Improved production of the non-native cofactor F_{420} in *Escherichia coli*

Mihir V. Shah1,2,4, Hadi Nazem-Bokaee1,2,4, James Antoney2,3, Suk Woo Kang2,3, Colin J. Jackson2,3 & Colin Scott1,2*

The deazaflavin cofactor F_{420} is a low-potential, two-electron redox cofactor produced by some Archaea and Eubacteria that is involved in methanogenesis and methanotrophy, antibiotic biosynthesis, and xenobiotic metabolism. However, it is not produced by bacterial strains commonly used for industrial biocatalysis or recombinant protein production, such as *Escherichia coli*, limiting our ability to exploit it as an enzymatic cofactor and produce it in high yield. Here we have utilized a genome-scale metabolic model of *E. coli* and constraint-based metabolic modelling of cofactor F_{420} biosynthesis to optimize F_{420} production in *E. coli*. This analysis identified phospho-enol pyruvate (PEP) as a limiting precursor for F_{420} biosynthesis, explaining carbon source-dependent differences in productivity. PEP availability was improved by using gluconeogenic carbon sources and overexpression of PEP synthase. By improving PEP availability, we were able to achieve a ~40-fold increase in the space–time yield of F_{420} compared with the widely used recombinant *Mycobacterium smegmatis* expression system. This study establishes *E. coli* as an industrial F_{420}-production system and will allow the recombinant in vivo use of F_{420}-dependent enzymes for biocatalysis and protein engineering applications.

Cofactor F_{420} is required for methanogenesis in Archaea1–5, anaerobic oxidation of methane by anaerobic methanotrophs6–9, and is involved in secondary metabolism in some Eubacteria10,11. As a deazaflavin, it is structurally similar to flavin, but given its lower redox mid-point potential (~360 mV for F_{420} cf. ~−230 mV for flavins) and obligate 2-electron transfer it functions analogously to NAD/NADP12,13. It has been suggested that cofactor F_{420}-dependent enzymes have significant potential as biocatalysts for the reduction of enoates, imines and ketones, and potentially for other unexplored reactions and processes14–17. However, the lack of a cost-effective production system for F_{420} is a major deterrent in exploring F_{420}-dependent reactions for biocatalytic applications. Development of a low-cost production system for F_{420} will be essential for advancing the application of F_{420}-dependent enzymes. An ultimate goal of this study is to devise the development of a system for F_{420} production as an end-product and at scale.

For research applications, cofactor F_{420} is produced via fermentation of organisms that naturally produce it, in particular several species of *Mycobacteria*15,18. However, *Mycobacteria* are not well suited for large-scale production of the cofactor because they are not generally recognised as safe (GRAS) organisms, tend to form dense aggregated “clumps”, and are slow growing19. Two 13-step chemical syntheses of F_{420} isomers have been reported, differing only in the peptide linkage between the two glutamate residues. Both reported extensive use of protecting groups and low overall yield20,21. While improved syntheses of the deazaflavin moiety F_{O} have been reported (Fig. 1)22, it is unlikely that a full chemical-synthesis route to F_{420} will be economical due to low yield, poor atom economy and the instability of several intermediates. Although some F_{420}-dependent enzymes have limited activity with F_{O}, it is unlikely to be a suitable substitute due to poor kinetics22. More recently, an efficient chemoenzymatic approach to producing the F_{120} analogue F_{120}P (phosphate group attached to F_{O}) was reported and F_{120}-dependent enzymes had substantially greater activity with F_{120}P than F_{O}, albeit still lower than with authentic F_{420}23.

The bacterial F_{420} biosynthesis pathway proposed by Bashiri et al.18 was further updated in this study. F_{O} synthase (CofGH), one of the first steps in the F_{420} biosynthesis pathway, was found to be a radical SAM enzyme.
capable of generating two molecules of 5'-deoxyadenosine. Therefore, F420 synthase accepts two molecules of S-adenosyl-L-methionine as substrate in addition to tyrosine, producing two molecules of L-methionine, one molecule of ammonia, and two molecules of 5'-deoxyadenosine (Fig. 1A).

Ideally, cofactor F420 would be produced via fermentation in a well-characterised microorganism, such as Escherichia coli, for which genetic and metabolic engineering tools are well developed and can be easily cultivated at large scale. Recently, we engineered the model laboratory bacterium E. coli to produce F420 via heterologous expression of biosynthetic genes sourced from Mycobacterium smegmatis (FbiD, FbiC, FbiB) and Methanosarcina mazei (CofD; equivalent to FbiA) (Fig. 1). The yield of cofactor F420 achieved in E. coli was 0.38 µmol/g DCW (grams of dry cell weight), which is comparable to yield obtained in wild type Mycobacterium smegmatis. Thermodynamic analysis of cofactor F420 biosynthesis revealed that the overall pathway is energetically favourable; with the final steps being effectively irreversible. This suggests that yields attained could be improved upon through metabolic engineering.

It has recently been shown that FbiD/CofC has species-specific substrate preferences (Fig. 1). FbiD from Paraburkholderia rhizoxinica prefers 3-phosphoglycerate (3PG)27 (Fig. 1, Step B3), M. smegmatis and M. mazei prefer phospho-enol pyruvate (PEP)28,29 (Fig. 1, Step B1), and Methanococcus jannaschii prefers 2-phospho-L-lactate (2PL)30 (Fig. 1, Step B2). It has been suggested that this step in the biosynthetic pathway may be particularly sensitive to the intramolecular concentration of its substrates, as thermodynamic analysis showed that this step is only just favourable in the forward direction. The diversity of substrates used in various organisms may reflect the adaptation of this step to use highly abundant metabolites. Preliminary findings suggest that variation of phospho-carbohydrate moiety has less effect on F420-dependent enzymes than variations of tail length.

Herein, we report that cofactor F420 biosynthesis in E. coli is heavily influenced by the carbon source. We used metabolic modelling to understand the underlying causes of carbon-source dependent differences in yield, which identified several potential bottlenecks. As a suitable genome-scale metabolic model was unavailable for production of F420 in E. coli, nor for any natural F420-producing organisms, we incorporated both the phospho-enol pyruvate-dependent and 3-phosphoglycerate-dependent F420 biosynthesis pathways into the iHK1487 genome-scale metabolic model for E. coli BL2131. The updated model (iEco-F420) was used to identify potential flux bottlenecks and to explore the theoretical limits of F420 production in this organism. Although the overall thermodynamics of the pathway are favourable, calculations revealed unfavourable energetics for the reaction catalysed by FbiD/CofC, which converts PEP and guanosine triphosphate into enolpyruvyl-diphospho-5'-guanosine. Several strategies were explored to improve availability of PEP for this reaction. Through a combination of metabolic engineering and rational carbon source selection, we were able to improve the yield of cofactor F420 from 0.28 to 1.60 µmol/g DCW. The highest productivity observed with E. coli was a yield of 1.60 µmol/g DCW and culture time of 13 h (equivalent to 123 nmol/h/g DCW); this space–time yield is fourfold higher than is in recombinant M. smegmatis, for which the highest published yield of cofactor F420 achieved was 3.0 µmol/g DCW with culture time of 96 h (equivalent to 31 nmol/h/gDCW).

