 harbouring pFCM, pET28a:: <i>rganmt_ecmetF</i> ,	This study
BL21- <i>fcM</i> - <i>rganmt</i> - <i>mmsahh</i>	BL21(GOLD)DE3 harbouring pFCM, pET28a:: <i>rganmt_mmsahh</i>	This study
Plasmids		
pET28a(+)	<i>Kan</i> <sup>R</sup> , <i>lacI</i> , <i>T7lacI</i> , <i>pBR322</i> , <i>RBS</i> <sub>77</sub>	Novagen, Cat. No. 69864
pET22b(+)	<i>Amp</i> <sup>R</sup> , <i>lacI</i> , <i>T7lacI</i> , <i>pBR322</i> , <i>RBS</i> <sub>77</sub>	Novagen, Cat. No. 69744
pCDF-DUET-1	<i>Str</i> <sup>R</sup> , <i>lacI</i> , <i>T7lac</i> , <i>CloDF13</i> , <i>RBS</i> <sub>77</sub>	Novagen, Cat. No. 71340
pKD3	<i>Amp</i> <sup>R</sup> <i>Cam</i> <sup>R</sup> , carries chloramphenicol resistance gene flanked by FRT ( <i>flippase</i> recognition target) sites	[52]
pKD4	<i>Amp</i> <sup>R</sup> <i>Kan</i> <sup>R</sup> , carries kanamycin resistance gene flanked by FRT ( <i>flippase</i> recognition target) sites	[52]
pDM4	<i>Kan</i> <sup>R</sup> , <i>Cam</i> <sup>R</sup> , <i>sacB</i> , <i>traJ</i> , <i>oriR6K</i>	[53,54]
pET28a:: <i>rganmt</i>	pET28a(+) vector containing <i>rganmt</i>	[26]
pET22b:: <i>rganmt</i>	pET22b(+) vector containing <i>rganmt</i>	This study
pET28a:: <i>rganmt_ecmetF</i>	pET28a(+) vector containing <i>rganmt</i> , <i>rbS</i> , <i>ecmetF</i>	This study
pET28a:: <i>rganmt_mmsahh</i>	pET28a(+) vector containing <i>rganmt</i> , <i>rbS</i> , <i>mmsahh</i>	This study
pCDF:: <i>ecmat</i>	pCDF-DUET-1 vector containing <i>ecmat</i> in MCS-1	[26]
pCDF:: <i>ecmtan_ecluxS</i>	pCDF-DUET-1 vector containing <i>ecmtan</i> , <i>rbS</i> , <i>ecluxS</i> in MCS-1	[26]
pCDF:: <i>ecmat_T7_ecmtan_ecluxS</i>	pCDF-DUET-1 vector containing <i>ecmtan</i> in MCS-1 and <i>ecmtan</i> , <i>rbS</i> , <i>ecluxS</i> in MCS-2	[26]
pCDF:: <i>ecmetF</i>	pCDF-DUET-1 vector containing <i>ecmetF</i> in MCS-1	This study
pFCM	pSC101 ori, <i>Strep</i> <sup>R</sup> , <i>P</i> <sub>pgi-mut</sub> <i>mexftfl</i> , <i>mexfch</i> , <i>mexmtda</i>	[40]
pDM4:SS9-C <sub>1</sub> M	pDM4 vector containing 600 bp up- and downstream homology to safe spot 9 <sup>ll55</sup> for the genome integration of <i>fcM</i> ( <i>P</i> <sub>STRONG</sub> - <i>RBS</i> <sub>C</sub> - <i>ftfl</i> - <i>RBS</i> <sub>C</sub> - <i>fch</i> - <i>RBS</i> <sub>C</sub> - <i>mtda</i> )	[13]

### ***E. coli* strain engineering**

All strains and plasmids used in this study are listed in Table 1. The strain C<sub>1</sub>-opt was constructed from the published C<sub>1</sub>-Aux (C<sub>1</sub>AΔ3) strain based on *E. coli* strain SIJ488 [50], a derivative of MG1655, engineered to carry inducible genes for λ-Red recombineering and flippase in its chromosome and C<sub>1</sub>S was directly constructed from *E. coli* strain SIJ488 [50]. Genome engineering was performed by λ-red recombineering [52] and P1 phage transduction [50,52]. In brief, the genes indicated in Table 1 for C<sub>1</sub>S and C<sub>1</sub>-opt were deleted using PCR-amplified linear DNA fragments carrying kanamycin or chloramphenicol resistance markers flanked by FRT-sites based on the pKD4 or pKD3 plasmids, respectively [52]. Resistance markers were amplified with primers carrying 50 bp overhangs serving as homologous regions flanking the target gene [56]. For removal of resistance markers from the chromosome, flippase was induced by addition of L-rhamnose as described previously [50]. Gene deletions and removal of resistance cassettes were verified by PCR.

For chromosomal insertion of the genes encoding the enzymes for formate assimilation (*MexFTFL*, *MexFCH* and *MexMTDA*, together referred to as FCM) [13,34], we used a previously described genome insertion protocol [53] and vector [13] containing the desired genes (“pDM4:SS9-C<sub>1</sub>M”, Table 1). In brief, a non-replicative plasmid (pDM4, R6K ori) was introduced into the respective recipient strain via conjugation from an *E. coli* ST18 donor strain, as described previously [48,49,53]. Kanamycin resistance was used to select for chromosomal insertion of the synthetic FCM operon (including the kanamycin resistance gene) via native homologous recombination based on 600 bp homology regions, with subsequent levansucrase (*sacB*) counter-selection.

For chromosomal insertion of the arabinose-inducible T7-RNA polymerase to create C<sub>1</sub>S and C<sub>1</sub>-opt, we transferred the corresponding genomic locus from

BL21-AI (Invitrogen) using P1 phage transduction. The transferred locus consists of an insertion replacing the *araB* gene, introducing genes encoding a tetracycline efflux pump (a selectable tetracycline resistance marker) and T7-RNA polymerase [57]. The latter is controlled by the endogenous P<sub>araBAD</sub> promoter, thus, achieving arabinose-inducible expression of T7-RNA polymerase. Successful transductants were selected via tetracycline resistance and the insertion site was verified by PCR.

### **Plasmid construction**

C<sub>1</sub>-opt already contains a kanamycin resistance, which is why the anthranilate *N*-MT from *Ruta graveolens* (*rganmt*) was cloned with the restriction enzymes *NdeI* and *HindIII* into an ampicillin resistance-containing pET22b(+) vector: Separately, 1 µg of the vectors pET28a::*rganmt* and pET22b(+) were digested with *NdeI* and *HindIII* at 37 °C for 2 h followed by separation by agarose gel electrophoresis (1% agarose, 100 V, 50 min) and clean-up of the excised bands of the cut *rganmt* insert and the cut pET22b(+) vector using the Promega Wizard™ Plus DNA Purification System. 50 ng of the linear vector were combined in a 1:3 ratio with the respective gene to be inserted and 2 µl of 5× In-Fusion buffer in a total volume of 10 µl nuclease free water. The mixture was incubated at 50 °C for 20 min followed by standard chemical transformation into *E. coli* Stellar cells, as recommended for the Takara Bio In-Fusion Cloning. The identity of each vector was confirmed by sequencing.

