A revised biosynthetic pathway for the cofactor F\textsubscript{420} in prokaryotes

Ghader Bashiri\textsuperscript{1}, James Antoney\textsuperscript{2,3}, Ehab N.M. Jirgis\textsuperscript{1}, Mihir V. Shah\textsuperscript{2}, Blair Ney\textsuperscript{2,3}, Janine Copp\textsuperscript{4}, Stephanie M. Stuteley\textsuperscript{1}, Sreevalsan Sreebhavan\textsuperscript{5}, Brian Palmer\textsuperscript{5}, Martin Middleditch\textsuperscript{1}, Nobuhiko Tokuriki\textsuperscript{4}, Chris Greening\textsuperscript{2,6}, Colin Scott\textsuperscript{2}, Edward N. Baker\textsuperscript{1} & Colin J. Jackson\textsuperscript{2,3}

Cofactor F\textsubscript{420} plays critical roles in primary and secondary metabolism in a range of bacteria and archaea as a low-potential hydride transfer agent. It mediates a variety of important redox transformations involved in bacterial persistence, antibiotic biosynthesis, pro-drug activation and methanogenesis. However, the biosynthetic pathway for F\textsubscript{420} has not been fully elucidated: neither the enzyme that generates the putative intermediate 2-phospho-L-lactate, nor the function of the FMN-binding C-terminal domain of the γ-glutamyl ligase (FbiB) in bacteria are known. Here we present the structure of the guanylyltransferase FbiD and show that, along with its archaeal homolog CofC, it accepts phosphoenolpyruvate, rather than 2-phospho-L-lactate, as the substrate, leading to the formation of the previously uncharacterized intermediate dehydro-F\textsubscript{420}-0. The C-terminal domain of FbiB then utilizes FMNH\textsubscript{2} to reduce dehydro-F\textsubscript{420}-0, which produces mature F\textsubscript{420} species when combined with the γ-glutamyl ligase activity of the N-terminal domain. These new insights have allowed the heterologous production of F\textsubscript{420} from a recombinant F\textsubscript{420} biosynthetic pathway in Escherichia coli.
Cofactor F420 is a deazalavin that acts as a hydride carrier in diverse redox reactions in both bacteria and archaea. While F420 structurally resembles the flavins flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD), it functions as an obligate two-electron hydride carrier and hence is functionally analogous to the nicotinamides NAD+ and NADP+. The lower reduction potential of the F420, relative to the flavins, results from the substitution of N5 of the isoalloxazine ring in the flavins for a carbon in F420. Originally characterized from methanogenic archaea in 1972, F420 is an important catabolic cofactor in methanogens and mediates key one-carbon transformations of methanogenesis. F420 has since been shown to be synthesized in a range of archaea and bacteria. In Mycobacterium tuberculosis, the causative agent of tuberculosis, F420 has been shown to contribute to persistence and to activate the new clinical antituberculitic prodrugs delamanid and pretomanid. There are also growing numbers of natural products that have been shown to be synthesized through F420-dependent pathways, including tetracyclines, lincosamides, and thiostrepton. F420-dependent enzymes have also been explored for bioremediation and biocatalytic applications.

The currently accepted F420 biosynthetic pathway consists of two branches (Fig. 1). In the first branch, tyrosine is condensed with 5-amino-6-ribitylaminol[1,3]-pyrimidine dione from the flavin biosynthetic pathway to generate the deazachlorophores. In the second branch, tyrosine is converted to F420, and the CofG/H pair (where “Cof” refers to mycobacterial proteins and “Fo” refers to archaeal homologs). In the second branch, it has been hypothesized that a 2-phospho-l-lactate guanylyltransferase (CofC in archaea and the putative enzyme FbiD in bacteria) catalyzes the guanylylation of 2-phospho-l-lactate (2-PL) using guanosine-5’-triphosphate (GTP), yielding l-lactyl-2-diphospho-5’-guanosine (LPPG). The two branches then merge at the reaction catalyzed by the transferase FbiA/CofD, where the 2-phospho-l-lactyl moiety of LPPG is transferred to Fo, forming F420-0. Finally, the γ-glutamyl ligase (FbiB/CofE) catalyzes the polyglutamylation of F420 to generate mature F420, with poly-γ-glutamate tail lengths of ~2–8, depending on species.

There are three aspects of the F420 biosynthetic pathway that require clarification. First, the metabolic origin of 2-PL, the proposed substrate for CofC, is unclear. It has been assumed that a hypothetical kinase (designated CofB) phosphorylates l-lactate to produce 2-PL. However, no such enzyme has been identified in bacteria or archaea, and our genomic analysis of F420 biosynthesis operons has failed to identify any candidate enzymes with putative l-lactate kinase activity. Second, the existence of FbiD has only been inferred through bioinformatics and genetic knockout studies and the enzyme has not been formally characterized. Finally, the bacterial γ-glutamyl ligase FbiB is a two-domain protein, in which the N-terminal domain is homologous to other F420-γ-glutamyl ligases (including the archaenal equivalent, CoFe) and the C-terminal domain adopts an FMN-binding nitroreductase (NTR) fold. Although both domains are required for full γ-glutamyl ligase activity, no function has been associated with either the C-terminal domain or the FMN cofactor, given no redox reactions are known to be involved in F420 biosynthesis.

Here we demonstrate that 2-PL is not required for F420 biosynthesis in prokaryotes and instead phosphoenolpyruvate (PEP), an abundant intermediate of glycolysis and gluconeogenesis, is incorporated into F420. Mass spectrometry (MS) and protein crystallography are used to demonstrate that PEP guanylylation is catalyzed by its CofD/CofE enzymes that were previously thought to act upon 2-PL. In bacteria, the incorporation of PEP in the pathway results in the production of the previously undetected intermediate dehydro-F420-0, which we identified by MS. We then showed, with the help of ligand docking, that this intermediate is then reduced by the C-terminal domain and polyglutamylated by the N-terminal domain. These findings result in a substantially revised pathway for F420 biosynthesis and have allowed us to heterologously express a functional F420 biosynthetic pathway in Escherichia coli, an organism that does not normally produce F420 at levels comparable to some native F420-producing organisms.