**Results and discussion**

**The effect of different carbon sources on cofactor F420 yield and growth of E. coli.** To investigate the effects of different carbon sources on the production of F420, we tested acetate, fumarate, glucose, glycerol, pyruvate, and succinate as carbon sources, as these carbon sources enter central metabolism at different points, have varied uptake mechanisms and therefore distinct biogenic consequences for the cell. Pyruvate and fumarate, (followed closely by acetate and succinate) supported the greatest F420 production per gram of dry cell weight (DCW; Fig. 2A, B; Table 1). However, it should be noted that these carbon sources did not support high levels of biomass formation (Fig. 2C). Indeed, the cell density (measured as OD600) varied significantly by carbon source (Fig. 2C). With respect to overall productivity of F420 production (expressed as µmol F420/L/h), glycerol was the most productive carbon source (Fig. 2B); F420 yield with pyruvate was 0.90 µmol/g DCW which is close to the yield of the cofactor NADPH in E. coli of 1.3 µmol/g DCW (Fig. 2A). High F420 Yield and productivity with pyruvate implicitly indicated the impact of this intracellular metabolite as well as its precursor, PEP, on F420 biosynthesis.

**Phospho-enol pyruvate (PEP) is a key metabolite for F420 biosynthesis.** To systematically understand the effect of F420 biosynthesis on the distribution of flux through the entire metabolic network of the engineered E. coli grown with different carbon sources, we created and utilized the iEco-F420 metabolic model (see Methods) to compare flux profiles. Figure 3 summarises Flux Balance Analysis (FBA) results for two main pathways; glycolysis and the TCA cycle, for in silico growth with glucose, glycerol, and succinate as sole carbon sources, which were selected because of their different F420 productivity profiles. FBA predicted assimilation...
of 72% of glucose, as the sole carbon source, via the phosphoenolpyruvate (PEP): phosphotransferase system (PTS); all enzymes involved in glycolysis were active. Given the defined criteria, FBA predicted no flux through PEP synthase (PPS) or PEP carboxykinase (PPCK) indicating tight control over the pool of PEP during in silico growth with glucose (Fig. 3). These simulation results suggest a key role for PEP during F420 biosynthesis.

The metabolic model indicated that growth on succinate results in activation of the gluconeogenesis pathway and PPCK. With glycerol as the sole carbon source the upper glycolytic pathway was turned off (Fig. 3B), resulting in up to 27% higher overall ATP generation. On the other hand, fumarate was predominantly metabolised through aspartase since the glyoxylate shunt was highly active when glycerol was the carbon source, which leads to a reduction in total flux through TCA cycle. These modelling results explain the higher growth (Fig. 2C) and higher capacity for F420 production when engineered E. coli cells expressing the F420 biosynthetic pathway are grown with glycerol compared with glucose or succinate. Interestingly, with succinate as the carbon source, the iEco-F420 model predicted that pyruvate was produced mainly through malate dehydrogenase (decarboxylating) (Fig. 3A), leaving the PEP pool more accessible for incorporation into F420 production, consistent with the experimental yields. These results are consistent with the empirical growth experiments and also indicate a key role for PEP in controlling flux through the F420 biosynthesis pathway.

The iEco-F420 model contains 35 reactions consuming PEP: 19 are PEP-dependent phosphotransferases, 10 reactions participate in central carbon metabolism, two occur in cell envelope metabolism, two in tyrosine metabolism, and one in F420 biosynthesis (Supplementary File 1; Table S1). In an effort to increase the PEP pool, we used the model to test whether any of these competing reactions were dispensable in silico. However, single
gene deletion in silico predictions suggested that removing the reactions involved in cell envelope and tyrosine metabolism would result in cell death.

We next performed flux variability analysis (FVA) for all reactions in the metabolic network, including the PEP-consuming reactions (Supplementary File 1; Tables S2–S13) to specifically explore flux variations in PEP-consuming/producing reactions as a result of maximization of flux through biosynthesis of F420. PEP hydratase (enolase) was chosen to interpret flux variations with respect to PEP availability for cellular growth versus F420 production. Figure 4 shows the flux profile of PEP hydratase using all six carbon sources. When glucose is the sole carbon source, PEP must be produced through glycolysis to meet cellular objective (i.e., maximizing growth). At maximum biomass (where the blue and red lines showing minimum and maximum fluxes meet in Fig. 4), PEP hydratase flux is positive, meaning that 2-phosphoglycerate (2-PG) is fully metabolized to PEP. One engineering objective for increasing the heterologous production of F420 requires more carbon to be diverted into the target product rather than biomass, up to the point where the growth of the host is so negatively affected that it becomes uneconomical. When biomass yield drops to 80% of its maximum, for example, the minimum and maximum fluxes through PEP hydratase are still both positive, meaning that essential cellular processes take priority. As a result, 2-PG needs to be metabolized to provide stoichiometric requirements of PEP. However, at 50% of maximum biomass yield, the minimum flux (Fig. 4) through PEP hydratase becomes negative, meaning that the system is more relaxed to divert a portion of PEP for other processes including F420 production.

Unlike the flux predictions for PEP hydratase using glucose, PEP is significantly more available for processes other than cellular growth when the carbon source is succinate, fumarate or pyruvate, even at maximum biomass yields (Fig. 4). This is consistent with reports that glucose uptake in E. coli occurs primarily via the PTS, consuming up to 50% of the available PEP in cell35,36, thereby reducing its availability for F420 biosynthesis. Gluconeogenic

| Plasmid used | Carbon source | F420 Yield (µmol/g DCW) | F420 Concentration (µmol/L) | F420 productivity (µM/h) | Comments |
|--------------|---------------|-------------------------|-----------------------------|--------------------------|----------|
| pF420        | Glucose       | 0.28 ± 0.01             | 0.39 ± 0.06                 | 0.037 ± 0.001            | PEP-dependent F420 biosynthesis pathway |
| pF420        | Glycerol      | 0.66 ± 0.04             | 1.07 ± 0.06                 | 0.096 ± 0.004            |         |
| pF420        | Pyruvate      | 0.90 ± 0.08             | 0.91 ± 0.08                 | 0.067 ± 0.004            |         |
| pF420        | Acetate       | 0.74 ± 0.07             | 0.25 ± 0.04                 | 0.023 ± 0.004            |         |
| pF420        | Succinate     | 0.75 ± 0.06             | 0.40 ± 0.03                 | 0.014 ± 0.000            |         |
| pF420        | Fumarate      | 0.97 ± 0.03             | 0.22 ± 0.01                 | 0.015 ± 0.001            |         |
| pF420+PPS-pRSF | Glycerol    | 0.26 ± 0.00             | 0.49 ± 0.00                 | 0.050 ± 0.004            |         |
| pF420+PPS-pRSF | Pyruvate    | 0.53 ± 0.10             | 0.76 ± 0.15                 | 0.078 ± 0.001            |         |
| pF420+PPS-pRSF | Acetate     | 1.61 ± 0.11             | 2.33 ± 0.16                 | 0.175 ± 0.023            |         |
| pF420+PPS-pRSF | Succinate  | 0.73 ± 0.01             | 0.19 ± 0.01                 | 0.019 ± 0.002            |         |
| pF420+PPS-pRSF | Fumarate   | 0.30 ± 0.02             | 0.76 ± 0.05                 | 0.076 ± 0.001            |         |
| pF420+PPS-pRSF | Glycerol   | 0.80 ± 0.03             | 0.94 ± 0.03                 | 0.086 ± 0.001            |         |
| pF420+PPS-pRSF | Pyruvate  | 0.97 ± 0.09             | 1.03 ± 0.10                 | 0.070 ± 0.002            |         |
| pF420+PPS-pRSF | Acetate   | 0.81 ± 0.06             | 0.43 ± 0.03                 | 0.024 ± 0.002            |         |
| pF420+PPS-pRSF | Succinate | 0.32 ± 0.00             | 0.17 ± 0.00                 | 0.016 ± 0.004            |         |
| pF420+PPS-pRSF | Fumarate | 0.27 ± 0.00             | 0.07 ± 0.00                 | 0.005 ± 0.000            |         |
| pF420+PPS-pRSF | Glycerol | 0.19 ± 0.01             | 0.40 ± 0.01                 | 0.049 ± 0.003            |         |
| pF420+PPS-pRSF | Pyruvate | 0.87 ± 0.04             | 1.56 ± 0.07                 | 0.094 ± 0.008            |         |
| pF420+PPS-pRSF | Acetate | 0.77 ± 0.05             | 1.47 ± 0.10                 | 0.047 ± 0.005            |         |
| pF420+PPS-pRSF | Succinate | 0.25 ± 0.02             | 0.27 ± 0.02                 | 0.005 ± 0.000            |         |
| pF420+PPS-pRSF | Fumarate | 0.22 ± 0.01             | 0.04 ± 0.00                 | 0.002 ± 0.000            |         |
| pF420+PPS-pRSF | Glycerol | 0.27 ± 0.01             | 0.52 ± 0.02                 | 0.035 ± 0.005            |         |
| pF420+PPS-pRSF | Pyruvate | 0.36 ± 0.06             | 0.70 ± 0.11                 | 0.047 ± 0.005            |         |
| pF420+PPS-pRSF | Acetate | 0.87 ± 0.07             | 1.19 ± 0.10                 | 0.067 ± 0.002            |         |
| pF420+PPS-pRSF | Succinate | 0.66 ± 0.04             | 0.95 ± 0.05                 | 0.029 ± 0.005            |         |
| pF420+PPS-pRSF | Fumarate | 0.27 ± 0.04             | 0.27 ± 0.04                 | 0.003 ± 0.000            |         |
| pF420+PPS-pRSF | Glycerol | 0.25 ± 0.03             | 0.04 ± 0.01                 | 0.002 ± 0.000            |         |
| pF420+PPS-pRSF | Pyruvate | 0.21 ± 0.01             | 0.53 ± 0.00                 | 0.034 ± 0.003            |         |
| pF420+PPS-pRSF | Acetate | 0.38 ± 0.01             | 0.82 ± 0.02                 | 0.048 ± 0.002            |         |
| pF420+PPS-pRSF | Glycerol | 0.54 ± 0.01             | 0.88 ± 0.01                 | 0.059 ± 0.000            |         |
| pF420+PPS-pRSF | Pyruvate | 0.32 ± 0.03             | 0.53 ± 0.02                 | 0.014 ± 0.000            |         |

