Elongation of the Poly-\(\gamma\)-glutamate Tail of F\(_{420}\) Requires Both Domains of the F\(_{420}\):\(\gamma\)-Glutamyl Ligase (FbiB) of Mycobacterium tuberculosis*

Received for publication, September 1, 2015, and in revised form, February 6, 2016. Published, JBC Papers in Press, February 9, 2016, DOI 10.1074/jbc.M115.689026

Ghader Bashiri†1, Aisyah M. Rehan‡1,2, Sreevalsan Sreebhavan‡, Heather M. Baker†, Edward N. Baker†, and Christopher J. Squire‡3

From the †Structural Biology Laboratory, School of Biological Sciences and Maurice Wilkins Centre for Molecular Biodiscovery, and the ‡Faculty of Medical and Health Sciences, University of Auckland, Auckland 1010, New Zealand

Cofactor F\(_{420}\) is an electron carrier with a major role in the oxido-reductive reactions of Mycobacterium tuberculosis, the causative agent of tuberculosis. A \(\gamma\)-glutamyl ligase catalyzes the final steps of the F\(_{420}\) biosynthesis pathway by successive additions of \(\gamma\)-glutamate residues to F\(_{420}\)-0, producing a poly-\(\gamma\)-glutamate tail. The enzyme responsible for this reaction in archaea (CofE) comprises a single domain and produces F\(_{420}\)-2 as the major species. The homologous M. tuberculosis enzyme, FbiB, is a two-domain protein and produces F\(_{420}\) with predominantly 5–7 \(\gamma\)-glutamate residues in the poly-\(\gamma\)-glutamate tail. The N-terminal domain of FbiB is homologous to CofE with an annotated \(\gamma\)-glutamyl ligase activity, whereas the C-terminal domain has sequence similarity to an FMN-dependent family of nitroreductase enzymes. Here we demonstrate that full-length FbiB adds multiple \(\gamma\)-glutamate residues to F\(_{420}\)-0 in vitro to produce F\(_{420}\)-5 after 24 h; communication between the two domains is critical for full \(\gamma\)-glutamyl ligase activity. We also present crystal structures of the C-terminal domain of FbiB in apo-, F\(_{420}\)-0-, and FMN-bound states, displaying distinct sites for F\(_{420}\)-0 and FMN ligands that partially overlap. Finally, we discuss the features of a full-length structural model produced by small angle x-ray scattering and its implications for the role of N- and C-terminal domains in catalysis.

The cofactor F\(_{420}\) is a flavin derivative that is sporadically distributed among microorganisms, mainly archaea and actinobacteria (including mycobacteria). F\(_{420}\) has been emerging as a new player in the biology of mycobacteria (1), with increasing numbers of F\(_{420}\)-utilizing proteins characterized from different mycobacterial species (2–8). This cofactor has been suggested to protect Mycobacterium tuberculosis, the causative agent of tuberculosis, against oxidative and nitrosative stress during pathogenesis (9–11). At the biochemical level, cofactor F\(_{420}\) functions as a hydride transfer agent in oxido-reductive reactions with a lower redox potential than that of NAD(P)\(^+\) (12).

The biosynthesis pathway of cofactor F\(_{420}\) has been investigated in both archaeal and mycobacterial species. In the current view of the proposed pathway, the first intermediate with the complete chromophore (7,8-didemethyl-8-hydroxy-5-deazariboflavin (FO))\(^4\) is produced by FO synthase (FbiC in mycobacteria (13) and CofGH in archaea (14)). A transferase enzyme (FbiA in mycobacteria (15) and CofD in archaea (16)) subsequently catalyzes the addition of a 2-phospho-\(\gamma\)-lactate moiety to FO to produce F\(_{420}\)-0 (F\(_{420}\) with no poly-\(\gamma\)-glutamate tail). The final step of the pathway is performed by a \(\gamma\)-glutamyl ligase (FbiB in mycobacteria (15) and CofE in archaea (17)) that catalyzes successive additions of \(\gamma\)-glutamate residues to F\(_{420}\)-0 (Fig. 1A).