**Results**

FbiD/CofC accept PEP, rather than 2-PL. The archaenal enzyme CofC has previously been suggested to catalyze the guanylylation of 2-PL to produce LPPG during F420 biosynthesis (Fig. 2a). Another study, using transposon mutagenesis, has shown that MSMEG 2392 of Mycobacterium smegmatis is essential in the biosynthesis of F420 from Fo. We have recently shown that homologs of this gene have sequence homology to CofC and belong to operons with other validated F420 biosynthetic genes in a wide variety of bacteria. In keeping with the bacterial naming system, we refer to this enzyme as FbiD. To test the function of this putative bacterial FbiD, we cloned the homologous Rv2983 gene from M. tuberculosis into a mycobacterial expression vector and purified heterologously expressed Mycobacterium tuberculosis FbiD (MtFbiD) from M. smegmatis mc24517 host cells. We also expressed and purified Mt-FbiA (the enzyme thought to transfer the 2-phospho-l-lactyl moiety of LPPG to Fo to produce F420-0) to use in coupled high-performance liquid chromatography-MS (HPLC-MS) enzymatic assays with Mt-FbiD. Surprisingly, we found that when Mt-FbiD and Mt-FbiA were included in an assay with 2-PL, GTP (or ATP), and Fo, no product was formed (Fig. 2b). We then tested whether Mt-FbiA and CofC from Methanocaldococcus jannaschii (Mj-CofC) could catalyze F420-0 formation under the same conditions, which again yielded no product (Fig. 2b).

Although 2-PL is hypothesized to be an intermediate in F420 biosynthesis, this has never been experimentally confirmed in bacteria. Additionally, no enzyme capable of phosphorylating l-lactate to 2-PL has been identified in F420-producing organisms, despite considerable investigation. 2-PL has been little studied as a metabolite and is only known to occur as a by-product of pyruvate kinase activity. 2-PL has not been implicated as a substrate in any metabolic pathway outside the proposed role in F420 biosynthesis; rather, it has been shown in vitro to inhibit several enzymes involved in glycolysis and amino acid biosynthesis. Our inability to detect activity with 2-PL led us to consider alternative metabolites that could potentially substitute for 2-PL, including the structurally analogous and comparatively abundant molecule PEP.

While there was no activity when 2-PL was used in the FbiD/CofC:FbiA-coupled assays, when these enzymes were incubated with PEP, GTP (or ATP), and Fo, a previously unreported intermediate in the F420 biosynthesis pathway, which we term “dehydro-F420-0,” was produced (Fig. 2b). The identity of this compound, which is identical to F420-0 except for a methylene group in place of the terminal methyl group, was verified by MS/MS (Fig. 2c). The only difference that we observed between the activities of Mt-FbiD and Mj-CoFe was that while Mt-FbiD exclusively utilizes GTP to produce dehydro-F420-0, Mj-CoFe can also catalyze the reaction with ATP, albeit to a lesser extent (Fig. 2b). Interestingly, in our experiments the FbiD/CofC enzymes were only active in the presence of FbiA. This was not unexpected given that the inferred intermediate (enolpyruvyl-diphospho-5’-guanosine (EPPG)) is expected to be unstable.
To understand the molecular basis of PEP recognition by *Mtb*-FbiD, we crystallized the protein and solved the structure by selenium single-wavelength anomalous diffraction (Se-SAD), and then used this selenomethionine-substituted structure to obtain the native FbiD structure by molecular replacement. The latter was then refined at 1.99 Å resolution (*R*/Rfree = 0.19/0.22) (Table 1). As expected, *Mtb*-FbiD adopts the same MobA-like nucleoside triphosphate transferase family protein fold as CoFC.
central 7-stranded β-sheet (six parallel strands and one antiparallel), with α-helices packed on either side (Fig. 3a). However, *Mtb*-FbiD lacks the protruding hairpin that is important for dimer formation in CofC. Superposition of CofC from *Methanosarcina mazei* (PDB code 2I5E) onto *Mtb*-FbiD gives a root mean square difference of 1.85 Å over 181 Ca atoms, with 25.4% sequence identity, establishing clear structural homology (Fig. 3c).

We also soaked PEP into pre-formed FbiD crystals to obtain an FbiD-PEP complex (2.18 Å resolution, R/Rfree = 0.22/0.26). FbiD has a cone-shaped binding cleft with a groove running across the base of the cone, formed by the C-terminal end of the central β-sheet (Fig. 3a). PEP binds in the cleft with its phosphate group anchored through two Mg2+ ions to three acidic side chains (D116, D188, and D190) (Fig. 3b). This three-aspartate motif is conserved amongst FbiD homologs as shown by a multiple sequence alignment (Supplementary Fig. 1). The PEP carboxylate group is hydrogen bonded to the hydroxyl group of S166 and the main chain NH groups of T148 and G163. All PEP binding residues are conserved in the CofC protein of *M. mazei* (PDB code 2I5E) (Supplementary Fig. 1), consistent with the enzymatic assays that showed PEP is the substrate for archaeal CofC, as well as FbiD. The reason that FbiD/CofC is active with PEP rather than 2-PL is most likely a consequence of the different stereochemistry of 2-PL (compared with the planar geometry of PEP) preventing a necessary structural rearrangement, or the binding of GTP, which would be required to attain a productive transition state, as suggested for PEP carboxykinase. In the GTP-bound structure of *E. coli* MobA (PDB code 1FRW), GTP binds in a characteristic surface groove, providing a structural framework for substrate binding and catalysis. In our enzyme assays, neither FbiD nor CofC showed activity in the absence of FbiA. Moreover, we did not observe GTP binding in either our co-crystallization or differential scanning fluorimetry experiments. We speculate that the GTP binding site is not formed until FbiD/CofC interacts with FbiA/CofD, enabling catalysis to proceed through to formation of dehydro-F420-0. This may provide an advantage by producing EPPG/EPPA only when both proteins are available, thereby overcoming the issue of intermediate instability.