Table 1. Summary of the results obtained for engineered E. coli producing PEP-dependent or 3PG-dependent F420 using different carbon sources with and without overexpression of either PPS or PPCK.
carbon sources such as pyruvate, succinate, and fumarate increase intracellular PEP levels compared to glucose as their uptake is PEP-independent. PEP hydratase flux variation with glycerol is the highest among other carbon sources, meaning that glycerol assimilation could potentially lead to greater flexibility in utilising PEP for F420 biosynthesis. However, glycerol uptake occurs through the glycolysis pathway and although its uptake requires half the energy (in form of ATP) of glucose, most of the PEP is still required for cellular activities rather than biosynthesis of F420. Nonetheless, glycerol remains a candidate carbon source for large-scale F420 production compared with glucose when maintaining high cell masses is essential because it allows for higher cellular mass yields while bypassing PTS-dependent PEP depletion. In the case of acetate, ATP-dependent acetate assimilation is the only route for producing acetyl-CoA, which is an essential precursor for the biosynthesis of most amino acids and fatty acids and therefore biomass yield drops significantly (Fig. 4). However, FVA for PEP hydratase indicates the feasibility of utilising PEP for non-cellular activities.

We measured intracellular PEP for engineered E. coli grown with glucose and glycerol to validate the model predictions. When glycerol was used as the sole carbon source, PEP and F420 levels were 1.43 and 1.82-fold higher, respectively, compared with when glucose was used as the carbon source. This difference was also borne out in Figure 3. Flux balance analysis of TCA cycle (A) along with glycolysis/gluconeogenesis (B) pathways predicted by iEco-F420 metabolic model of E. coli for independent simulations using glucose, glycerol, or succinate as sole carbon sources (60 C-mol of carbon source). Objective is maximizing F420 production while maintaining growth at 30% of its max. Maintenance ATP requirements is fixed at 5.17 mmol/g DCW. Colormap shows absolute flux values in mmol/gDCW/hr. Fructose-bisphosphate aldolase, Triose-phosphate isomerase, Glyceraldehyde-3-phosphate dehydrogenase, and phosphoenolpyruvate hydratase are active in favor of gluconeogenesis pathway with succinate as the carbon source.
the simulation data (Supplementary Table S1). These results, collectively, demonstrate that the choice of carbon source directly affects intracellular availability of PEP, which, in turn, influences F420 levels.

**Using 3PG as an alternative to PEP.** As PEP is likely to be a flux-limiting metabolite, we explored the possibility of using an alternative metabolite in its place. Three different metabolites have been proposed to be incorporated in the sidechain of F420: PEP, 2-phospho-l-lactate and 3-phospho-d-glycerate18,27,37. While 2-phospho-l-lactate has not been observed in \textit{E. coli}18, 3-phosphoglycerate (3PG) is a glycolytic pathway intermediate present in \textit{E. coli} at 10 times the concentration of PEP38. Moreover, in the context of the iEco-F420 model, PEP-dependent F420 biosynthesis requires an additional FMN-dependent reduction step (the FbiB-dependent conversion of dehydro-F420-0 into F420-0) indicating that additional carbon would need to be diverted into FMN biosynthesis28,39. Preliminary evidence suggests that 3PG-F420, unlike FO and FOP, is accepted as a cofactor by F420-dependent enzymes with similar kinetics to standard F42030. Given the relative abundance of 3PG, we investigated it as an alternative to PEP by substitution of \textit{M. smegmatis} FbiD with that of \textit{P. rhizoxinica}.

Although 3PG is present at a higher intracellular concentration than PEP (1.5 mM cf. 0.18 mM)38 and is predicted to provide relatively similar maximum theoretical F420 yields (Supplementary Fig. S1), the experimentally determined yield of F420-3PG was found to be lower than for F420-PEP (Fig. 5). Moreover, no F420-3PG formation was observed with either succinate or fumarate as carbon source. This contrasts with the model predictions of feasible theoretical yields for F420-3PG with all carbon sources tested (Supplementary Fig. S1). It is possible that the \textit{P. rhizoxinica} Fbd product, glyceryl-2-diphospho-5′-guanosine, 3PG-F420-0 and/or its polyglutamated derivatives are poor substrates for the enzymes catalysing subsequent steps in F420 biosynthesis, which had been sourced from \textit{Mycobacteria} and may have low specificity for 3PG containing F420 metabolites (Fig. 5).

**Over-expression of PEP synthase increases the yield of F420.** Given that PEP is a limiting metabolite in F420 biosynthesis, we investigated whether production of PEP could be increased. Growth on fumarate and succinate is known to increase the expression of PEP-producing enzymes PPS and PPCK40 (Fig. 3A). Indeed, overexpression of PPS has been used to increase PEP concentrations in vivo41,42 to improve the yield of shikimic acid43, aromatic amino acids42,44 and lycopene45 biosynthesis. However, overexpressing PPS has been reported to negatively affect cell growth due to the excretion of pyruvate and acetate12.