The length of the poly-\(\gamma\)-glutamate tail varies between archaeal and mycobacterial species; in archaea, two \(\gamma\)-glutamate residues are seen (18), whereas in mycobacteria, up to nine residues are present (3, 19). There exists an intriguing difference between the enzymes responsible for this reaction in these microorganisms, with CofE having only one domain, whereas FbiB is a two-domain protein. The N-terminal domain of FbiB is annotated as a \(\gamma\)-glutamyl ligase with sequence similarities to CofE, whereas the C-terminal domain has sequence similarity to an FMN-dependent family of nitroreductase enzymes. Functional homology to nitroreductases, however, seems unlikely, and it is reasonable to hypothesize that the C-terminal domain of FbiB facilitates elongation of the poly-\(\gamma\)-glutamate tail of cofactor F\(_{420}\) in vitro.

Here we describe the structural and functional characterization of the FbiB protein from M. tuberculosis, demonstrating that the full-length enzyme is capable of adding multiple \(\gamma\)-glutamate residues to F\(_{420}\)-0 in vitro. Our results also indicate that communication between the two domains of FbiB is critical for full \(\gamma\)-glutamyl ligase activity, because the N-terminal domain is capable of producing only F\(_{420}\)-1. We have also determined the crystal structure of the C-terminal domain of FbiB in apo-

\(\text{FO, 7,8-didemethyl-8-hydroxy-5-deazariboflavin; SAXS, small angle x-ray scattering; WAXS, wide angle x-ray scattering; TEV, tobacco etch virus.}\)
F$_{420}$-0-, and FMN-bound states (1.9, 2.05, and 2.1 Å resolution, respectively). These structures indicate specific and distinct binding sites for F$_{420}$-0 and FMN ligands that are adjacent and partially overlapped. Despite the presence of a conserved FMN binding site in this domain, FMN does not appear to have a role in the reaction catalyzed by FbiB and is probably a remnant of the ancestral FMN-dependent nitroreductase enzyme. A full-length structural model of FbiB produced by small angle x-ray scattering (SAXS), using our x-ray crystal structure of the C-terminal domain and an N-terminal homology model, shows that the two domains are separated in space and are linked by a flexible α-helical segment. We discuss the implications of this full-length model for our functional understanding of the C-terminal domain.

**Experimental Procedures**

**PCR Amplification and Cloning**—The open reading frame encoding the full-length FbiB protein (Rv3262) (20) was amplified from *M. tuberculosis* H37Rv genomic DNA using the primers outlined in Table 1. The cloning was conducted using the Gateway® cloning system (21). The amplified PCR product was used to produce entry clones by performing a BP reaction. Positive entry clones were selected on LB agar medium supplemented with 50 µg/ml kanamycin and were then verified using BsrGI digestion and sequencing. The resulting entry clones were used to clone the full-length construct into pDEST17 (21) and pDESTsmg (22) vectors using an LR reaction. The expression construct for pDEST17 was selected on LB agar plates containing 100 µg/ml ampicillin. Selection of pDESTsmg construct was performed on low salt LB agar plates, pH 8.0, containing 50 µg/ml hygromycin B. All expression constructs were subsequently verified using BsrGI digestion. Two additional constructs were also prepared, using the same protocol, to express N-terminal (FbiB(11–249)) and C-terminal (FbiB(249–448)) domains of FbiB in pDESTsmg vector (Table 1).

**Expression and Purification**—FbiB constructs were expressed in *Escherichia coli* BL21(DE3)pRP and *Mycobacterium smegmatis* mc² 4517 (23) cells. In both cases, protein expression was performed in autoinduction medium as described previously (24). Protein expression in *E. coli* was started at 37 °C for 4 h, followed by overnight incubation at 18 °C. *M. smegmatis* cultures were supplemented with 0.05% (v/v) Tween 80, and protein expression was carried out for 3–4 days at 37 °C (24).