The C-terminal domain of FbiB reduces dehydro-F420-0 to F420-0. Dehydro-F420-0 would yield F420-0 upon reduction of the terminal methylene double bond. However, no masses corresponding to F420-0 were identified in any of the liquid chromatography-MS (LC-MS) traces from the FbiD:FbiA-coupled assays, suggesting that an enzyme other than FbiD or FbiA catalyzes dehydro-F420-0 reduction. We have previously shown that full-length FbiB consists of two domains: an N-terminal...
domain that is homologous to the archaeal γ-glutamyl ligase CoE\textsuperscript{a,b,35}, and a C-terminal domain of the NTR fold\textsuperscript{36} that binds to FMN and has no known function, but is essential for extending the poly-γ-glutamate tail\textsuperscript{36}.

We tested whether FbiB could use dehydro-F\textsubscript{420}-0 as a substrate with a three enzyme assay in which FbiB and L-glutamate were added to the FbiD:FbiA-coupled assay. 

\textit{Mtb}-FbiB was observed to catalyze the addition of L-glutamate residues to dehydro-F\textsubscript{420}-0, forming dehydro-F\textsubscript{420} species with one ([M + H\textsuperscript{+}], m/z of 643.40) and two ([M + H\textsuperscript{+}], m/z of 772.40) glutamate residues (Supplementary Fig. 2). We then tested the hypothesis that the orphan function of the FMN-binding C-terminal domain could in fact be a dehydro-F\textsubscript{420}-0 reductase. We used a four-enzyme in vitro assay where \textit{E. coli} NAD(P)H:flavin oxidoreductase\textsuperscript{37} (Fre), FMN, NADH, and 10 mM dithiothreitol (to maintain reducing conditions and generate reduced FMNH\textsubscript{2}) were added to the FbiD:FbiA:FbiB assay and the reaction was performed in anaerobic conditions (to prevent reaction of FMNH\textsubscript{2} with oxygen). We found that F\textsubscript{420}-1, that is, the fully reduced and mature glutamylated cofactor, was produced, but only in the presence of both FbiB and Fre/FMNH\textsubscript{2} (Fig. 4a).

Thus, dehydro-F\textsubscript{420}-0 is a bona fide metabolic intermediate that can be converted to mature F\textsubscript{420} by FbiB in an FMNH\textsubscript{2}-dependent fashion. These results demonstrate that bacterial FbiB is a bifunctional enzyme, functioning as a dehydro-F\textsubscript{420}-0 reductase and as a γ-glutamyl ligase (Fig. 4b).

When the previously published crystal structures of \textit{Mtb}-FbiB are analyzed in the context of these results, the molecular basis for this activity becomes clear. Crystal structures of the \textit{Mtb}-FbiB C-terminal domain with F\textsubscript{420} bound (PDB ID: 4XOQ) and FMN bound (PDB ID: 4XOO) have been solved\textsuperscript{38}, and when these are overlaid it is apparent that the FMN molecule is ideally situated to transfer a hydride to the terminal methylene of dehydro-F\textsubscript{420}-0 (assuming dehydro-F\textsubscript{420}-0 binds in a similar fashion to F\textsubscript{420}). Interestingly, the phospholactyl group of F\textsubscript{420} appears to be disordered in these crystal structures, suggesting it may adopt multiple conformations. To test this, we docked dehydro-F\textsubscript{420}-0 into the FMNH\textsubscript{2}-bound structure and performed an energy minimization using the OPLS\textsuperscript{3e} forcefield to allow them to attain a low-energy conformation. This forcefield was chosen based on its improved torsional parameters for small molecules and more accurate descriptions of the potential energy surfaces of ligands, allowing for improved docking poses over other forcefields\textsuperscript{38}. The results show that in this ternary complex the two molecules can adopt ideal positions and orientations for the reduction of dehydro-F\textsubscript{420}-0 (Fig. 4c). The methylene group of dehydro-F\textsubscript{420}-0 is accommodated by a small hydrophobic pocket mostly comprised of P289 and M372 allowing it to be positioned above the N5 atom of FMNH\textsubscript{2}, in a plausible Michaelis complex for hydride transfer (Fig. 4c). We therefore suggest that the phosphonolpyruvyl (analogous to the phospholactyl) group of dehydro-F\textsubscript{420}-0 most likely samples conformations within this pocket where it can be reduced.

Interestingly, the γ-glutamyl ligase CoE from archaea is a single domain enzyme; there is no homology to the C-terminal NTR-fold domain of FbiB. In an analogous situation, Fo synthesis is performed by two single domain enzymes in archaea, CoH, and CoG\textsuperscript{2}, whereas in bacteria this reaction is catalyzed by a two-domain protein, FbiC (with N- and C-terminal domains homologous to CoH and CoG, respectively). Previous analysis

---

**Table 1 Data collection, processing, and refinement statistics**

|                | SeMet-FbiD\textsuperscript{a} | Apo-FbiD\textsuperscript{a} (PDB 68WG) | FbiD-PEP\textsuperscript{a} (PDB 68WH) |
|----------------|-------------------------------|----------------------------------------|---------------------------------------|
| Data collection| Space group                   | C222\textsubscript{1}                  | C222\textsubscript{1}                |
|                | Cell dimensions (Å)           | a, b, c                                  | a, b, c                                  |
|                | Resolution (Å)               | 47.17–2.33                              | (2.14–2.33)                           |
|                | i/φ                          | 21.3 (8.1)                              | 8.8 (1.2)                              |
|                | Completeness (%)             | 99.99 (99.9)                            | 99.99 (99.9)                           |
|                | Redundancy (%)               | 14.6 (14.2)                             | 14.8 (13.5)                            |
| Refinement     | Resolution (Å)               | 83.25–1.99                              | 82.95–1.98                             |
|                | Rmerge/RFom                  | 19.1/22.8                              | 22.8/26.2                              |
|                | No. reflections              | 621728 (38169)                          | 477504 (35875)                         |
|                | No. atoms                    | 4769                                    | 4643                                   |
|                | Protein                      | 36                                      | 36                                     |
|                | Water                        | 180                                     | 4643                                   |
|                | B-factors                    | 24.12                                   | 48.78                                  |
|                | Mg\textsuperscript{2+} (n = 6) | 44.33                                   | 48.46                                  |
|                | Water                        | 48.46                                   | 48.46                                  |
|                | R.m.s. deviations            | 1.416                                   | 1.383                                  |
|                | Bond lengths (Å)             | 0.011                                   | 0.01                                   |
|                | Bond angles (°)              | 1.416                                   | 1.383                                  |