---

**Figure 4.** Flux variability analysis showing PEP utilization capacity represented by the flux through PEP hydratase as a function of cellular biomass yield for different carbon sources predicted by the iEco-F420 metabolic model of \textit{E. coli}. Minimum and maximum fluxes are shown as blue and red, respectively. PEP: phosphor-enol pyruvate; 2PG: 2-phospho glycerate. Horizontal dashed lines do not correspond to any actual flux values; instead, they serve as hypothetical lines above which glycolysis drives carbon flow (2PG is metabolised to PEP).
In an attempt to increase intracellular PEP concentrations, we overexpressed PPS and PPCK from an IPTG-inducible expression plasmid and studied the effect on F420 yield. Consistent with previous reports, overexpression of PPS resulted in growth inhibition. Therefore, to improve final biomass concentration, PPS was only induced once cell density (OD600) was greater than 1.0, which resulted in significant improvement in F420 yield. We tested overexpression of PPS when grown on different carbon sources, as shown in Fig. 6. Overexpression of PPS improved the yield of F420 from 0.27 to 0.54 µmol/g DCW using glucose and from 0.53 to 0.80 µmol/g DCW using glycerol.

When grown on pyruvate, an F420 yield of 1.60 µmol/g DCW was observed without the addition of IPTG. With the addition of IPTG, the yield of F420 yield decreased to 0.90 µmol/g DCW. The yield of F420 also decreased after PPS induction when grown on succinate or fumarate (Fig. 6A). The pyruvate:PEP node of E. coli metabolism is highly regulated at both the transcriptional and metabolic levels, it is possible that PPS is metabolically regulated during gluconeogenesis or that the reversable flux through PPS is being driven thermodynamically towards pyruvate formation when grown on gluconeogenic carbon sources. With glucose and glycerol, induction of PPS with IPTG resulted in significant improvement in the yield of F420 as compared to non-induced PPS. On the contrary, with pyruvate, non-induced PPS resulted in significantly higher yield and productivity of F420 compared to IPTG induced PPS. It may be that optimal PPS expression levels differ with different carbon sources. The highest yield of F420 obtained was 1.60 µmol/g DCW, with a productivity of 0.17 µmol/h, using pyruvate as carbon source with leaky expression of PPS.

The impact of PPCK overexpression on F420 yield was also studied (Fig. 7). Unlike the expression of PPS, no improvement in F420 production was observed during PPCK overexpression. We confirmed the protein was expressed in soluble form (Supplementary Fig. S6). It is quite likely that we saw no difference in F420 concentration when PPCK was over-expressed because E. coli PPCK activity is metabolite controlled, either by the cellular PEP concentration or PEP:pyruvate ratio. We therefore investigated the potential of uncontrolled PPCK overexpression using the iEco-F420 model. We explored the overall capability of the metabolic network to improve flux through FbiB (i.e., production of mature F420) by simulating over-expression of PPS or PPCK. The results, shown in Supplementary Fig. S2, indicate that by forcing a higher flux through PPS or PPCK, the maximum FbiB flux (shown by black arrows) drops unless it occurs at a non-zero flux through PPS or PPCK. These results indicate the maximum stoichiometric capacity for F420 biosynthesis as a result of over-expressing PPS or PPCK; however, the overall kinetics of the system and regulatory mechanisms for growth with different carbon sources would significantly influence F420 yields, in vivo. The experimental results (Fig. 6) confirmed improved F420 biosynthesis when using glucose and pyruvate as a result of PPS overexpression, in agreement with the simulation results shown in Supplementary Fig. S2 for these carbon sources. It should be noted that the simulation results of Supplementary Fig. S2 also demonstrate the potential impact of the type of transporter on the flux through CofE when over-expressing PPS or PPCK.

The effect of time and carbon source on polyglutamate chain length. The final step in F420 biosynthesis is the addition of between one and nine glutamate residues to the F420-0 intermediate to yield F420-n (n: number of glutamate residues). What influences the tail length of F420 is still not clear, although in vitro analysis of F420-hexaglutamate ligases from different organisms has revealed that they typically produce F420 species with polyglutamate chain lengths consistent with F420 obtained from the native organisms. The number of glutamate residues influences the cofactor affinity of some F420-dependent enzymes; for example, the F420-dependent oxidoreductases MSMEG_2027, MSMEG_0777 and MSMEG_3380 from M. smegmatis reportedly having a high affinity for long chain F420 rather than shorter-chain F420. Similar effects are seen with polyglutamylated folates and folate mimics. Interestingly, F420-n composition changes with different growth conditions.
phases of *Methanosarcina barkeri* and *M. mazei*. We therefore investigated the composition of F420 during different growth phases of *E. coli*. The composition of F420 at various time points is shown in Fig. 8. When grown with glucose or glycerol as the carbon source, *E. coli* initially produced short chain F420-(1–4) in higher proportions, which shifted over time to predominantly longer chain F420-(5–8) (Fig. 8A, B). CofE from *M. smegmatis* (the enzyme used in this system) has been shown to produce predominantly longer F420 species (5–8) in stationary phase50.

Interestingly, we found that the tail length distribution at the end of the exponential phase was influenced by the carbon source used (Fig. 8C, D). Growth on succinate yielded the highest proportion of long chain F420 with F420-(5–8) comprising >90%. Glycerol has the next highest proportion of F420-(5–8) at >80%, with glucose and acetate with the lowest levels of F420-(5–8) produces (<30% and <25%, respectively) (Fig. 8C). The iEco-F420 model was used to guide interpretation of carbon source-dependent tail length distribution. According to the cofactor biosynthesis pathway shown in Fig. 1, two molecules of GTP per molecule of glutamate are required to metabolise an F420 molecule with only one glutamate residue. Likewise, in an ideal case where all incoming carbon to the *E. coli* has to end up in F420 with only one glutamate residue, the iEco-F420 model predicted that the ratio of sum of fluxes through all glutamate-producing reactions (\(\nu_{\text{glu}}\)) to sum of fluxes through all GTP-producing reactions (\(\nu_{\text{gtp}}\)) has to be equal to two regardless of the type of carbon source. However, when the model was used to simulate F420 biosynthesis with chain length compositions observed experimentally, flux predictions suggested deviations in the ratio of \(\nu_{\text{glu}}\) to \(\nu_{\text{gtp}}\) which depends on the type of carbon source. Interestingly, the ratio of \(\nu_{\text{glu}}\) to \(\nu_{\text{gtp}}\) was predicted to be 1.731 and 1.772 using succinate and glycerol, respectively, showing the largest deviation for a ratio of two. On the other hand, the ratio of \(\nu_{\text{glu}}\) to \(\nu_{\text{gtp}}\) was predicted by the model to be 1.994, 1.960, and 1.873 using glucose, acetate, and pyruvate, respectively, explaining why the lowest proportion of long chain F420 was observed with these carbon sources.

---

**Figure 6.** Effect of PPS over-expression on (A) F420 yield with and without IPTG (0.1 mM) induction; (B) F420 productivity (µM/h) with and without IPTG (0.1 mM). Leaky expression of PPS was observed without the addition of IPTG (Supplementary Fig. S3). *E. coli* was cultivated in a minimal media with different sole carbon sources as shown. Error bars show standard errors of the mean of at least two replicates.
3PG-F₄₂₀ yielded significantly higher fraction of short chain F₄₂₀-(1–4) > 70% (Fig. 8D) compared to PEP derived F₄₂₀ irrespective of the carbon source used. This could be due to the difference in the kinetics of the enzymes for 3PG-F₄₂₀ and PEP-F₄₂₀.