For co-overexpression on pET28a vector with BL21, the *ecmetF* gene and the SAH hydrolase gene from *Mus musculus* (*mmsahh*) were cloned behind *rganmt* by digesting pET28a::*rganmt* with *XhoI* and cloning of the respective insert into the linearised vector using the seamless In-Fusion Cloning technique (Table 2). When co-overexpression was conducted with pET28a and pCDF-DUET-1 vector in C<sub>1</sub>S, *ecmetF* was cloned into the first MCS of pCDF-DUET-1 (Table 2). For this purpose, the respective

**Table 2** Utilised oligonucleotides

Utilisation	Sequence
Linearisation of pCDF DUET-1 MCS-1 slot <i>fwrD</i>	TCTACTAGCGCAGCTTAA
Linearisation of pCDF DUET-1 MCS-1 slot <i>rev</i>	TATGTATATCTCCTTCTTATACTT
Cloning of a gene from pET28a into MCS-1 of pCDF-DUET-1 <i>fwrD</i>	CTTTCTGTTTCGACTTGCCGGATCTCAGTGGTGG
Cloning of a gene from pET28a into MCS-1 of pCDF-DUET-1 <i>rev</i>	AAGGAGATATACCATAAGAAGGAGATATACCATGGGCAGC
Cloning of <i>ecmetF</i> or <i>mmsahh</i> behind <i>rganmt</i> in pET28a after digestion with <i>XhoI</i> <i>fwrD</i>	GGTGGTGGTGCTCGATTTGTTAGCAGCCGGATCTCAG
Cloning of <i>ecmetF</i> or <i>mmsahh</i> behind <i>rganmt</i> in pET28a after digestion with <i>XhoI</i> <i>rev</i>	TGCGGCCGCACTCGATTAAGAAGGAGATATACCATGGGC
Constitutive promoter P <sub>pgi-mut</sub> from pFCM [40]	GTGGAATTATAGCATTTTAGCCTTTAATTGTCAATAGGT
	CTCGAGGTGAAGACGAAAGGGCCTCG
Constitutive promoter P <sub>STRONG</sub> from pDM4:SS9-C <sub>1</sub> M [13,53,58]	AATACTTGACATATCACTGTGATTCACATATAATATGCG
RBS <sub>C</sub> [53]	AAGTTAAGAGGCAAGA

inserts were amplified from a pET28a vector and the pCDF-DUET-1 vector was linearised via polymerase chain reaction (PCR, Table 2):

Primers with 15 bp overhangs complementary to the vector were ordered at Eurofins as lyophilised powder. The PCR was set up with 3  $\mu$ l fwd primer ( $c=10$  pM), 3  $\mu$ l rev primer ( $c=10$  pM, Table 2), 1  $\mu$ l of template (pET28a vector with the gene of interest or yeast gDNA), 18  $\mu$ l nuclease free water, and 25  $\mu$ l 2 $\times$ Phusion master mix (Thermo Fisher). The PCR was conducted with 5 min initial denaturation at 98  $^{\circ}$ C, followed by 29 cycles of 20 s denaturation at 98  $^{\circ}$ C, 20 s annealing at 55  $^{\circ}$ C and 35 s elongation at 72  $^{\circ}$ C. Afterwards, 5 min terminal elongation at 72  $^{\circ}$ C followed. The size of the gene was confirmed by agarose gel electrophoresis (1% agarose, 100 V, 50 min), followed by extraction of the band and clean up with the Promega Wizard<sup>TM</sup> Plus DNA Purification System. For insertion of the gene, the vectors were linearised with the respective primer pairs or with *Xho*I (Table 2). 50 ng of the linear vector were combined in a 1:3 ratio with the respective gene to be inserted and 2  $\mu$ l of 5 $\times$ In-Fusion buffer in a total volume of 10  $\mu$ l nuclease free water. The mixture was incubated at 50  $^{\circ}$ C for 20 min followed by standard chemical transformation into *E. coli* Stellar cells, as described for the Takara Bio In-Fusion Cloning. Identity of each vector was confirmed by sequencing.

#### Assay procedure to perform whole-cell methylation

For whole-cell methylation experiments, the *E. coli* strains were chemically (co-)transformed with the respective plasmids and then grown overnight at 37  $^{\circ}$ C on agar plates supplemented with the required antibiotics. A single colony was picked and used to inoculate 5 ml LB-medium (+ antibiotic) followed by overnight incubation at 37  $^{\circ}$ C, 170 rpm. The preculture was used to inoculate 25 ml LB-medium (+ antibiotic) in a 50 ml falcon tube in a ratio of 1:100. The culture was grown to an OD<sub>600</sub> of 0.5 – 0.7, overexpression was induced with 1 mM IPTG and took place at 20  $^{\circ}$ C, 140 rpm for 20 h (the lower temperature and shaking velocity was chosen to lower the stress level during protein production).

After overexpression, the cultures were chilled on ice for 10 min followed by centrifugation for 15 min, 4  $^{\circ}$ C, 2,500 g. The pellet was washed with 5 ml of ice cold M9-medium (48 mM Na<sub>2</sub>HPO<sub>4</sub>, 22 mM KH<sub>2</sub>PO<sub>4</sub>, 8.6 mM NaCl, 1.8 mM NH<sub>4</sub>Cl, 2 mM MgSO<sub>4</sub>, 0.1 mM CaCl<sub>2</sub>, and 22 mM glucose as

carbon source. NOTE: for experiments conducted for Figure SI 5 & SI 6 20 mM glycerol or 13 mM xylose were used as carbon source instead) followed by centrifugation (15 min, 4  $^{\circ}$ C, 2,500 g). The pellet was then resuspended to an OD<sub>600</sub> of approximately 4.5 in ice cold M9-medium containing 1 mM IPTG and required antibiotics. Assays were then pipetted on ice in 1.5 ml reaction tubes in a total volume of 1 ml followed by incubation at 37  $^{\circ}$ C and 170 rpm for 24 h. The whole-cell catalysts were used in a final OD<sub>600</sub> of 3.0 with 0.75 mM of the substrate to be methylated (2,5-ANP or DHN), 2.5–75 mM <sup>13</sup>C-formate and optionally 2 mM glycine. To terminate the assay after 24 h, 100  $\mu$ l of the assay mixture were taken and mixed with 35  $\mu$ l of 10% perchloric acid followed by centrifugation at 4  $^{\circ}$ C and 18,000 g for 30 min. The supernatant was taken and stored at –20  $^{\circ}$ C until analysis. For determination of conversion rates, HPLC–UV ( $\lambda=280$  nm) analysis was used and the share of <sup>13</sup>C-labelled methylated product was elucidated with LC–MS/MS.

#### Statistics

A t-test was performed to compare the mean product formation within 24 h across different conditions. A two-sided independent samples t-test assuming equal variances was applied. Statistical significance was defined as  $p < 0.05$ . All experiments were conducted in at least three biological replicates.

#### HPLC–UV analysis

HPLC–UV analysis ( $\lambda=280$  nm) was used to detect MT substrates and products generated during whole-cell methylation experiments and to calculate product formation. For analysis, 10  $\mu$ l of the sample were injected into an Agilent 1260 infinity II System equipped with a quaternary pump and the samples were separated on an ISAspher 100-3-C18 column (150 $\times$ 4.0 mm) with an elution gradient method using 0.1% formic acid in MilliQ water as mobile phase A and acetonitrile as mobile phase B at a flow rate of 1.0 ml  $\cdot$ min<sup>–1</sup> (s. Table SI 2). The method was initiated at 3% mobile phase B and held for 4 min, followed by an increase to 70% mobile phase B within the next 3 min. Within the following 0.1 min, mobile phase B was increased to 100% and kept at this level for another 2.9 min, followed by a decrease to 3% mobile phase B within 1 min and equilibration for another 3 min. Conversion was calculated by Eq. 1.