Values within parentheses are for highest-resolution shell

*Number of crystals = 1

---

**Fig. 3** Crystal structure of \textit{Mtb}-FbiD. a Electrostatic surface representation of the \textit{Mtb}-FbiD structure in complex with phosphoenolpyruvate (PEP), shown as a ball-and-stick model. b The phosphate group of PEP binds to three aspartic acid side chains through two Mg\textsuperscript{2+} ions (shown in cyan). PEP is shown in 2F\textsubscript{o} – F\textsubscript{c} omit density contoured at 2.0σ, and drawn as ball-and-stick model. Water molecules are shown as red spheres and hydrogen bond interactions are outlined as dashed lines. c Superposition of \textit{Mtb}-FbiD (wheat ribbon) and \textit{Methanothermobacter jannaschii} (Mj-CoFC) (green ribbon), indicating 1.85 Å root mean square difference (rmsd) over 181 superimposed Ca. PEP is shown as a ball-and-stick model.
The yield of purified F420 was ~27 nmol L\(^{-1}\) of culture, which is comparable to physiological levels of several F420-producing species. Ultraviolet-visible (UV–Vis) and fluorescence spectra of the purified F420 matched literature values (Supplementary Fig. 5). Finally, we confirmed that the purified cofactor was functional by measuring enzyme kinetic parameters with F420-dependent glucose-6-phosphate dehydrogenase activity.

Expression of \textit{fbiABCD} is sufficient to produce F420 in \textit{Escherichia coli}. Cofactor F420 is only produced by certain bacterial species; the majority of bacteria, including \textit{E. coli}, lack the genes required for F420 biosynthesis. Our in vitro assay results suggest that the first enzyme in the F420 biosynthetic pathway is a lactate kinase (CoFb) and the second enzyme is a NADPH-dependent glucose-6-phosphate dehydrogenase (CoFb/A). To test our hypothesis, we generated a recombinant plasmid expressing the entire \textit{fbiABCD} operon from \textit{M. mazei}. Western blot analysis of the cell lysate revealed the production of F420, as detected by anti-FLAG antibodies. Finally, we confirmed the functional activity of the purified F420 by measuring enzyme kinetic parameters with F420-dependent glucose-6-phosphate dehydrogenase activity.
dehydrogenase (FGD) from *M. smegmatis* (Fig. 5c). The apparent Michaelis constant was within error of that observed with FGD and F420 produced from *M. smegmatis*, while the $k_{\text{cat}}$ was approximately half that of the *M. smegmatis* F420 (Fig. 5c), which could result from slight differences in the distribution of tail lengths, as previously reported (Supplementary Table 2)\(^43\). These results confirm that the heterologous production of F420 was achieved with a biosynthetic pathway containing only CofD (a homolog of FbiA)/FbiB/FbiC/FbiD.

Having demonstrated that our recombinant *E. coli* strain could produce mature F420 that was functional ex vivo, we tested whether co-expression of F420-dependent enzymes could facilitate F420-dependent transformations in vivo. We have previously shown that F420H$_2$ can decolorize malachite green nonenzymatically, and that several F420-dependent reductases accelerate this reaction\(^44\). As summarized in Supplementary Table 3, *E. coli* were unable to decolorize malachite green without induction of F4$_{420}$ biosynthesis. When the cells additionally expressed FGD and the F$_{420}$-dependent oxidoreductase MSMEG\_2027, malachite green decolorization was significantly increased ($p < 0.001$, Student’s paired $t$-test, two-sided). These results demonstrate that incorporation of the F$_{420}$ biosynthetic pathway into *E. coli* allows F$_{420}$-dependent enzymes to function in vivo upon heterologous expression.

**Discussion**

It has become widely accepted within the field that one of the essential initial steps in F$_{420}$ biosynthesis involves a hypothetical L-lactate kinase that produces 2-PL, which is subsequently incorporated into F$_{420}$ through the activities of CofC and CofD. However, neither bioinformatics nor genetic knockout studies have identified plausible candidate genes for a L-lactate kinase\(^2,18,23\). Furthermore, 2-PL has been shown to inhibit several enzymes involved in central metabolism\(^28–30\). In terms of pathway flux, this makes 2-PL an unusual starting point for biosynthesis of an abundant metabolite such as F$_{420}$, which can exceed 1 $\mu$M in some species\(^42\). The results presented in this paper unequivocally demonstrate that PEP, rather than 2-PL, is the authentic starting metabolite in bacteria. These results reconcile the previously problematic assumptions that are required to include 2-PL within the biosynthetic pathway and establish a revised pathway (Fig. 6) that is directly linked to central carbon metabolism (via PEP) through the glycolysis and gluconeogenesis pathways.

Our observation that CofC accepts PEP (and not 2-PL) in vitro, appear to contradict previous studies in which Mj-CofC was reported to use 2-PL as substrate\(^19,35\). However, this discrepancy...
is most likely due to the supplementation of the coupled CofC/CofD reaction in these studies with pyruvate kinase and 2 mM PEP, a strategy that was used to overcome apparent product inhibition by GMP\textsuperscript{19,35}. Regardless, we cannot explain how a pathway starting from PEP can produce mature (i.e., not dehydro-F\textsubscript{420}) F\textsubscript{420} in archaea given that their equivalent of FbiB (CofE) lacks the C-terminal domain dehydro-F\textsubscript{420}-0 reductase domain seen in FbiB. One possibility is that an unknown dehydro-F\textsubscript{420}-0 reductase exists elsewhere in the genome (remote from CofE). Further studies are required to resolve this step in archaean F\textsubscript{420} biosynthesis.