The iEco-F₄₂₀ metabolic model additionally provided some insights into the energetic differences in F₄₂₀ biosynthesis with only one glutamate tail as well as with varying number of glutamate tails. For all carbon sources examined, yields were higher for F₄₂₀ with only one glutamate than those for a mixture of F₄₂₀ molecules with different chain-lengths. This is because at a fixed growth rate (i.e., constant cell mass yield), total energy production (in the form of ATP) is higher for biosynthesis of F₄₂₀ with one glutamate than that for biosynthesis of a mixture of F₄₂₀ molecules (Fig. 9). Based on the results illustrated in Fig. 9, glucose maintains the highest cell mass per mole of ATP produced by cells, which explains the low F₄₂₀ yield from glucose compared to other carbon sources as shown in Fig. 2A. Assimilation of acetate as the sole carbon source requires the activation of ATP-dependent acetate kinase. Therefore, cells have to produce ATP in order to uptake carbon source for survival, which results in low growth rates (Fig. 2C) and maintaining the lowest cell mass yield per mole ATP produced among other carbon sources (Fig. 9) but, relatively high F₄₂₀ yields (Fig. 2A; Supplementary Fig. S1). According to the modelling predictions, acetate might provide benefits from industrial perspective because, cells would be forced to produce ATP for fueling F₄₂₀ production rather than for their growth.

**Conclusion**

This work establishes that intracellular PEP concentration is the key limiting metabolic bottleneck for heterologous F₄₂₀ biosynthesis in E. coli, at least when using biosynthetic enzymes, such as those from M. smegmatis, that natively use PEP. An updated metabolic model of E. coli incorporating the recombinant F₄₂₀ biosynthetic pathway was developed and used to identify differences in metabolic flux distribution through the entire metabolic network.
Figure 8. HPLC-FLD data for different glutamate residues attached to F420 at different time points after induction of F420 pathway using Glucose (A) and Glycerol (B). HPLC-FLD data for different glutamate residues attached to F420 (C) and 3PG-F420 (D) using different carbon sources at the end of exponential phase.

Figure 9. Theoretical biomass yields (at 30% maximum growth) with respect to total energy produced in the form of ATP predicted by the iEco-F420 metabolic model of E. coli BL21 simulated with different carbon sources. F420-1 indicates the yields for cells synthesizing cofactor F420 with only one glutamate tail, whereas F420-1 to 8 indicate those for cells synthesizing a mixture of F420 molecules with varying number of glutamate tails. The uptake of C-source was fixed to 60 C-mol in all simulations to account for differences in number of carbon atoms in C-sources.
network including the central carbon metabolism using various carbon sources. This allowed us to identify, test and rationalize a number of approaches to improve F420 yield. Table 1 summarises the F420 yield, concentration, and productivity obtained using different strains and conditions in this study. In terms of productivity, glycerol was found to be the best carbon source for F420 production amongst those tested as it allowed the best balance between optimizing PEP concentration (compared to glucose) vs. slower growth (compared with acetate, succinate, etc.). We also examined whether alternative substrates for FbiD could be used to remove the reliance on PEP, showing that although 3PG is a viable substrate and is more abundant in E. coli, the use of the P. rhizoxinica FbiD did not result in higher 3PG-F420 yields (compared with F420), presumably because of the downstream enzymes, which are sourced from M. smegmatis, have a preference for F420 metabolites over 3PG-F420 metabolites. Replacing other enzymes in the pathway with those from P. rhizoxinica may be a worthwhile strategy to improve 3PG-F420 yield, as the theoretical yield of 3PG-F420 is similar to that of PEP-derived F420 and strategies have already been developed to improve 3PG availability by increasing glycolytic flux by knocking out the zwf gene involved in first step of pentose phosphate pathway[37]. Other strategies that we tested to increase F420 yield included increasing PEP production through over expression of PPS; this produced the highest productivity with pyruvate. Our results also indicate that F420 composition and concentration in E. coli is comparable to some of the best natural producers such as M. smegmatis. The F420 yield obtained with M. smegmatis is approximately 0.30 µmol/g DCW in wild type[27] and 3.0 µmol/g DCW in engineered M. smegmatis[26]. However, cultivation of M. smegmatis in shake flasks takes up to 96 hours[40], while engineered E. coli cultivation takes 16 h (this study). This results in space–time yield of 114 nmol/h/gDCW for the recombinant E. coli system, vs. 31 nmol/h/gDCW for the recombinant M. smegmatis system[33]. This increased yield is in addition to the many other advantages of using E. coli, including less expensive antibiotics (ampicillin vs. hydromycin), reduced safety risks and thus greater accessibility to the technology, and the fact that E. coli is a much more widely used strain for protein engineering and DNA modification. By systematically optimizing growth and production conditions in E. coli, we have created a system that should make production of F420 more economical at an industrial scale and the study of F420-dependent enzymes more accessible, improving on the previous attempts for F420 production through the use of more exotic and challenging bacterial species, such as M. smegmatis.

**Materials and methods**

**Bacterial strain, vector and media composition.** E. coli BL21 DE3 (New England Biolabs) was used for protein expression and DH5α (New England Biolabs) strain was used for plasmid propagation. LB media consisting of 10 g/L tryptone, 5 g/L yeast extract and 10 g/L of NaCl was used for the plasmid propagation and cloning. 15 g/L of agar was added to prepare LB agar plates. E. coli protein expression and F420 production studies were done in M9 minimal media consisting of 6.78 g/L Na2HPO4, 3 g/L KH2PO4, 1 g/L NH4Cl, 2.5 g/L NaCl, 100× trace elements solution (1.667 g/L FeCl3·6H2O, 0.018 g/L ZnSO4·7H2O, 0.012 g/L CuCl2·2H2O, 0.012 g/L MnSO4·H2O, 0.018 g/L CoCl2·6H2O, and 2.225 g/L Na2EDTA·2H2O), 0.246 g/L MgSO4·7H2O, 0.011 g/L CaCl2, and 5 g/L of glucose or 0.167 C·mol of different carbon sources with the final pH of 7.3. Chloramphenicol (25 µg/ml), kanamycin (50 µg/ml) and ampicillin (100 µg/ml) were added where appropriate.

E. coli was cultivated in shake flasks at 30 °C with shaking at 200 rpm. The F420 biosynthesis genes in pSB1C3 plasmid were induced with 200 ng/ml of tetracycline. E. coli with pRSF duet plasmid was induced with IPTG (0.1 mM) for the expression of PPS and PEP carboxykinase (PPCK) enzymes. The biomass was measured using optical density at 600 nm.

**Plasmid construction.** The synthesis of the plasmid expressing the F420 biosynthesis pathway operon was described in Bashiri et al.[15]. The pathway consists of CoD (Accession number: Q8PVT6), CoE (A0QTG1), CoF (A0QUZ4) and CoGH (NC_008596) genes under the control of the tetracycline-inducible promoter Bba_R0040[29] and the artificial terminator Bba_B1006[30]. The F420 biosynthesis operon had been previously synthesized by GenScript and cloned into pSB1C3 containing the constitutive tetracycline repressor cassette Bba_K145201 with EcoRI/XbaI and PstI/SpeI restriction enzymes, plasmid construction is explained in Bashiri et al.[15]. This construct, hereafter referred to as pF420, enables production of F420 to be induced by the addition of tetracycline. All the gene sequences except CoD were obtained from *Mycobacterium smegmatis*. CoD gene sequence was obtained from *Methanosarcina mazei*. All the genes were codon optimized for expression in E. coli BL21 DE3 strain. All F420 proteins were FLAG tagged and soluble expression of F420 pathway proteins was confirmed using western blot[15]. In order to produce 3PG-F420, the *M. smegmatis* CoF homologue was replaced by the homologue from *Paraburkholderia rhizoxinica* (ESASS2). This F420 biosynthesis operon was synthesised (Biomatik) in two fragments (P_rhizo_3PG-F420-1 and P_rhizo_3PG-F420-2). P_rhizo_3PG-F420-1 was a 3.2 kb flanked with EcoRI and BamHI restriction sites in pUC57, P_rhizo_3PG-F420-2 was a 3.6 kb flanked with BamHI and PstI restriction sites in vector pUC57. Three-way ligation was performed to ligate P_rhizo_3PG-F4201 (cut with EcoRI and BamHI), P_rhizo_3PG-F4202 (BamHI and PstI) and pF420 (cut with EcoRI and PstI). The ligation resulted in plasmid pF420-3PG (Supplementary Table S2).