All three FbiB constructs were cloned with an N-terminal His$_6$ tag on both pDEST vectors to facilitate the subsequent purification steps. The His$_6$ tag on both pDEST vectors is cleavable using tobacco etch virus (TEV) protease. All constructs were purified from *E. coli* and *M. smegmatis* cells using the same procedure, as described below. The cells were harvested and resuspended in 20 mM HEPES, pH 7.0, 150 mM NaCl, 20 mM imidazole, 1 mM β-mercaptoethanol. The cells were then lysed 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 × g to separate the insoluble material. The recombinant proteins were first purified using an immobilized metal affinity chromatography step by loading the supernatant onto a HisTrap FF 5-ml nickel-affinity column (GE Healthcare) that had been pre-equilibrated in the lysis buffer. The column was washed with the lysis buffer, and the protein was subsequently eluted using a gradient of imidazole in the buffer. Appropriate protein fractions were pooled and dialyzed at 4 °C in the presence of rTEV protease (25) to remove the N-terminal His$_6$ tag. The His$_{6-}$tagged rTEV protease, encoded in a pProEX HTa expression vector, was produced earlier from *E. coli* Rosetta TM (DE3) pLysS cells (Novagen), as described previously (25). After overnight incubation of the purified FbiB proteins with rTEV protease, a subtractive immobilized metal affinity chromatography step was performed to remove the cleaved protein from uncut protein and rTEV protease. The resulting protein fraction was concentrated and then injected onto a size exclusion Superdex 200 10/300 column (GE Healthcare) pre-equilibrated in 20 mM HEPES, pH 7.0, 150 mM NaCl, 1 mM β-mercaptoethanol.

**Binding Assays**—Fluorescence spectroscopy was performed using an EnSpire® multimode plate reader (PerkinElmer Life Sciences). The fluorescence intensity was measured using a black 96-well plate with total reaction volumes of 100 µl in triplicates. Excitation and emission wavelengths of 420 and 480 nm, respectively, were used to monitor the intrinsic F$_{420}$ fluorescence. For FMN binding assays, the excitation wavelength was set at 445 nm, and emission wavelength was set at 525 nm. To determine the dissociation constant, the protein samples (0.1 µM) were incubated with either F$_{420}$ (0.001–20 µM) or FMN (0.01–100 µM) and left for 30 min at ambient temperature before fluorescence measurements. The binding reactions contained 20 mM HEPES, pH 7.0, 150 mM NaCl, 1 mM β-mercaptoethanol and were corrected against a control lacking the FbiB protein. Ligand binding data were fitted using one- or two-site binding models (SigmaPlot version 12.5).

**Activity Assays**—The γ-glutamyl ligase activity (17) was measured in 50-µl reactions containing different FbiB constructs (1...
The optimized reaction mixture included 50 mM HEPES, pH 8.5, 100 mM NaCl, 5 mM MnCl₂, 10 mM L-glutamate, 5 mM GTP, and 2 μM F₄₂₀-0. The F₄₂₀-0 substrate was prepared by enzymatic hydrolysis of the poly-γ-glutamate tail of F₄₂₀ using carboxypeptidase G (in 50 mM Tris-HCl, pH 7.5, 0.1 M NaCl, 0.2 mM ZnSO₄) (17). The possible effect of FMN on the polyglutamylation reaction was also tested by including a range of FMN concentrations (0–1 mM) in duplicate activity assays. The reactions were incubated at 37 °C and stopped using 20 mM EDTA after various time points.

HPLC Analysis—Separation of F₄₂₀ species was performed on an Agilent HP 1100 HPLC system equipped with photo-diode 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 mm, 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 and a run time of 30 min. The gradient profile was performed as follows: 0–25 min 95–80% B, 25–26 min 80% B, 26–27 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 (bandwidth 20 nm) using a reference of 550 nm (bandwidth 50 nm). The wavelengths used for the fluorescence detector were 420 nm (excitation) and 480 nm (emission).

LC-MS Analysis—After HPLC separation, F₄₂₀ samples were analyzed using an Agilent series 1200 liquid chromatography instrument coupled with an Agilent 6460 jet stream triple quadrupole mass spectrometer (Agilent Technologies, Santa Clara, CA). The injection volume and sample separation were performed as described above, with the mobile phase consisting of 100% methanol with 0.02% ammonia (A) and 25 mM ammonium acetate buffer, pH 6.0 (B). Ionization was achieved using electrospray ionization in either negative or positive mode.

Crystalization and Data Collection—The C-terminal domain of FbiB (FbiB(249–448)) was crystallized as described previously (26). The best diffracting crystals were obtained in 25% (w/v) PEG 3350, 0.35 M Li₂SO₄. The crystals were cryo-protected in 70% N-paratone and 30% mineral oil (v/v) before flash-freezing in liquid nitrogen.