Equation 1. Calculation of conversion to methylated products from HPLC-peak area (area under curve, AUC).

$$\text{Conversion [\%]} = \frac{\text{AUC (product)}}{\text{AUC (product)} + \text{AUC (substrate)}} * 100 \quad (1)$$



introduction of the spectinomycin resistance encoded on pFCM, an adenylyltransferase [26]. Hereby enhanced ATP regeneration increases ATP availability for SAM synthesis and, hence, improves product formation [26]. Supplementation with 5 mM  $^{13}\text{C}$ -formate did not yield higher conversion for BL21-*rganmt* or BL21-*fcm-rganmt* (Fig. 3). Analysis with LC-MS/MS, however, revealed that 67% of the produced M-2,5-ANP were  $^{13}\text{C}$ -labelled in the experiment with BL21-*fcm-rganmt* (Fig. 3, Figure SI 1). This incorporation of formate-derived methyl groups did not take place in BL21-*rganmt* highlighting the necessity of pFCM for the assimilation of formate.

Despite introduction of formate-derived methyl groups and the, therefore, increased availability of  $\text{C}_1$ -building blocks in BL21-*fcm-rganmt* fed with  $^{13}\text{C}$ -formate, conversion rates did not improve as expected from literature reports [28]. As shown in our previous study, supplementation of BL21-*rganmt* with 5 mM methionine strongly increased product formation; therefore, neither MT acceptor substrate permeation into the bacterial cell, nor activity of RgANMT seemed to be the major limiting factors. We rather suspected a kinetic bottleneck either in the SAM regeneration cycle or in the reduction of methylene- $\text{H}_4\text{F}$  to methyl- $\text{H}_4\text{F}$  through EcMETE. We co-overexpressed *ecmat* and the SAH hydrolase gene from *Mus musculus* (*mmsahh*), which should improve SAH degradation through cleavage to adenosine and homocysteine; as well as the *ecmetF* gene together with *rganmt* encoded on a pET28a vector. Curiously, BL21-*fcm-rganmt-ecmat* did not produce any methylated product. An increased metabolic burden through overproduction of two enzymes could be one reason for the loss of methylation activity of this strain. Co-overexpression of *mmsahh* in BL21-*fcm-rganmt-mmsahh* lowered conversion rates slightly in combination with a decreased share of labelled methylated product. A small increase in conversion with BL21-*fcm-rganmt-ecmetF* (21% conversion) supplemented with 5 mM  $^{13}\text{C}$ -formate compared to BL21-*fcm-rganmt* (19% conversion) was observed (Fig. 3). Reduction of methylene- $\text{H}_4\text{F}$  to methyl- $\text{H}_4\text{F}$  by EcMETE, hence, seems to be a minor bottleneck for whole-cell biotransformation of formate to methyl groups. Nevertheless, the increase in methylated product was rather small, which is why we omitted co-overexpression in the following experiments. These first results demonstrated straight-forward conversion of formate into transferable methyl groups by whole-cell biotransformation using *E. coli*.

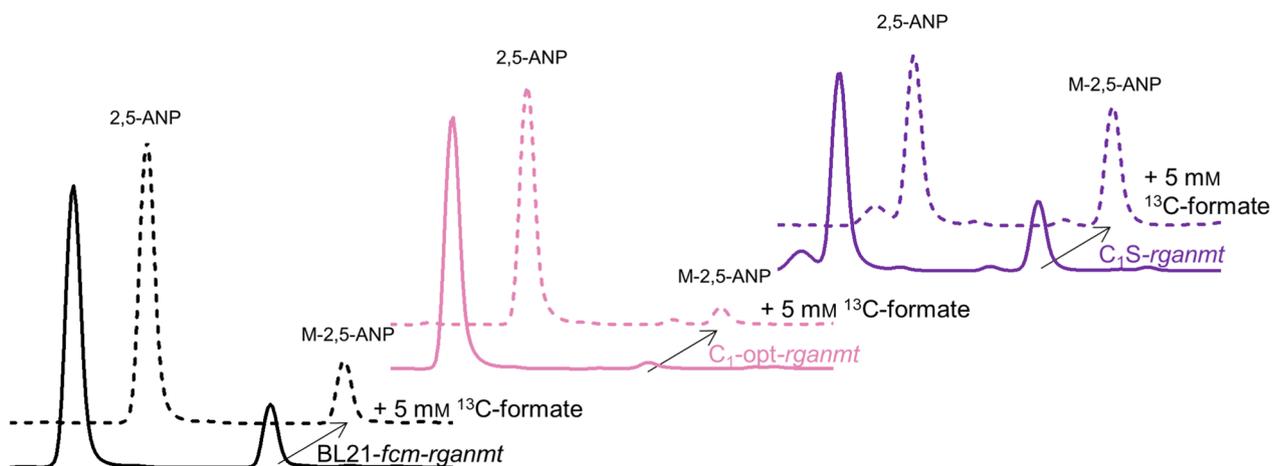
#### Chemoselective C- and O-methylation with formate-derived methyl groups

We further analysed, whether other MTs could also be used to methylate with formate-derived SAM. The

caffeic acid O-MT from *Prunus persica* (PpCaOMT) also accepts 2,5-ANP, however, catalyses an O-methylation to form 2,4-methoxy nitroaniline (2,4-MNA, Figure SI 2). While classic chemical methods struggle to selectively N- or O-methylate 2,5-ANP, the here described MTs were demonstrated to be highly selective [45]. As a third model reaction, the C-MT from *Streptomyces rishiriensis* (CouO) was chosen for the transfer of a methyl group onto the unnatural substrate 2,7-dihydroxynaphthalene (DHN), producing 1-methyl-DHN (M-DHN, Figure SI 2). The selectively methylated, functionalised MT-products are interesting as building blocks for chemical synthesis of small molecules such as active pharmaceutical ingredients. After the biotransformation experiment with supply of 5 mM  $^{13}\text{C}$ -formate, shares of  $^{13}\text{C}$ -labelled 2,4-MNA and M-DHN were 51% and 81%, respectively; again, without a substantial change in conversion rates (Figure SI 2). In a previous study, co-overproduction of EcMAT and CouO was found to strongly increase conversion rates; this was also the case here: Co-overproduction of EcMAT together with FCM and CouO increased product formation about 3.0-fold with and without the addition of formate. When 5 mM  $^{13}\text{C}$ -formate were supplied, a share of 68% of M-DHN was  $^{13}\text{C}$ -labelled (Figure SI 2). These results show the flexible applicability of whole-cell methylation with formate-derived methyl groups, using BL21-*fcm* in combination with different MTs.