The *ppsA* gene encoding PPS from *E. coli* (P23538) and *pck* gene encoding PPCK from *E. coli* (B5YT3) were synthesized and cloned into pETC22[30] plasmid flanked by Ndel and BamHI restriction sites (Twist Bioscience) to produce PPS-pETC22 and PPCK-pETC22 plasmids (Supplementary Table S2). pF420 have mutated pUC57 origin of replication which causes high copy number[38] and PETCC2 have pBR322 origin of replication, which are incompatible for co-transformation. Therefore, genes encoding PPS and PPCK were cloned into pRSF duet in order to have replicative compatibility with pF420 vector. PETCC2 and pRSF duet plasmids were digested with restriction enzymes Ndel and BamHI, the genes were gel purified from PETCC2 digestion and were ligated to pRSF duet plasmid. The pRSF duet plasmid containing the genes encoding PPS (PPS-pRSF) or PPCK (PPCK-pRSF) were used to transform BL21 DE3 containing pF420 plasmid using electroporation. Transformed cells
were selected on LB agar plates supplemented with chloramphenicol and kanamycin. Supplementary Table S2 summarises the list of plasmids used in this study.

Expression of F420 pathway proteins in pF420 had been confirmed previously using immunoblotting. The procedure is explained in detail by Bashiri et al.18. Soluble expression of PPS, PPCK and F420-3PG operon proteins was confirmed through SDS-PAGE gel (Supplementary Fig. S3).

**Analytical methods**

F420 detection, quantification and chain length measurement. Production of F420 in E. coli and its detection was confirmed using LC–MS, as previously reported18. For the quantification of F420 from E. coli, 1 ml of sample from shake flask cultivation was taken and centrifuged at 10,000g for 2 min. The pellet was resuspended in 120 µl of 75% ethanol, boiled for 3 min at 94 °C to lyse the cells, resuspension was centrifuged at 10,000g for 2 min and the fluorescence of 100 µl supernatant was measured in a SpectraMax M3 (Molecular Devices) 96-well plate spectrophuorometer (excitation at 420 nm and emission at 480 nm). Fluorescence correlated directly with F420 concentration in the cell lysate. F420 amount relative to the biomass concentration was obtained in fluorescence (Units)/Biomass (OD600). The correlation between Biomass (OD600) and dry cell weight (DCW) was 0.56 g/L DCW for OD600 of 1.0. There was a linear correlation between absorbance at 420 nm and fluorescence (420 nm excitation and 480 nm emission). The fluorescence values of the cell lysate were converted to absorbance, and the extinction coefficient of 41.4 mM−1/cm was used60 to convert fluorescence unit of cell lysate to mM of F420 in cell lysate. Using these parameters fluorescence (units)/Biomass (OD600) was converted to µmol F420/g DCW.

Analytical separation of F420 species based on the length of its glutamate tail is shown in Supplementary Fig. S4 and was achieved with an ion-paired reverse phase HPLC-FLD protocol as reported previously64. The supernatant was run on an Agilent 1200 series HPLC system equipped with an Agilent fluorescence detector and an Agilent Poroshell 120 EC-C18 2.1 × 50 mm 2.7 mm column. The system was run at a flow rate of 0.5 ml/min and the samples were excited at 420 nm and emission was detected at 480 nm. A linear gradient of two buffers was used: Buffer A, containing 20 mM ammonium phosphate, 10 mM tetraethylammonium phosphate, pH 7.0. Buffer B, 100% acetonitrile. A gradient was run from 25% to 40% buffer B as follows: 0–1 min 25%, 1–10 min 25%–35%, 10–13 min 35%, 13–16 min 35–40%, 16–19 min 40%–25%.

Intracellular PEP was measured using single ion monitoring method (SIM) in single quad (Agilent 6120). Two buffers were used, Buffer A containing 10 mM ammonium formate with pH adjusted to 4.0 and Buffer B containing 100% acetonitrile, isocratic flow with 25% buffer B was used with flow rate of 0.5 ml/min using Agilent Poroshell 120 EC-C18 2.1 × 50 mm 2.7 mm column. 10 µl of sample was injected and SIM method was used to detect mass of 167 m/z in negative mode in order to estimate PEP concentration in cell lysate.

**Metabolic network model of E. coli BL21.** The iHK1487 metabolic model of E. coli31, containing 1487 genes, 2701 reactions, and 1164 metabolites, was used as the scaffold to integrate F420 biosynthesis pathway in this study. 360 reactions were found to be mass or charge imbalanced in the model, of which 338 reactions were demand or exchange reactions and one was biomass reaction, and therefore, were excluded from further corrections. Of the remaining 21 imbalanced reactions, 12 were involved in the lipopolysaccharide and cell envelope biosynthesis/recycling pathways, four were in alternate carbon metabolism pathway, three were in capsular polysaccharide biosynthesis/recycling pathway, and two were transporters (Supplementary File 1; Table S14). 14 imbalanced reactions were corrected by balancing protons, correcting chemical formula, or modifying participating metabolite(s) (see Supplementary File 1 for details). For example, four new metabolites were amended to the metabolic model to represent balanced core oligosaccharide lipid A molecules. The remaining seven imbalanced reactions could not be further resolved due to the lack of metabolite and/or enzyme specificity, rendered non-essential after flux analysis.

The bacterial F420 biosynthesis pathway proposed by Bashiri et al.18 was added to the original metabolic model to enable prediction of F420 production by engineered E. coli. This updated model was named iEco-F420. For synthesis in the iEco-F420 model was designed by adding two molecules of S-adenosyl-l-methionine to the reactant side, and adding two molecules of l-methionine, one molecule of ammonia, and two molecules of 5’-deoxyadenosine to the product side of the catalysing reaction. Moreover, the updated model allows for analysing the biosynthesis of cofactor F420 with up to eight glutamate residues (through l-glutamate:coenzyme F420 ligase, CofE).

For that, either an F420 molecule with one glutamate residue or a stoichiometric combination of F420 molecules with varying glutamate tails can be set as the target to analyse their flux and biosynthesis profile. Furthermore, the iEco-F420 model contains F420-dependent formate dehydrogenase62, F420-dependent G6P dehydrogenase63, F420:NADP oxidoreductase64, F420-dependent oxidoreductase65, and F420-reducing hydrogenase65 allowing for the analysis of cofactor F420 recycling and regeneration within the metabolic network of engineered F420-producing E. coli.