The discovery of dehydro-F\textsubscript{420}-0 as the product of FbiD:FbiA activity in mycobacteria indicated that another enzyme must be required to reduce dehydro-F\textsubscript{420}-0 and produce F\textsubscript{420}-0. This allowed us to define the function of the orphan C-terminal domain of mycobacterial FbiB\textsuperscript{30}. The N-terminal domain is homologous with the archaean γ-glutamyl ligase CofE\textsuperscript{35}, but is only capable of catalyzing glutamylation of F\textsubscript{420}-1\textsuperscript{39}. The C-terminal domain, which binds both F\textsubscript{420} and FMN, is essential for extending the poly-γ-glutamate tail of F\textsubscript{420}-0\textsuperscript{20}, but we could not explain the possible function of the FMN cofactor. Here, we show that FbiB can catalyze both poly-glutamylation and reduction of dehydro-F\textsubscript{420}-0, via the activities of the N- and C-terminal domains, respectively.

The increasing recognition of the importance of F\textsubscript{420} in a variety of biotechnological, medical, and ecological contexts underlines the need for making the compound more widely accessible to researchers; however, our inability to produce F\textsubscript{420} recombinantly in common laboratory organisms has been a major barrier to wider study. Here, we confirm the results of our in vitro experiments by showing that recombinant expression of the four characterized F\textsubscript{420} biosynthesis genes allows production of F\textsubscript{420} in E. coli. These findings should now facilitate the use of F\textsubscript{420} in a variety of processes with recombinant organisms, such as biocatalysis using a bio-orthogonal cofactor, directed evolution of F\textsubscript{420}-dependent enzymes, recombinant production of antibiotics for which F\textsubscript{420} is a required cofactor, and metabolic engineering.

Methods

**Bacterial strains and growth conditions.** Protein expression utilized either *M. smegmatis* mc\textsuperscript{2}4517\textsuperscript{45}, E. coli BL21(DE3), or LOBST-R BL21(DE3)\textsuperscript{46} cells. For growth of *M. smegmatis*, media were supplemented with 0.05% (v/v) Tween-80. *Mycobacterium smegmatis* cells were grown in ZYM-5052\textsuperscript{47} or modified auto-induction media Terrific Broth (TB) 2.0 (2.0% tryptone, 0.5% yeast extract, 0.5% NaCl, 22 mM KH\textsubscript{2}PO\textsubscript{4}, 42 mM Na\textsubscript{2}HPO\textsubscript{4}, 0.6% glucose, 0.2% lactose)\textsuperscript{14}. For selenomethionine labelling, cells were grown in PASM-5052 media\textsuperscript{47}. *Escherichia coli* cell expressions were conducted in either the above-modified auto-induction media or TB medium modified for auto-induction of protein expression (1.2% tryptone, 2.4% yeast extract, 72 mM KH\textsubscript{2}PO\textsubscript{4}, 17 mM KH\textsubscript{2}PO\textsubscript{4}, 2 mM MgSO\textsubscript{4}, 0.8% glycerol, 0.05% glucose, 0.5% lactose, 0.375% aspartic acid), grown for 4 h at 37 °C followed by overnight incubation at 18 °C.

**Protein expression and purification.** Mtb-FbiD: The ORF encoding FbiD (Rv2983) was obtained by PCR from *M. tuberculosis* H37Rv genomic DNA (Supplementary Table 1). The pYUBDuet-fbiABD co-expression construct was then prepared by cloning fbiD into pYUBDuet\textsuperscript{45} using BamHI and HindIII restriction sites, followed by cloning the fbiAB operon using NdeI and PdiI restriction sites. This construct expresses FbiD with an N-terminal His\textsubscript{6}-tag, whereas the FbiA and FbiB proteins are expressed without any tags. The pYUBDuet-fbiABD vector was transformed into *M. smegmatis* mc\textsuperscript{2}4517 strain\textsuperscript{30} for expression. The cells were grown in a fermenter (BioFlo\textsuperscript{®} 415, New Brunswick Scientific) for 4 days before harvesting. The cells were lysed in 20 mM HEPES, pH 7.5, 150 mM NaCl, 20 mM imidazole, and 1 mM β-mercaptoethanol using a cell disruptor (Microfluidizer M-110P) in the presence of Complete protease inhibitor mixture mini EDTA-free tablets (Roche Applied Science). The lysate was centrifuged at 20,000 x g to separate the insoluble material. Recombinant FbiD was separated from other proteins by immobilized metal affinity chromatography on a HiTrap FF 5-mL Ni-NTA column (GE Healthcare), eluted with imidazole, and further purified by size-exclusion chromatography (SEC) on a Superdex 75 10/300 column (GE Healthcare) pre-equilibrated in 20 mM HEPES, pH 7.5, 150 mM NaCl, and 1 mM β-mercaptoethanol.

FbiA: The ORF encoding *M. tuberculosis* FbiA was commercially synthesized and cloned into pRSET-A (Invitrogen). The pYUB82b-fbiA construct used for expression in *M. smegmatis* mc\textsuperscript{2}4517 was prepared by

---

**Fig. 6** The revised bacterial F\textsubscript{420} biosynthesis pathway. The revised pathway is a modified scheme showing that phosphoenolpyruvate (PEP) acts as the substrate for the FbiD/CofC enzymes to produce enolpyruvyl-diphospho-5'-guanosine (EPPG) or enolpyruvyl-diphospho-5'-adenosine (EPPA) (in the case of CofC). The immediate reaction product formed from Fo and EPPG/EPPA is dehydro-F\textsubscript{420}-0, which is reduced to F\textsubscript{420}-0 through the newly described reductase activity of the C-terminal domain of FbiB in mycobacteria. A separate enzyme in archaea and some bacteria is expected to catalyze this reduction step (CofX). FbiB/CofE subsequently adds a poly-γ-glutamate tail to form F\textsubscript{420}. "Fbi" refers to bacterial proteins, whereas "Cof" represents archaean ones.
subcloning fBiA into pYUB28b41 using an Ndel site introduced by overlap PCR utilizing the HindIII restriction site present on both vectors amplified with HindIII (Supplementary Table 1). The resulting plasmid pYUB28b-fBiA construct expresses FBiA with a N-terminal His6-tag. The protein was expressed in M. smegmatis mc4517 in ZYM-3052 media auto-induction media47,48 in a fermentor (BioFlo® 415, New Brunswick Scientific) for 4 days. The protein was then purified using Ni-NTA and SEC steps, as described above, in 20 mM HEPES, pH 7.5, 200 mM NaCl, 5% glycerol, and 1 mM β-mercaptoethanol.