#### Selecting engineered *E. coli* strains for improving methylation through the supply of formate

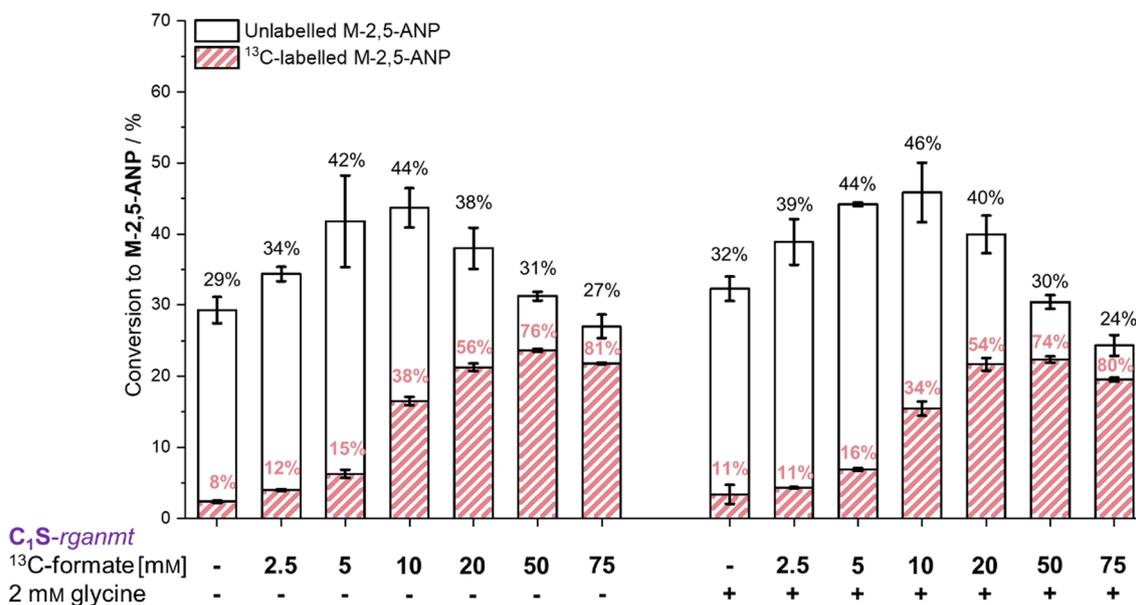
Introduction of formate-derived carbon into methyl groups worked successfully; however, conversion rates did not improve by the supply of formate. Initially, we suspected regulation through the transcription repressor EcMETJ to hamper reduction of formate to methyl groups by decreasing the expression of *ecmetF*, *ecmetE*, *ecmetH*, *ecmat* and genes involved in the synthesis of homocysteine [61,62]. Experiments with BL21 $\Delta$ *metJ*-*fcm* in combination with the MTs did not result in an increase of conversion rates when formate was supplemented (Figure SI 3). Furthermore, growth of the cultures and overall conversion decreased in comparison to BL21-*fcm*, suggesting that BL21 $\Delta$ *metJ*-*fcm* was not the optimal host to continue the investigations of transforming formate into methyl groups. Instead, we engineered strains based on *E. coli* K-12 MG1655. The strains' design was performed in accordance with the results of a previous study where Yishai et al. [40] (i) introduced the *fcm* genes on a plasmid under control of a constitutive promoter, and (ii) deleted genes involved in the native generation of  $\text{C}_1$ - $\text{H}_4\text{F}$  compounds, leading to excellent results in converting formate into the methyl group of methionine [40]. Yishai et al. achieved a  $\text{C}_1$ -auxotrophy



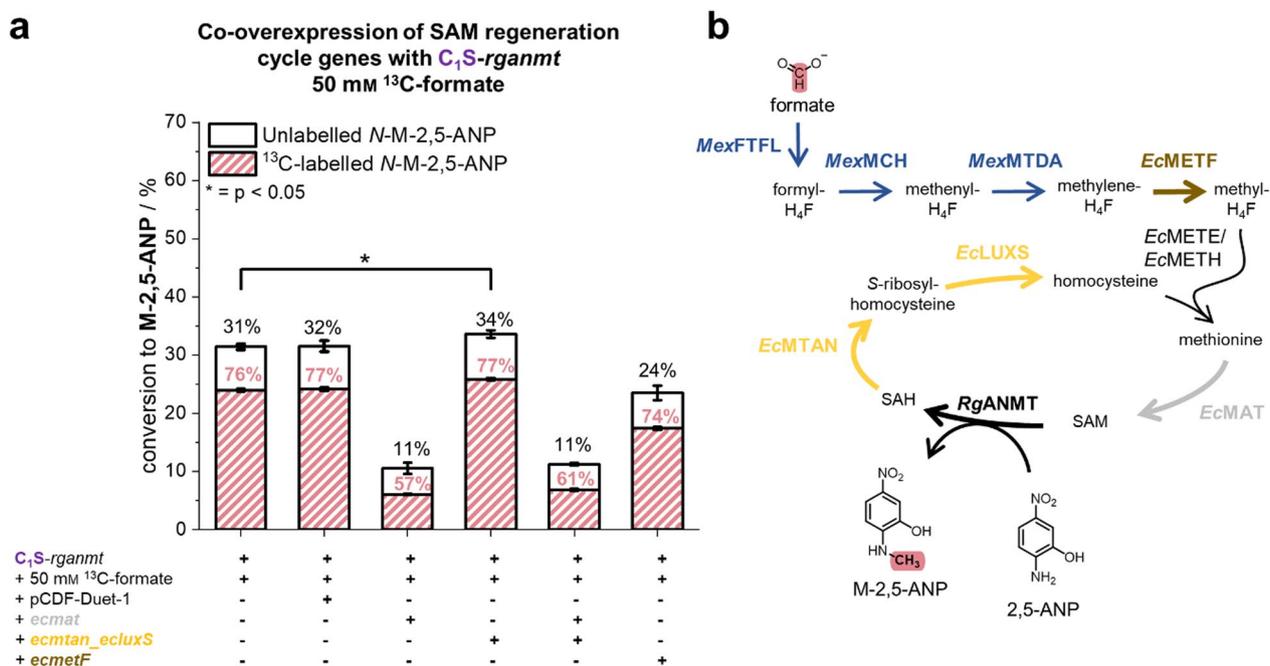
**Fig. 4** Whole-cell methylation by BL21-*fcm-rganmt*, *C*<sub>1</sub>-*opt-rganmt* and *C*<sub>1</sub>-*S-rganmt*. HPLC–UV (λ = 280 nm) chromatogram of whole-cell methylation experiments using BL21-*fcm-rganmt*, *C*<sub>1</sub>-*opt-rganmt* and *C*<sub>1</sub>-*S-rganmt* without (unbroken line) and with the addition of 5 mM <sup>13</sup>C-formate (dotted line). Experiments were conducted in M9-medium (22 mM glucose) with a final OD<sub>600</sub> of 3.0, 0.75 mM 2,5-ANP, and 5 mM <sup>13</sup>C-formate at 37 °C, 170 rpm for 24 h. 2,5-ANP - 2,5-aminonitrophenol, M-2,5-ANP - N-methyl-2,5-aminonitrophenol, *fcm* formate assimilation machinery genes from *M. extorquens*, *rganmt* - anthranilate N-MT gene from *R. graveolens*

in *E. coli* by deletion of the glycine cleavage system (*EcGCS*, encoded by *ecgcvTHP*) and serine hydroxymethyltransferase (*ecglyA*), which represent the main supply pathways for C<sub>1</sub>-H<sub>4</sub>F metabolites, forcing assimilation of formate for the generation of C<sub>1</sub>-H<sub>4</sub>F compounds [40]. In this study, *ecglyA* and *ecgcvTHP* were deleted to

construct the C<sub>1</sub>-auxotrophic strain *C*<sub>1</sub>-*opt*. Additionally, the serine-C<sub>1</sub> auxotrophic strain *C*<sub>1</sub>S was constructed, in which *ecgcvTHP* was deleted, *ecglyA* is still intact but the *de novo* biosynthesis of serine is not possible due to deletion of the 3-phospho-glycerate dehydrogenase (*ecserA*) [40]. For both strains, formate assimilation is enabled by