Integrating the F420 biosynthesis pathway along with correcting imbalances resulted in the iEco-F420 metabolic model of E. coli with 26 new metabolites and 43 new reactions. The updated metabolic model is available in Excel format in Tables S15 and S16 of Supplementary File 1. All reaction fluxes are in mmol/g DCW except for the reaction representing cell biomass formation that is expressed in h−1. The M9 minimal medium composition was designed to constrain the input of nutrients in the updated model. Independent simulations were run using glucose, glycerol, pyruvate, acetate, fumarate, or succinate as sole carbon sources (60 C-mol of carbon source). The uptake of C-source was fixed to 60 C-mol in all simulations to account for differences in number of carbon atoms in C-sources. The objective function was to maximize F420 production, while maintaining growth at 30% of its maximum to represent in vivo growth conditions. Maintenance ATP requirements were fixed at 5.17 mmol/g DCW and the minimum oxygen uptake was set to 18.5 mmol/g DCW/h. For simulating in vivo
over-expression of PPS and PPCK experiments, flux through PPS or PPCK was fixed in each run by constraining their lower and upper bounds to a value between zero and 70 mmol/g DCW/h, whereas zero represents no over-expression. Only one carbohydrate transporter was allowed to be active in each of these simulation runs. The model was assembled in a format compatible for flux balance analysis. IBA optimization problems were solved by GNU Linear Programming Kit (GLPK) (http://www.gnu.org/software/glpk/) solver in MATLAB using COBRA toolbox. Flux variability analysis (FVA) was performed to obtain range of fluxes under optimal growth conditions as described previously.

Data availability

The datasets generated during and/or analysed during the current study are available either as supplementary files or from the corresponding author on reasonable request.

Received: 30 May 2021; Accepted: 21 October 2021
Published online: 05 November 2021

References

1. Daniels, L., Fuchs, G., Thauer, R. K. & Zeikus, J. G. Carbon monoxide oxidation by methanogenic bacteria. J. Bacteriol. 132, 118–126 (1977).

2. Eirich, L. D., Vogels, G. D. & Wolfe, R. S. Proposed structure for coenzyme F420 from Methanococcus jannaschii. Biochemistry 17, 4583–4593. https://doi.org/10.1021/bi00154a002 (1978).

3. Eirich, L. D., Vogels, G. D. & Wolfe, R. S. Distribution of coenzyme F420 and properties of its hydrolytic fragments. J. Bacteriol. 140, 20–27 (1979).

4. Ferry, J. G. Methanogenesis: Ecology, physiology, biochemistry and genetics. Springer, Berlin, (2012).

5. Ferry, J. G. Biochemistry of methanogenesis. Crit. Rev. Biochem. Mol. 27, 473–503. https://doi.org/10.3109/104092920982570 (1992).

6. Knittel, K. & Boetius, A. Anaerobic oxidation of methane: Progress with an unknown process. Annu. Rev. Microbiol. 63, 311–334. https://doi.org/10.1146/annurev.micro.61.080706.093130 (2009).

7. Nazem-Bokaee, H., Yan, Z., Maranas, C. D. & Ferry, J. G. in Methane Bicatalysis: Paving the Way to Sustainability 183–197 (Springer, Berlin, 2018).

8. Wang, E. P. et al. Methanothrophic archaea possessing diverging methane-oxidizing and electron-transporting pathways. ISME J. 8, 1069–1078. https://doi.org/10.1038/ismej.2013.212 (2014).

9. Thauer, R. K. Anaerobic oxidation of methane with sulfate: On the reversibility of the reactions that are catalyzed by enzymes also involved in methanogenesis from CO2. Curr. Opin. Microbiol. 14, 292–299. https://doi.org/10.1016/j.mib.2011.03.003 (2011).

10. Reuke, B., Korn, S., Eisenreich, W. & Bacher, A. Biosynthetic precursors of deazaflavins. J. Bacteriol. 174, 4042–4049. https://doi.org/10.1128/JB.174.12.4042-4049.1992 (1992).

11. Neussner, D. et al. The genes Imb1F and Imb2 of Streptomyces lincolnensis encode enzymes involved in the conversion of L-tyrosine to propylpropionate during the biosynthesis of the antibiotic lincomycin A. Arch. Microbiol. 169, 322–332. https://doi.org/10.1007/s002030050578 (1998).

12. Jacobson, F. & Walsh, C. Properties of 7,8-dimethyl-8-hydroxy-5-deazaflavins relevant to redox coenzyme function in methanogenic metabolism. Biochemistry 23, 979–988. https://doi.org/10.1021/bi00300at28 (1984).

13. Walsh, C. Naturally-occurring 5-deazaflavin coenzymes—Biological redox roles. Acc. Chem. Res. 19, 216–221. https://doi.org/10.1021/ar00127a004 (1986).

14. Taylor, M. C. et al. Identification and characterization of two families of F420-dependent reductases from Methylocystis that catalyze aflatoxin degradation. Mol. Microbiol. 78, 561–575. https://doi.org/10.1111/jsm.1365-2958.2010.07356.x (2010).

15. Shah, M. V. et al. Cofactor F420-dependent enzymes: An under-explored resource for asymmetric redox biocatalysis. Catalysts 9, 868 (2019).

16. Yan, Z., Wang, M. & Ferry, J. G. A Ferredoxin- and F420H2-dependent, electron-bifurcating, heterodisulfide reductase with homologs in the domains bacteria and archaea. BioBio. https://doi.org/10.1128/mBio.02285-16 (2017).

17. Nazem-Bokaee, H. & Maranas, C. D. A prospective study on the fermentation landscape of gaseous substrates to biorenewables using Methanococcus jannaschii metabolic model. Front. Microbiol. 9, 1855. https://doi.org/10.3389/fmicb.2018.01855 (2018).

18. Bashiri, G. et al. A revised biosynthetic pathway for the cofactor F420 in prokaryotes. Nat. Commun. 10, 1558. https://doi.org/10.1038/s41467-019-09534-x (2019).

19. DePats, W. H., Bergkessel, M. & Newman, D. K. Aggregation of nontuberculous mycobacteria is regulated by carbon-nitrogen metabolic model. J. Bacteriol. 10, e01715–01719. https://doi.org/10.1128/mBio.01715-19 (2019).

20. Kimachi, T., Kawase, M., Matsuki, S., Tanaka, K. & Yoneda, F. First total synthesis of coenzyme factor 420. J. Chem. Soc. Perkin Trans. 1, 253–256. https://doi.org/10.1039/P19900000253 (1990).

21. Kimachi, T., Tanaka, K. & Yoneda, F. Synthesis of a proposed isomer of F420 having a-glutaminyl bonding. J. Heterocycl. Chem. 28, 439–443. https://doi.org/10.1002/jhet.5570280244 (1991).

22. Hossain, M. S. et al. Convenient synthesis of deazaflavin cofactor F2 and its activity in F2-dependent NADP reductase. Org. Biomol. Chem. 13, 5082–5085. https://doi.org/10.1039/C5OB00365B (2015).

23. Drenth, J., Trájkovic, M. & Fraaije, M. W. Chemoenzymatic synthesis of an unnatural deazaflavin cofactor that can fuel F2-dependent enzymes. ACS Catal. 9, 6435–6443. https://doi.org/10.1021/acscatal.9b01506 (2019).

24. Decamps, L. et al. Biosynthesis of F2 precursor of the F420 cofactor, requires a unique two radical-SAM domain enzyme and tyrosine as substrate. J. Am. Chem. Soc. 134, 18173–18176. https://doi.org/10.1021/ja307762b (2012).

25. Isabelle, D., Simpson, D. R. & Daniels, L. Large-scale production of coenzyme F2-5-b by using Mycobacterium smegmatis. Appl. Environ. Microbiol. 68, 5750–5755. https://doi.org/10.1128/AEM.68.11.5750-5755.2002 (2002).

26. Braga, D. et al. Metabolic pathway rerouting in Paraburkholderia rhizoxinica evolved long-overlooked derivatives of coenzyme F420. ACS Chem. Biol. https://doi.org/10.1021/acschembio.9b00605 (2019).

27. Braga, D. et al. Metabolic pathway rerouting in Paraburkholderia rhizoxinica evolved long-overlooked derivatives of coenzyme F420. ACS Chem. Biol. 14, 2088–2094. https://doi.org/10.1021/acschembio.9b00605 (2019).