Mj-CoC: The ORF encoding Methanolococcus jannaschii CoC (Mj0887)23 was synthesized (GenScript) and cloned into pYUB28b vector41 using Ndel and HindIII restriction sites. The protein was expressed in E. coli in TB auto-induction media and purified using Ni-NTA and SEC steps, as described above, in 20 mM HEPES, pH 7.5, 200 mM NaCl 5% glycerol, and 1 mM β-mercaptoethanol.

Ec-Fre: The E. coli flavin reductase47 was cloned into pProEX-HTB using KaI and HindIII restriction sites (Supplementary Table 1). Protein expression and purification was performed similar to Mj-CoC.

**Construction of synthetic F420 biosynthesis operon.** Ribosome binding sites were individually optimized for each of the codon-optimized F420 biosynthesis genes using the Ribosome Binding Site Calculator server99,100. Multiple operon designs were analyzed using the server’s operon calculator and modified to remove unnecessary codon products and RNA and ribosome bloating while maintaining predicted translation initiation rates for all coding sequences within an order of magnitude. The final design placed the operon under the control of the tetracycline-inducible promoter Bba_R0040 and the artificial terminator Bba_B1006. For subsequent assembly the operon was flanked by BioBrick restriction sites and a PEP-BiD sequence. The operon was synthesized by GenScript and cloned into pSB1C3 containing the constitutive tetracycline repressor cassette Bba_K145201 using the standard BioBrick assembly protocol with EcoRI and Bbap/Sel restriction enzymes53. This construct, hereafter referred to as F420-op, enables production of F420 to be induced by the addition of anhydrotetracycline. For Western blot analysis as second version of the operon with single C-terminal FLAG tags on all four genes was likewise synthesized.

To test for in vivo F420-dependent reductase/oxidase activity, the ORF of M. smegmatis FGD (MSMEG_0777) was codon optimized and commercially synthesized by GenScript and cloned into multiple cloning site 1 of pCOLAduet-1 (Novagen) with Ncol and HindIII sites. Subsequently MSMEG_2392 was subcloned from pETMCSIII-MSMEG_2292 into multiple cloning site 2 using Ndel and Kpnl to give pFGD_2027.

**Mtb-fBiaB crystallization and structure determination.** Apo-fBiaB (20 mg mL$^{-1}$) in 20 mM HEPES, pH 7.5, 150 mM NaCl, 1 mM β-mercaptoethanol was crystallized using the sitting drop vapor diffusion method in 30% polyethylene glycol 1500, 3% MPD (2-methyl-2,4-pentanediol), 0.2 M MgSO4, and 0.1 M sodium acetate, pH 5.0. For experimental phasing, selenomethionine-substituted FBiB crystals were grown using protein produced in M. smegmatis cells56. Seleno replacement using PHASER 63 with the apo-FbiD structure as a search model.

**Docking dehydro-F420$^+$ into fBiB C-terminal domain.** The structure of M. smegmatis FbiD (MSMEG_2392) was solved by molecular replacement using the server Desmond (Schrödinger Release 2018-4: Desmond Molecular Dynamics). The protein was expressed overnight at 30 °C in auto-induction media and purified similarly to Mj-CoC.

Multiple sequence alignments were performed using Expresso from the ESPript 3.0 web server106.

**HPLC assays.** FBiD/CoC-FbiA-coupled activity was monitored in a reaction mixture containing 100 mM HEPES, pH 7.5, 2 mM GTP, 0.1 mM FmN, 5 mM MgCl2, 1 mM 2-PL or PEP, 1 μM FbiD, and 5 μM FbiA. The reactions were incubated at 37 °C and stopped using 20 mM EDTA at various time points. Separation of F420 in the reaction mixture was performed on an analytical HP 1100 HPLC system equipped with photodiode array and fluorescence detectors (Agilent Technologies). Samples were kept at 4 °C, and the injection volume was 20 μL. Samples were separated on a Phenomenex Luna C18 column (150 × 3 mm2, 5 μm) with a 0.2 μm in-line filter that was maintained at 30 °C. The mobile phase consisted of 100% methanol (A) and 25 mM sodium acetate buffer, pH 6.0 (B), with a gradient elution at a flow rate of 0.5 mL·min$^{-1}$ and a run time of 20 min. The gradient profile was performed as follows: 0–25 min 95%–80% B, 25–30 min 80% B, 26–37 min 95% B, 27–30 min 95% B, and a post-run of 2 min. The wavelengths used for photodiode array were 280 and 420 nm (20 nm bandpass) using a detector with 550 nm (bandwidth 50 nm). The wavelength was set for the fluorescence detector were 420 nm (excitation) and 480 nm (emission).

**LC-MS characterization of dehydro-F420$^+$ species.** Enzymic reactions were set up as described above. Ten microliters aliquots were injected onto a C18 trap cartridge (LC Packings, Amsterdam, The Netherlands) for desalting prior to chromatographic separation on a 0.3 × 100 mm2, 3.5 μm Zorbax 300SB C18 Strong cation column (Agilent Technologies, Santa Clara, CA, USA) using the following gradient at 6 μL·min$^{-1}$: 0–2 min 10% B; 2 min 10% B to 25% B, 25 min 40% B; 25 min 40% B to 90% B, 30 min 99% B; 30 min 99% B, and a post-run of 2 min. The gradients used for the photodiode array were 280 and 420 nm (20 nm bandpass) using a detector with 550 nm (bandwidth 50 nm). The wavelength was set for the fluorescence detector were 420 nm (excitation) and 480 nm (emission).