**Fig. 5** Increasing whole-cell methylation with formate in *C*<sub>1</sub>-*S-rganmt*. Given is the conversion of 2,5-ANP to M-2,5-ANP by *C*<sub>1</sub>-*S-rganmt* (black) and the corresponding share of <sup>13</sup>C-labelled product (red) after supply with 2.5–75 mM <sup>13</sup>C-formate with and without the addition of 2 mM glycine. Experiments were conducted in M9-medium (22 mM glucose) with a final OD<sub>600</sub> of 3.0, 0.75 mM 2,5-ANP, 2.5–75 mM <sup>13</sup>C-formate, and optionally 2 mM glycine at 37 °C, 170 rpm for 24 h. NOTE: About 8% of M-2,5-ANP is naturally <sup>13</sup>C-labelled, therefore respective shares are indicators, not absolutes for formate-derived methyl groups. No significant difference was found between supply of the same formate concentration with and without the addition of glycine. M-2,5-ANP - N-methyl-2,5-aminonitrophenol, *rganmt* - anthranilate N-MT gene from *R. graveolens*



**Fig. 6** Co-overexpression of SAM regeneration cycle genes in *C<sub>1</sub>S-rganmt*. **a** Co-overexpression of *rganmt* and adjuvant genes encoded on a pCDF-DUET-1 vector in *C<sub>1</sub>S*. Given are the conversion rates (black) and the relative share of <sup>13</sup>C-labelled product (red). **b** Scheme of the metabolic pathway from formate to methyl groups with co-overproduced proteins highlighted. Experiments were conducted in M9-medium, with 22 mM glucose, a final OD<sub>600</sub> of 3.0, 0.75 mM 2,5-ANP and 50 mM <sup>13</sup>C-formate at 37 °C, 170 rpm for 24 h. NOTE: About 8% of M-2,5-ANP is naturally <sup>13</sup>C-labelled, therefore respective shares are indicators, not absolutes for formate-derived methyl groups. \*p < 0.05, 2,5-ANP - 2,5-aminonitrophenol, H<sub>4</sub>F - tetrahydrofolate, M-2,5-ANP - N-methyl-2,5-aminonitrophenol, SAM - S-adenosylmethionine, SAH - S-adenosylhomocysteine, EcLUXS - S-ribosylhomocysteine lyase from *E. coli*, EcMAT - methionine adenosyltransferase from *E. coli*, EcMETE/ EcMETH - methionine synthase from *E. coli*, EcMETF - methylene-H<sub>4</sub>F reductase from *E. coli*, EcMTAN - S-methyl-5'-thioadenosine/SAH nucleosidase from *E. coli*, MexFTFL - formate-H<sub>4</sub>F ligase from *M. extorquens*, MexMCH - formyl-H<sub>4</sub>F cyclohydrolase from *M. extorquens*, MexMTDA - methylene-H<sub>4</sub>F dehydrogenase A from *M. extorquens*, RgANMT - anthranilate N-MT from *R. graveolens*.

genomic introduction of the *fcM* genes under control of a constitutive promoter ( $P_{STRONG}$ , Table 1), a genomic introduced arabinose-inducible T7 RNA-polymerase system allows overproduction of MTs from the pET28a vector (Fig. 2, Table 1) [57].

To elucidate basic conversion rates, we focussed on N-methylation with RgANMT and compared BL21-*fcM-rganmt*, *C<sub>1</sub>-opt-rganmt* and *C<sub>1</sub>S-rganmt*. Following the overexpression phase in LB-media, all strains methylate 2,5-ANP to M-2,5-ANP without the addition of formate as analysed with HPLC-UV and LC-MS/MS (Fig. 4, Table SI 1). Product formation with *C<sub>1</sub>S-rganmt* was 1.5-fold higher than with BL21-*rganmt*, while with *C<sub>1</sub>-opt-rganmt* only about 2% of the substrate were converted to the methylated product (Fig. 4, Table SI 1). Growth of the cultures in LB-medium presumably allowed a certain part of metabolites (e.g. methionine, SAM, or serine for *C<sub>1</sub>S*) to be carried over or reside in the cells enabling whole-cell methylation without the addition of formate. Supplying *C<sub>1</sub>-opt-rganmt* with 5 mM <sup>13</sup>C-formate led to a twofold increase of methylation

with a share of 54% <sup>13</sup>C-labelled product. As conversion rates were still only about 5%, this strain was not further investigated, as it was unable to compete with BL21-*fcM-rganmt* in terms of share of formate-derived methyl groups or total product formation. Product formation of *C<sub>1</sub>S-rganmt* increased 1.4-fold (42% conversion) when 5 mM <sup>13</sup>C-formate were added, clearly outperforming BL21-*fcM-rganmt-ecmetF* (21% conversion). To our surprise, in the experiment with *C<sub>1</sub>S-rganmt* only 15% of M-2,5-ANP were <sup>13</sup>C-labelled, indicating only a fraction of transferred methyl groups to be derived from formate. Due to the clear increase in conversion rate, a much higher share was expected. Still, in comparison to BL21-*fcM-rganmt*, whole-cell methylation with *C<sub>1</sub>S-rganmt* benefitted from the addition of formate which is why we continued focussing on this *C<sub>1</sub>*-auxotrophic strain.

#### Improving methylation with formate-derived methyl groups using *C<sub>1</sub>S-rganmt*

To evaluate the influence of formate concentration on product formation during the biotransformation phase,

formate concentrations ranging from 2.5 mM to 75 mM were screened (Fig. 5, Figure SI 4). Substantial changes in conversion started from 2.5 mM and peaked at 10 mM  $^{13}\text{C}$ -formate supply with a 2.1-fold higher product formation compared to BL21-*fcm-rganmt-metF* supplemented with 5 mM  $^{13}\text{C}$ -formate (Table SI 1). Interestingly, shares of labelled product with 10 mM  $^{13}\text{C}$ -formate did not exceed 38%. Further increase of  $^{13}\text{C}$ -formate supply lowered product formation with *C<sub>1</sub>S-rganmt* whereas the shares of labelled product increased, indicating that the reported toxic effects of formate outweigh its benefit at such high concentrations [63,64]. When 50 mM  $^{13}\text{C}$ -formate were supplied, 31% conversion with 76%  $^{13}\text{C}$ -labelled methylated product were detected, giving the highest total amount of labelled product.

In the *C<sub>1</sub>S* strain, *EcGLYA* is still present and could use methylene- $\text{H}_4\text{F}$  to synthesise serine from glycine when an excess of glycine is supplemented. On one hand, this reaction could withdraw  $\text{C}_1$ -compounds from the  $\text{C}_1$ - $\text{H}_4\text{F}$  metabolism, leading to decreased availability of  $\text{C}_1$ -carbon units for whole-cell methylation. On the other hand, allowing serine synthesis could increase cysteine concentrations and through the transsulfuration pathway also methionine availability, supporting whole-cell methylation [65]. To test this, we repeated the screening with 2.5–75 mM  $^{13}\text{C}$ -formate in combination with supply of 2 mM glycine during the biotransformation phase in M9 media (Fig. 5). Co-feeding of  $^{13}\text{C}$ -formate and glycine increased conversion rates slightly (2–5% higher conversion, not statistically significant) for every condition up to 20 mM formate. With further increase of the  $\text{C}_1$ -compound concentration, the small benefits of supplied glycine subsided. In most conditions, the share of  $^{13}\text{C}$ -labelled product was slightly lower when glycine was added, indicating that some of the  $^{13}\text{C}$ -labelled  $\text{C}_1$ -compound was indeed withdrawn from the  $\text{C}_1$ - $\text{H}_4\text{F}$  metabolism. While the differences observed were not statistically significant, the supply of 10 mM  $^{13}\text{C}$ -formate and 2 mM glycine still yielded the highest total conversion rate within this study (46% conversion with a share of 34%  $^{13}\text{C}$ -M-2,5-ANP). We assume that residing methionine from LB-media of the overexpression phase is responsible for the non-linear increase of labelled product with the addition of  $^{13}\text{C}$ -formate. Increasing the share of formate-derived methyl groups using moderate formate concentrations by exchanging media and carbon sources during overexpression did, however, not improve the total amount of formate-derived methyl groups (Figure SI 5 & Figure SI 6). We therefore focussed on improving product formation with *C<sub>1</sub>S-rganmt* and 50 mM  $^{13}\text{C}$ -formate to maintain a high incorporation of formate into the methylated product.