28. Bashiri, G. & Baker, E. N. Convergent pathways to biosynthesis of the versatile cofactor F420. Curr. Opin. Struct. Biol. 65, 9–16. https://doi.org/10.1016/j.sbi.2020.05.002 (2020).

29. Grochowski, L. L., Xu, H. & White, R. H. Identification and characterization of the 2-phospho-1-lactate guanylyltransferase involved in coenzyme F420 biosynthesis. Biochemistry 47, 3033–3037. https://doi.org/10.1021/bi070247t (2008).

30. Braga, D. et al. Metabolic pathway rerouting in Paraburkholderia rhizoxinica evolved long-overlooked derivatives of coenzyme F420. bioRxiv. https://doi.org/10.1101/670455 (2019).

31. Kim, H., Kim, S. & Yoon, S. H. Metabolic network reconstruction and phenotype analysis of the industrial microbe, Escherichia coli BL21 (DE3). PLOS ONE. https://doi.org/10.1371/journal.pone.0204373 (2018).
35. Purwantini, E., Loganathan, U. & Mukhopadhyay, B. Coenzyme F420-dependent glucose-6-phosphate dehydrogenase coupled with l-phenylalanine dehydrogenase in M. jostii RHA1. J. Biol. Chem. 275, 5139–5148. https://doi.org/10.1074/jbc.9200142 (1992).

36. Noeck, B. et al. Structure of an amide bond forming F420-glutamyl ligase from Archaeoglobus fulgidus—A member of a new family of non-ribosomal peptide synthases. J. Mol. Biol. 372, 456–469. https://doi.org/10.1016/j.jmb.2007.06.063 (2007).

37. Liao, J. C. A global regulatory role of gluconeogenic genes in Escherichia coli revealed by transcriptome network analysis. J. Biol. Chem. 280, 36079–36087. https://doi.org/10.1074/jbc.M110.168430 (2005).

38. Besson, V., Rebeille, F., Neuburger, M., Douce, R. & Cosin, E. A. Effects of tetrahydrololate polylactamates on the kinetic parameters of serine hydroxymethyltransferase and glycine decarboxylase from pea leaf mitochondria. Biochem. J. 292, 425–430. https://doi.org/10.1042/bj2920425 (1993).

39. Peck, M. W. Changes in concentrations of coenzyme F420 analogs during batch growth of Methanosarcina barkeri and Methanothermobacter saccharolyticus. Appl. Environ. Microbiol. 65, 940–945 (1999).

40. Hoque, M. A. et al. Comparison of dynamic responses of cellular metabolites in Escherichia coli to pulse addition of substrates. Biologia 66, 954. https://doi.org/10.2478/bi-2015-0036 (2015).

41. Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in Escherichia coli. Biophys. Res. Commun. 59, 1204–1210. https://doi.org/10.1006/birc.1992.1200 (1992).

42. Schauer, N. L., Ferry, J. G., Honek, J. F., Orme-Johnson, W. H. & Walsh, C. Mechanistic studies of the coenzyme F420 reducing formate dehydrogenase from Methanobacterium formicicum. Biochemistry 25, 7163–7168. https://doi.org/10.1021/bi0200459 (1986).

43. Nguyen, Q.-T., Trincio, G., Bindia, C., Mattei, A. & Fraanje, M. W. Discovery and characterization of an F420-dependent glucose-6-phosphate dehydrogenase (Rh-FGD1) from Rhodococcus jostii RHA1. Appl. Microbiol. Biotechnol. 101, 2831–2842. https://doi.org/10.1007/s00253-016-8038-y (2017).

44. Kao, K. C., Tran, L. M. & Liao, J. C. A global regulatory role of gluconeogenic genes in Escherichia coli revealed by transcriptome network analysis. J. Biol. Chem. 280, 36079–36087. https://doi.org/10.1074/jbc.M110.168430 (2005).

45. Nolte, V. et al. Structure of an amide bond forming F420-glutamyl ligase from Archaeoglobus fulgidus—A member of a new family of non-ribosomal peptide synthases. J. Mol. Biol. 372, 456–469. https://doi.org/10.1016/j.jmb.2007.06.063 (2007).

46. Peck, M. W. Changes in concentrations of coenzyme F420 analogs during batch growth of Methanosarcina barkeri and Methanothermobacter saccharolyticus. Appl. Environ. Microbiol. 65, 940–945 (1999).

47. Hoque, M. A. et al. Comparison of dynamic responses of cellular metabolites in Escherichia coli to pulse addition of substrates. Biologia 66, 954. https://doi.org/10.2478/bi-2015-0036 (2015).

48. Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in Escherichia coli. Biophys. Res. Commun. 59, 1204–1210. https://doi.org/10.1006/birc.1992.1200 (1992).

49. Schauer, N. L., Ferry, J. G., Honek, J. F., Orme-Johnson, W. H. & Walsh, C. Mechanistic studies of the coenzyme F420 reducing formate dehydrogenase from Methanobacterium formicicum. Biochemistry 25, 7163–7168. https://doi.org/10.1021/bi0200459 (1986).

50. Nguyen, Q.-T., Trincio, G., Bindia, C., Mattei, A. & Fraanje, M. W. Discovery and characterization of an F420-dependent glucose-6-phosphate dehydrogenase (Rh-FGD1) from Rhodococcus jostii RHA1. Appl. Microbiol. Biotechnol. 101, 2831–2842. https://doi.org/10.1007/s00253-016-8038-y (2017).

51. Kamar, H., Nguyen, Q.-T., Bindia, C., Mattei, A. & Fraanje, M. W. Isolation and characterization of a thromostable F420-NADPH dehydrogenase from Thermobifida fusca. J. Biol. Chem. 292, 10123–10130. https://doi.org/10.1074/jbc.M117.778754 (2017).

52. Hendrickson, E. L. & Leigh, J. A. Roles of coenzyme F420-reducing hydrogenases and hydrogen- and methylene-tetrahydromethanopterin dehydrogenases in reduction of F420 and production of hydrogen during methanogenesis. J. Bacteriol. 190, 4818–4821. https://doi.org/10.1128/jbc.2008.00255 (2008).

53. Orth, J. D., Thiele, I. & Palsson, B. O. What is flux balance analysis?. Nat. Biotechnol. 28, 245–248. https://doi.org/10.1038/nbt.1614 (2010).

54. Heiderndt, L. et al. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v3.0. Nature Protocols 14, 639–702. https://doi.org/10.1038/nprot.2018.008 (2019).

55. Mahadevan, R. & Schilling, C. H. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. Metab. Eng. 5, 265–276. https://doi.org/10.1016/j.ymben.2003.09.002 (2003).
Acknowledgements
We would like to thank the CSIRO Synthetic Biology Future Science Platform for funding this work and providing Fellowships for MVS and HNB, and PhD Top-up funding to SWK and JA. We would also like to thank Adjunct Professor Claudia Vickers and Dr. Andrew Warden for their helpful comments during the preparation of this manuscript.

Author contributions
M.V.S., conceptualization, methodology, formal analysis, investigation, writing - original draft; H.N.-B., conceptualization, methodology, formal analysis, investigation, simulations, writing - original draft, figures; J.A., investigation; S.W.K., investigation; C.I.J., conceptualization, formal analysis, supervision, writing - review & editing, funding acquisition; C.S., conceptualization, formal analysis, supervision, writing - review & editing, project administration, funding acquisition.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-01224-3.

Correspondence and requests for materials should be addressed to C.S.

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

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

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/.

© The Author(s) 2021