**MS/MS confirmation of reduction of dehydro-F420$^+$ species.** To confirm the reduction of the PEP moiety in vitro assays were prepared in 50 mM HEPES pH 7.5, 100 mM KCl, 5 mM MgCl2, 2 mM GTP, 0.1 mM Fo, 1 mM FbiD, 5 μM FbiA, 6.5 μM MgS-SFbiD, 10 mM DTT, 20 μM FMN, 0.2 mM NADH, 0.1 mM Ec-Flr, 2 μM Mtb-fBiaB and 1 mM L-glutamate. To minimize futile oxidation of FMN$^+$ by oxygen the reaction mixture was repeatedly evacuated and purged with nitrogen and maintained under a nitrogen atmosphere. Samples were incubated at ambient temperature for up to 36 h and stopped by addition of 20 mM EDTA. Samples were desalted using Bond Elut Elution (Escherichia coli (E. coli) host cells49. Se-SAD M. smegmatis media 47,49 in a fermenter (BioFlo® 1100, Hospira, USA). After staining and de-staining, the membrane was blocked with 3% skim milk. Samples were incubated at ambient temperature for up to 36 h and stopped by addition of 20 mM EDTA. Samples were desalted using Bond Elut Elution (Escherichia coli (E. coli) host cells49. Se-SAD M. smegmatis media 47,49 in a fermenter (BioFlo® 1100, Hospira, USA). After staining and de-staining, the membrane was blocked with 3% skim milk.
solution and blotted with anti-FLAG antibodies conjugated with HRP (DYKDDDK Tag Monoclonal Antibody, Thermo Fisher Scientific, MA1-91878-HRP, x1000 dilution).

For detection of F430 in E. coli lysate by HPLC-FLD, cells were grown overnight in media containing 2.0% tryptone, 0.5% yeast extract, 0.5% NaCl, 22 mM KH₂PO₄, 42 mM NaHPO₄, 100 ng mL⁻¹ anhydroteratocline, and 34 μg mL⁻¹ chloramphenicol at 30 °C. Cells from 500 mL of culture were pelleted by centrifugation at 16,000 × g for 10 min at 4 °C. Cell debris was pelleted by centrifugation at 16,000 × g and filtered through a 0.22 μm PVDF filter. Analysis was conducted as described previously. pSB1C3 containing only BBa_K142501 was used as a control.

Purification and analysis of Escherichio coli-derived F430. For F430 extraction, 1 L of cell culture was centrifuged at 5000 × g for 15 min. The cell pellet was resuspended in 50 mL of 75% ethanol and boiled at 90 °C in a water bath for 6 min for cell lysis. The cell extract was again centrifuged at 5000 × g for 15 min to remove cell debris and the supernatant was lyophilized. The lyophilized cell extract was re-dissolved in 1 mL of 75% ethanol and further re-dissolved in 30 mL of 75% ethanol and boiled at 90 °C in a water bath for 6 min. The cell extract was purified for F430 using a 5 mL HitTrap QFF column (GE Healthcare) as previously described. The purified F430 solution was further desalted by passing it through C18 extract column (6 mL, HyperSep C18 cartridge, Thermo Fisher Scientific). The filtrated cell extract was purified for F430 using a 5 mL HitTrap QFF column (GE Healthcare) as previously described. UV and fluorescence spectra were collected on a Varian Cary 60 and a Varian Cary Eclipse, respectively, in a 10 mm quartz cell (Hellma Analytics). Samples were buffered to pH 7.5 with 50 mM HEPS and scanned from 250 to 600 nm. For fluorescence, the excitation wavelength was 420 nm and the emission scanned from 435 to 460 nm.

Activity assays with M. smegmatis and E. coli-derived F430 were conducted with M. smegmatis FGD expressed and purified as described previously. Assays were performed in 50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 50 mM FGD, 5 μM F430, and 0–100 μM of lincomycin. The reaction was measured by following loss of F430 fluorescence at 470 nm. Apparent Kₘ and Kₘ values were calculated using the GraphPad Prism 7.04 (GraphPad Software, La Jolla, CA, USA).

2-PL synthesis. In the absence of a commercial source, 2-PL was chemically synthesized by a slight modification of the method of Ballou and Fischer. Briefly, benzy1 lactate was condensed with chlorodiphenyl phosphate in pyridine, with cooling, to give benzylidenyphosphonyl lactate. Hydrogenolysis of this material in 70% aqueous tetrahydrofuran over 10% Pd-C gave phospholactic acid as a colorless, viscous oil, which was characterized by proton, carbon, and phosphorus nuclear magnetic resonance (NMR) spectroscopy, and by MS. ¹³C NMR (DMSO-δ₆) δ 11.68 (br, 3H), 4.53 (m, 1H), 1.36 (d, J = 6.8 Hz, 3H), 1.3C NMR (DMSO-δ₆) δ 172.50 (d, Jp = 0.05 Hz), 69.56 (d, Jp = 0.04 Hz), 19.27 (d, Jp = 0.04 Hz). ¹³P NMR (DMSO-δ₆) δ 1.64. APCLI-MS found: [M + H⁺]=171.1, [M – H⁻]=169.1.

Fo purification. Fo was purified from M. smegmatis culture medium overexpressing Mib-Fbsc as described previously.