To do so, the genes of the SAM regeneration cycle as well as *ecmetF* were co-overexpressed on a separate pCDF-DUET-1 vector in *C<sub>1</sub>S-rganmt* (Fig. 6, Figure SI 7). Increasing SAM synthesis by co-overexpression of *ecmat* lowered conversion, as was seen during the experiments with BL21-*rganmt*. Similar results were obtained for the co-overexpression of *ecmetF*. Co-overexpression of *ecmtan* and *ecluxS* led to a slight increase in product formation (~2% higher conversion), indicating that removal of SAH and regeneration to homocysteine could be a minor factor limiting whole-cell methylation in *C<sub>1</sub>S-rganmt*.

## Discussion

In this study we report that *E. coli* can be used to transform formate into methyl groups by introducing the formate assimilation machinery of *M. extorquens*. Introduction of formate-derived carbon into methylated products was highly efficient in BL21-*fcm-rganmt* supplied with 5 mM formate while conversion was not improved. Product formation with *C<sub>1</sub>S-rganmt*, in contrast, benefited from the addition of 5–10 mM formate, while the shares of formate-derived methyl groups peaked at 50 mM formate supply with slight decrease in conversion. Due to the energy-intensive production of renewable formate, low supply of the  $\text{C}_1$ -compound is desirable, with the aim to (i) improve product formation and to (ii) incorporate formate-derived methyl groups into the products. The so far best performing strains *C<sub>1</sub>S-rganmt* or *C<sub>1</sub>S-rganmt-ecmtan-ecluxS* will be the starting point for future adjustments towards application.

We suspect carbon source-derived unlabelled  $\text{C}_1$ -compounds to be the cause for the non-linear incorporation of formate-derived methyl groups with increased formate supply using *C<sub>1</sub>S*. Suppressing the formation of unlabelled  $\text{C}_1$ -compounds should, hence, improve efficiency of whole-cell biotransformation of formate to methyl groups. One source of unlabelled formate could be the native pyruvate formate-lyase (*EcPFL*). *EcPFL* is an oxygen-sensitive enzyme and native expression of *ecpfl* occurs under anaerobic conditions providing a further electron sink for fermentation by cleavage of pyruvate to acetyl-CoA and formate [66]. Another possibility could be the production of intracellular formaldehyde from the added carbon source, which can be used by the unspecific threonine aldolase (*EcLTAE*) to synthesise unlabelled serine, representing a substrate for the production of unlabelled methylene- $\text{H}_4\text{F}$  by *EcGLYA* [67]. Deletion of the respective genes is planned for follow-up experiments to improve efficiency of formate incorporation at low formate concentrations using *C<sub>1</sub>S*.

The highest total conversion achieved with *C<sub>1</sub>S-rganmt* was 46%, which has to be improved in order to use the strain on a larger scale. In future experiments, time courses showing the conversion of the MT substrate to the (<sup>13</sup>C-)methylated product in combination with analysis of the labelled and unlabelled *C<sub>1</sub>-H<sub>4</sub>F* metabolites, methionine, and SAM will be an important prerequisite to carry out flux analyses for identification of substantial bottlenecks. Based on this, overexpression or introduction of auxiliary pathways to improve product formation can be further explored.

The assimilation of formate, its reduction to methyl groups and the synthesis of SAM is an energy-consuming process, demanding two equivalents of ATP and two equivalents of NAD(P)H. Synthesis of homocysteine from aspartate further requires three equivalents of ATP and two equivalents of NADH. This increased demand for energy can stress *E. coli* and induce an imbalance in cofactor supply [27]. Increasing NADH regeneration by implementation of, e.g. formate dehydrogenase could improve the intracellular availability of reducing equivalents in *E. coli* and, as a result, also conversion [13].

## Conclusion

Our study demonstrates the introduction of the renewable *C<sub>1</sub>*-compound formate into SAM-derived methyl groups. Utilisation of pFCM in BL21-*fcm-rganmt-ecmetF* allowed methylation with 67% of <sup>13</sup>C-labelled product, using the MTs *PpCaOMT* and *CouO* more than 50% of introduced methyl groups were <sup>13</sup>C-labelled. BL21-*fcm-rganmt* offers a simple opportunity to introduce formate's carbon into value-added small molecules. Supply of formate to the engineered serine-dependent *C<sub>1</sub>*-auxotrophic *E. coli* strain *C<sub>1</sub>S-rganmt* successfully increased conversion by 57% with a share of 34% of <sup>13</sup>C-labelled product; with this, formate can be used as a source for methyl groups. Starting from this, further bottlenecks of whole-cell methylation (e.g. time courses and flux analysis) can be investigated, and by the addition of auxiliary enzymes (e.g. formate dehydrogenase), and other optimisation steps (e.g. deletion of *ecpfl* or *ecfiae*), it should be possible to further increase the performance of the strains. In summary, we describe a simple strategy on how to engineer *E. coli* to use the renewable *C<sub>1</sub>*-compound formate to support SAM-dependent methylation for future applications.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12934-025-02674-4>.

Supplementary material 1.

### Acknowledgements

We thank Aaliya Afandi Lee and Adelheid Nagel for their skilful technical support. We thank Prof. Dr Nicholas Turner for providing the *E. coli* BL21Δ*metJ* strain. We thank Dr Hai He for providing progenitor strains used in the construction of *C<sub>1</sub>-opt* and *C<sub>1</sub>S*.

### Author contributions

MKFM and JNA designed the research idea, AS, SLM and TJE further refined the research idea, MKFM performed and analysed in vivo methylation experiments, AS prepared the T7 RNA polymerase arabinose inducible *C<sub>1</sub>-opt* and *C<sub>1</sub>S* strains. MKFM wrote and together with JNA revised the manuscript, all authors contributed together to the finalisation of the manuscript.

### Funding

Open Access funding enabled and organized by Projekt DEAL. This work was funded by a PhD scholarship of the Deutsche Bundesstiftung Umwelt (DBU, Grant Number 20021/742) to MKFM and the European Research Council (ERC, Grant Number 716966). JNA acknowledges funding through a Heisenberg Grant of the Deutsche Forschungsgemeinschaft (DFG/527572100).

### Availability of data and materials

All data generated or analysed during this study are included in this published article and its supplementary information files. The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare no competing interests.