Genomic context analysis. A non-redundant CoEl database of 4813 sequences was collected using the Pfam identifier PF01996 and the InterPro classifications (IPR008325, IPR028247, and IPR023639). Archaeal CoE sequences were extracted from this dataset, resulting in a set of 1060 sequences that included representatives across 12 phyla: Crenarchaeota, Euryarchaeota, Thaumarchaeota, Candidatus Bathycarchaeota, Candidatus Diapherotrites, Candidatus Heimdallarchaeota, Candidatus Korarchaeota, Candidatus Lokarchaeota, Candidatus Marsearchaeota, Candidatus Primarchaeota, Candidatus Primarchaeota, Candidatus Odnarchaeota, and Candidatus Thoarchaeota. The genomic context (10 upstream and 10 downstream genes) of each archaeal coe was analyzed for

References
1. Greening, C. et al. Physiology, biochemistry, and applications of F430 and Fo-dependent redox reactions. Microbiol. Mol. Biol. Rev. 80, 451–493 (2016).
2. Ney, B. et al. The methanogenic redox cofactor F430 is widely synthesized by aerobic soil bacteria. ISME J. 11, 125–137 (2017).
3. Walsh, C. Naturally occurring 3-deazafolin coenzymes: biological redox roles. Acc. Chem. Res. 19, 216–221 (1986).
4. Cheeseman, P., Toons-Wood, A. & Wolfe, R. S. Isolation and properties of a fluorescent compound, FactorF430, from Methanobacterium strain M.O.H. J. Bacteriol. 112, 527–531 (1972).
5. Eirich, L. D., Vogels, G. D. & Wolfe, R. S. Proposed structure for coenzyme F430 from Methanobacterium. Biochemistry 17, 4583–4593 (1978).
6. Thauer, R. K., Kaster, A.-K., Seedorf, H., Buckel, W. & Hedderich, R. Methanogenic archaea: ecologically relevant differences in energy conservation. Nat. Rev. Microbiol. 6, 579–591 (2008).
7. Lackner, G., Peters, E. E., Helfrich, E. J. N. & Piel, J. Insights into the lifestyle of uncultured bacterial natural product factories associated with marine sponges. Proc. Natl. Acad. Sci. USA 114, E347–E356 (2017).
8. Jay, Z. J. et al. Marsearchaeota are an aerobic archaeal lineage abundant in geothermal iron oxide microbial mats. Nat. Microbiol. 3, 752–740 (2018).
9. Purwantini, E. & Mukhopadhyay, B. Conversion of NO₂ to NO by reduced M. smegmatis cytochrome P450. Biochemistry 57, 10633–10638 (2008).
10. Gurumurthy, M. et al. A novel F420-dependent anti-oxidant mechanism demonstrates a requirement for M. tuberculosis cytochrome P450 in intracellular NO release. Science 322, 1392–1395 (2008).
11. Wang, P., Bashiri, G., Gao, X., Sawaya, M. R. & Yang, U. Uncovering the enzymes that catalyze the final steps in oxytetracycline biosynthesis. J. Am. Chem. Soc. 135, 7138–7141 (2013).
12. Coats, J. H., Li, G. P., Ming-ST, K. & Yurek, D. A. Discovery, production, and biological assay of an unusual flavenoid cofactor involved in lincomycin biosynthesis. J. Antibiot. 42, 472–474 (1989).
13. Ishikawa, H., Bashiri, G. & Kelly, W. L. Biosynthesis of the thiopeptins and identification of an F430H₃-dependent dehydropeptidase reducetase. J. Am. Chem. Soc. 140, 10749–10756 (2018).
14. Taylor, M. C. et al. Identification and characterization of two families of F430H₃-dependent reductases from Mycobacteria that catalyse aflatoxin degradation. Mol. Microbiol. 78, 561–575 (2010).
15. Greening, C. et al. Mycobacterial F430H₃-dependent reductases promiscuously reduce diverse compounds through a common mechanism. Front. Microbiol. 8, https://doi.org/10.3389/fmicb.2017.01000 (2017).

21. Li, H., Graupner, M., Xu, H. & White, R. H. Cofactor tail length modulates catalysis of bacterial F420–

22. Graupner, M. & White, R. H. Biosynthesis of the phosphodiester bond in coenzyme F420 in the Methanobacca. Biochemistry 40, 10859–10872 (2001).

23. Guerra-Lopez, D., Daniels, L. & Rawat, M. Mycobacterium smegmatis mc2 155 fbc1.ett, B. DMEG_2392 involved in dihydrophosphate dye descolorization and coenzyme F420 biosynthesis. Microbiology 153, 2724–2723 (2007).

24. Rifat, D. et al. Mutations in Rv2983 as a novel determinant of resistance to

25. Cole, S. T. et al. Deciphering the biology of

26. Ash, D. E., Goodhart, P. J. & Reed, G. H. ATP-dependent phosphorylation of

27. Izui, K., Matsuda, Y., Kameshita, I., Katsuki, H. & Woods, A. E.

28. Bashiri, G., Rehan, A. M., Greenwood, D. R., Dickson, J. M. J. & Baker, E. N.

29. Andersen, K. R., Leksa, N. C. & Schwartz, T. U. Optimized

30. Fieschi, F., Nivière, V., Frier, C. & Décout, J.-I. & Fontecave, M.

31. Nocek, B. et al. Structure of an amide bond forming F420-

32. Ash, D. E., Goodhart, P. J. & Reed, G. H.

33. Lake, M. W., Temple, C. A., Rajagopalan, K. V. & Schindelin, H. The crystal

34. Lake, M. W., Temple, C. A., Rajagopalan, K. V. & Schindelin, H.

35. Fieschi, F., Nivière, V., Frier, C. & Décout, J.-I. & Fontecave, M.

36. Evans, P. Scaling and assessment of data quality. Acta Crystallogr. D 62, 72–82 (2006).

37. Fieschi, F., Nivière, V., Frier, C. & Décout, J.-I. & Fontecave, M.

38. Collier, F. et al. A conserved phosphatase destroys toxic glycolytic side products in mammals and yeast. Nat. Chem. Biol. 12, 601–607 (2016).

39. Walker, S. R., Cumming, H. & Parker, E. J. Substrate and reaction intermediates inhibit F420

40. Izui, K., Matsuda, Y., Kameshita, I., Katsuki, H. & Woods, A. E.

41. Bashiri, G., Rehan, A. M., Greenwood, D. R., Dickson, J. M. J. & Baker, E. N.

42. Collard, F. et al. Conceived project, designed experiments, performed experiments, analyzed results, and co-wrote the manuscript; J.A., the Australian Research Council, and the National Health and Medical Research Council, an NHMRC New Investigator Grant (C.G.; 1142699), an ARC DECRA Fellowship (C.G.; 1601010374), a Sir Charles Hercus Fellowship (G.B.), technical assistance. This research is supported by a Sir Charles Hercus Fellowship (G.B.) and an NHMRC Project Grant (1170242).