#### Author details

<sup>1</sup>Institute of Pharmaceutical Sciences, University of Freiburg, Albertstr. 25, 79104 Freiburg, Germany. <sup>2</sup>Max Planck Institute for Terrestrial Microbiology, Karl-Von-Frisch-Straße 10, 35043 Marburg, Germany. <sup>3</sup>Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, 14476 Potsdam-Golm, Germany. <sup>4</sup>Department of Biochemistry, Charité Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität, Charitéplatz 1, 10117 Berlin, Germany. <sup>5</sup>LOEWE Center for Synthetic Microbiology (SYNMIKRO), Philipps University of Marburg, Marburg, Germany.

Received: 29 November 2024 Accepted: 9 February 2025

Published online: 07 March 2025

### References

- Klooy Y, Nilsson LJ, Palm E. Reaching net-zero in the chemical industry - a study of roadmaps for industrial decarbonisation. *Renew Sustain Energy Transit.* 2024;5:100075.
- Wollensack L, Budzinski K, Backmann J. Defossilization of pharmaceutical manufacturing. *Curr Opin Green Sustain Chem.* 2022;33:100586.

3. Bazzanella A, Krämer D. Technologies for sustainability and climate protection – chemical processes and use of CO<sub>2</sub>. DECHEMA eV. 2017.
4. Sheldon RA, Woodley JM. Role of biocatalysis in sustainable chemistry. *Chem Rev*. 2018;118(2):801–38.
5. Schaub T. CO<sub>2</sub>-based hydrogen storage: CO<sub>2</sub> hydrogenation to formic acid, formaldehyde and methanol. *Phys Sci Rev*. 2018;3(3):3.
6. Germer P, Andexer JN, Müller M. Biocatalytic one-carbon transfer – a review. *Synthesis*. 2022;54(20):4401–25.
7. Yishai O, Lindner SN, Gonzalez De La Cruz J, Tenenboim H, Bar-Even A. The formate bio-economy. *Curr Opin Chem Biol*. 2016;35:1–9.
8. Ünlü A, Duman-Özdamar ZE, Çaloğlu B, Binay B. Enzymes for efficient CO<sub>2</sub> conversion. *Protein J*. 2021;40(4):489–503.
9. Calzadiaz-Ramirez L, Meyer AS. Formate dehydrogenases for CO<sub>2</sub> utilization. *Curr Opin Biotechnol*. 2022;73:95–100.
10. Nattermann M, Wenk S, Pfister P, He H, Lee SH, Szymanski W, et al. Engineering a new-to-nature cascade for phosphate-dependent formate to formaldehyde conversion in vitro and in vivo. *Nat Commun*. 2023;14(1):2682.
11. Anastas PT. Green chemistry and the role of analytical methodology development. *Crit Rev Anal Chem*. 1999;29(3):167–75.
12. Anastas PT, Warner JC. Green chemistry: theory and practice. Oxford: Oxford University Press; 2000.
13. Kim S, Lindner SN, Aslan S, Yishai O, Wenk S, Schann K, et al. Growth of *E. coli* on formate and methanol via the reductive glycine pathway. *Nat Chem Biol*. 2020;16(5):538–45.
14. Struck AW, Thompson ML, Wong LS, Micklefield J. S-adenosyl-methionine-dependent methyltransferases: highly versatile enzymes in biocatalysis, biosynthesis and other biotechnological applications. *ChemBioChem*. 2012;13(18):2642–55.
15. Mordhorst S, Andexer JN. Round, round we go – strategies for enzymatic cofactor regeneration. *Nat Prod Rep*. 2020;37(10):1316–33.
16. Abdelraheem E, Thair B, Varela RF, Jockmann E, Popadić D, Hailes HC, et al. Methyltransferases: functions and applications. *ChemBioChem*. 2022. <https://doi.org/10.1002/cbic.202200212>.
17. Jones PA, Takai D. The role of DNA methylation in mammalian epigenetics. *Science*. 2001;293(5532):1068–70.
18. Cheung P, Lau P. Epigenetic regulation by histone methylation and histone variants. *Mol Endocrinol*. 2005;19(3):563–73.
19. Schönherr H, Cernak T. Profound methyl effects in drug discovery and a call for new CH methylation reactions. *Angew Chem Int Ed*. 2013;52(47):12256–67.
20. Kirman CR, Sweeney LM, Gargas ML, Kinzell JH. Evaluation of possible modes of action for acute effects of methyl iodide in laboratory animals. *Inhal Toxicol*. 2009;21(6):537–51.
21. Broderick JB, Broderick WE, Hoffman BM. Radical SAM enzymes: Nature's choice for radical reactions. *FEBS Lett*. 2023;597(1):92–101.
22. Galman JL, Parmeggiani F, Seibt L, Birmingham WR, Turner NJ. One-pot biocatalytic in vivo methylation-hydroamination of bioderived lignin monomers to generate a key precursor to L-DOPA. *Angew Chem Int Ed*. 2022;61(8):e202112855.
23. Liu Q, Lin B, Tao Y. Improved methylation in *E. coli* via an efficient methyl supply system driven by betaine. *Metab Eng*. 2022;72:46–55.
24. David B, Wolfender JL, Dias DA. The pharmaceutical industry and natural products: historical status and new trends. *Phytochem Rev*. 2015;14(2):299–315.
25. Luo ZW, Cho JS, Lee SY. Microbial production of methyl anthranilate, a grape flavor compound. *Proc Natl Acad Sci*. 2019;116(22):10749–56.
26. Mohr M, Benčić P, Andexer JN. Doping in vivo alkylation in *E. coli* by introducing the direct sulfurylation pathway of *S. cerevisiae*. *Angew Chem Int Ed*. 2024. <https://doi.org/10.1002/anie.202414598>.
27. Lv Y, Chang J, Zhang W, Dong H, Chen S, Wang X, et al. Improving microbial cell factory performance by engineering SAM availability. *J Agric Food Chem*. 2024;72(8):3846–71.
28. Okano K, Sato Y, Inoue S, Kawakami S, Kitani S, Honda K. Enhancement of S-adenosylmethionine-dependent methylation by integrating methanol metabolism with S-methyl-tetrahydrofolate formation in. *Catalysts*. 2020. <https://doi.org/10.3390/catal10091001>.
29. Gonzalez JC, Banerjee RV, Huang S, Sumner JS, Matthews RG. Comparison of cobalamin-independent and cobalamin-dependent methionine synthases from *Escherichia coli*: two solutions to the same chemical problem. *Biochem*. 1992;31(26):6045–56.
30. Simic P, Willuhn J, Sahn H, Eggeling L. Identification of glyA (Encoding Serine Hydroxymethyltransferase) and its use together with the exporter ThrE to increase l-threonine accumulation by *Corynebacterium glutamicum*. *Appl Environ Microbiol*. 2002;68(7):3321–7.
31. Kikuchi G. The glycine cleavage system: Composition, reaction mechanism, and physiological significance. *Mol Cell Biochem*. 1973;1(2):169–87.
32. Locasale JW. Serine, glycine and the one-carbon cycle: cancer metabolism in full circle. *Nat Rev Cancer*. 2013;13(8):572–83.
33. Sahr T, Ravel S, Rébeillé F. Tetrahydrofolate biosynthesis and distribution in higher plants. *Biochem Soc Trans*. 2005;33(4):758–62.
34. Wenk S, Rainaldi V, Schann K, He H, Bouzon M, Döring V, et al. Evolution-assisted engineering of *E. coli* enables growth on formic acid at ambient CO<sub>2</sub> via the serine threonine cycle. *Metab Eng*. 2025;88:14–24.
35. Kim S, Lee SH, Seo H, Kim KJ. Biochemical properties and crystal structure of formate-tetrahydrofolate ligase from *Methylobacterium extorquens* CM4. *Biochem Biophys Res Commun*. 2020;528(3):426–31.
36. Ochsner AM, Sonntag F, Buchhaupt M, Schrader J, Vorholt JA. Methylobacterium extorquens: methylotrophy and biotechnological applications. *Appl Microbiol Biotechnol*. 2015;99(2):517–34.
37. Vorholt JA, Chistoserdova L, Lidstrom ME, Thauer RK. The NADP-dependent methylene tetrahydromethanopterin dehydrogenase in *Methylobacterium extorquens* AM1. *J Bacteriol*. 1998;180(20):5351–6.
38. Yilmaz S, Kanis B, Hogers RAH, Benito-Vaquero S, Kahnt J, Glatter T, et al. System-level characterization of engineered and evolved formatotrophic *E. coli* strains. *bioRxiv*. 2024. <https://doi.org/10.1101/2024.11.11.622960>.
39. Sheppard CA, Trimmer EE, Matthews RG. Purification and properties of NADH-dependent 5, 10-methylenetetrahydrofolate reductase (MetF) from *Escherichia coli*. *J Bacteriol*. 1999;181(3):718–25.
40. Yishai O, Goldbach L, Tenenboim H, Lindner SN, Bar-Even A. Engineered assimilation of exogenous and endogenous formate in *Escherichia coli*. *ACS Synth Biol*. 2017;6(9):1722–31.
41. Berger B, Knodel M. Characterisation of methionine adenosyltransferase from *Mycobacterium smegmatis* and *M. tuberculosis*. *BMC Microbiol*. 2003;3(1):12.
42. Parveen N, Cornell KA. Methylthioadenosine/S-adenosylhomocysteine nucleosidase, a critical enzyme for bacterial metabolism: involvement of MTA/SAH nucleosidase in bacterial metabolism. *Mol Microbiol*. 2011;79(1):7–20.
43. Winzer K, Hardie KR, Burgess N, Doherty N, Kirke D, Holden MTG, et al. LuxS: its role in central metabolism and the in vitro synthesis of 4-hydroxy-5-methyl-3(2H)-furanone. *Microbiology*. 2002;148(4):909–22.
44. Shimizu S, Shiozaki S, Ohshiro T, Yamada H. Occurrence of S-adenosylhomocysteine hydrolase in prokaryote cells. *Eur J Biochem*. 1984;141(2):385–92.
45. Jockmann E, Subrizi F, Mohr MKF, Carter EM, Hebecker PM, Popadić D, et al. Expanding the substrate scope of N- and O-methyltransferases from plants for chemoselective alkylation\*\*. *ChemCatChem*. 2023;15(22):e202300930.
46. Reiter MA, Bradley T, Büchel LA, Keller P, Hegedis E, Gassler T, et al. A synthetic methylotrophic *Escherichia coli* as a chassis for bioproduction from methanol. *Nat Catal*. 2024. <https://doi.org/10.1038/s41929-024-01137-0>.
47. Kunjapur AM, Hyun JC, Prather KLJ. Dereglulation of S-adenosylmethionine biosynthesis and regeneration improves methylation in the *E. coli* de novo vanillin biosynthesis pathway. *Microb Cell Fact*. 2016;15(1):61.
48. Jackson SA, Fellows BJ, Fineran PC. Complete Genome Sequences of the *Escherichia coli* Donor Strains ST18 and MFDpir. *Microbiol Resour Announc*. 2020. <https://doi.org/10.1128/mra.01014-20>.
49. Thoma S, Schobert M. An improved *Escherichia coli* donor strain for diparental mating. *FEMS Microbiol Lett*. 2009;294(2):127–32.
50. Jensen SI, Lennen RM, Herrgård MJ, Nielsen AT. Seven gene deletions in seven days: fast generation of *Escherichia coli* strains tolerant to acetate and osmotic stress. *Sci Rep*. 2015;5(1):17874.
51. He H, Noor E, Ramos-Parra PA, Garcia-Valencia LE, Patterson JA, Díaz de la Garza RI, et al. In vivo rate of formaldehyde condensation with tetrahydrofolate. *Metabolites*. 2020;10(2):65.
52. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci*. 2000;97(12):6640–5.

53. Wenk S, Yishai O, Lindner SN, Bar-Even A. An engineering approach for rewiring microbial metabolism. In: Scrutton N, editor. *Methods in enzymology*. Cambridge: Academic Press; 2018. p. 329–67.
54. Milton DL, O'Toole R, Horstedt P, Wolf-Watz H. Flagellin A is essential for the virulence of *Vibrio anguillarum*. *J Bacteriol*. 1996;178(5):1310–9.
55. Bassalo MC, Garst AD, Halweg-Edwards AL, Grau WC, Domaille DW, Mutalik VK, et al. Rapid and efficient one-step metabolic pathway integration in *E. coli*. *ACS Synth Biol*. 2016;5(7):561–8.
56. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, et al. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol*. 2006. <https://doi.org/10.1038/msb4100050>.
57. Bhawasinghka N, Glenn KF, Schaaper RM. Complete genome sequence of *Escherichia coli* BL21-A1. *Microbiol Resour Announc*. 2020. <https://doi.org/10.1128/mra.00009-20>.
58. Braatsch S, Helmark S, Kranz H, Koeblmann B, Ruhdal JP. *Escherichia coli* strains with promoter libraries constructed by red/ET recombination pave the way for transcriptional fine-tuning. *Biotechniques*. 2008;45(3):335–7.
59. Lin B, Tao Y. Whole-cell biocatalysts by design. *Microb Cell Fact*. 2017;16(1):106.
60. de Carvalho CCCR. Whole cell biocatalysts: essential workers from Nature to the industry. *Microb Biotechnol*. 2017;10(2):250–63.
61. Marincs F, Manfield IW, Stead JA, McDowall KJ, Stockley PG. Transcript analysis reveals an extended regulon and the importance of protein–protein co-operativity for the *Escherichia coli* methionine repressor. *Biochem J*. 2006;396(2):227–34.
62. Shoeman R, Coleman T, Redfield B, Greene RC, Smith AA, Saint-Girons I, et al. Regulation of methionine synthesis in *Escherichia coli*: Effect of metJ gene product and S-adenosylmethionine on the in vitro expression of the metB, metL and metJ genes. *Biochem Biophys Res Commun*. 1985;133(2):731–9.
63. Nicholls P. Formate as an inhibitor of cytochrome c oxidase. *Biochem Biophys Res Commun*. 1975;67(2):610–6.
64. Warnecke T, Gill RT. Organic acid toxicity, tolerance, and production in *Escherichia coli* biorefining applications. *Microb Cell Fact*. 2005;4(1):25.
65. Aitken SM, Lodha PH, Morneau DJK. The enzymes of the transsulfuration pathways: active-site characterizations. *Biochim Biophys Acta BBA - Proteins Proteomics*. 2011;1814(11):1511–7.
66. Sawers G, Böck A. Novel transcriptional control of the pyruvate formate-lyase gene: upstream regulatory sequences and multiple promoters regulate anaerobic expression. *J Bacteriol*. 1989;171(5):2485–98.
67. He H, Höper R, Dodenhöft M, Marlière P, Bar-Even A. An optimized methanol assimilation pathway relying on promiscuous formaldehyde-condensing aldolases in *E. coli*. *Metab Eng*. 2020;60:1–13.

## Publisher's Note

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