For experimental phasing, preformed apo-crystals were soaked for 10 min in 0.5 M KBr solution that was prepared in the cryo-solution (25% PEG 3350 (w/v), 0.35 M LiSO₄, 25% glycerol (v/v)) before being flash-cooled in liquid nitrogen. Bromide-multiwavelength anomalous diffraction data sets were collected at three wavelengths at the Australian Synchrotron. Data collection statistics are summarized in Table 2.

Crystals of the C-terminal domain of FbiB in complex with ligands were obtained by soaking preformed apo crystals in precipitant solutions containing the ligands. The FMN complex was obtained by overnight soaking of apo-crystals in the cryo-solution (25% PEG 3350 (w/v), 0.35 M LiSO₄, 25% glycerol (v/v)) containing 5 mM FMN. For preparation of crystals in complex with F₄₂₀-0, apo-crystals were soaked with F₄₂₀-0 that was prepared for activity assays as mentioned earlier. F₄₂₀-0 solution in water was concentrated to near dryness and redissolved in the cryo-solution (25% PEG 3350 (w/v), 0.35 M Li₂SO₄, 25% glycerol (v/v)) to a final concentration of 4.5 mM, after which the crystals were soaked for 2 h.

Structure Determination and Refinement—All data sets were indexed and processed using XDS (27), re-indexed using POINTLESS, and scaled with SCALA from the CCP4 program suite (28, 29). For structure determination using bromide-multiwavelength anomalous diffraction, the bromide sites and their occupancies were found using SHELXD (30), as implemented in the autoSHARP program suite (31). SHARP was then used for substructure refinement and phasing (32), followed by phase improvements by the SOLOMON density modification program (33). Cycles of automatic model building by ARP/WARP (34) resulted in a partial protein model that was then used for further automated model building with the application of non-crystallographic symmetry restraints on the four molecules of the asymmetric unit. The final model was completed manually using COOT (35). Water molecules were identified by their spherical electron density and appropriate hydrogen bond geometry with the surrounding structure. Following each round of manual model building, the model was refined using
REFMAC5 (36), against the data to 1.9 Å resolution. The auto-BUSTER refinement program (37) was used in the final stages of refinement, both with and without TLS refinement parameters. Full refinement statistics are shown in Table 3.

Structures of the C-terminal domain of FbiB in complex with FMN and F_{420}-0 were solved by molecular replacement using PHASER (38) with the apo-FbiB structure as a search model. All structures were refined by cycles of manual building using Phaser (38) with the apo-FbiB structure as a search model. All models, as described fully below.

Full-length Model of FbiB from SAXS and Homology Modeling—Protein samples were extensively dialyzed against 20 mM HEPES, pH 7.5, 150 mM NaCl, 5% glycerol (v/v), and 1 mM tris-[2-carboxyethyl]phosphine. This buffer was also used as the buffer control and to dilute protein samples. SAXS data were collected on the Australian Synchrotron SAXS/WAXS beamline and processed using the scatterBrain software package (39). In brief, protein samples and buffer controls were diluted and aliquoted into 96-well plates and rubber lids placed on the wells to prevent evaporation. The 96-well plates were mounted on the SAXS/WAXS beamline on a temperature-controlled mount at 283 K for autosampling and capillary flow data acquisition. Solutions were flowed through thin walled quartz capillaries, and data acquisition comprised consecutive 1-s x-ray exposures of the flowing samples to minimize radiation damage. Appropriate images were combined, and buffer scattering was carefully subtracted using the scatterBrain software (39). Data quality and derived parameters were calculated using the ATSAS software package (40). Details of SAXS data collection parameters and statistics are shown in Table 4.

A structure of the N-terminal domain of FbiB was produced by homology modeling using the Phyre2 software (41) and the Protein Data Bank coordinate set 2PHN. The N-terminal homology model and the C-terminal crystal structure were combined to produce a full-length model using the SAXS data and a rigid body minimization protocol using SASREF (40) with a distance restraint to keep the termini of the two domains in appropriate proximity, effectively linked. Numerous possible models were assessed against the SAXS scattering profile using the CRYSOl software (40) before, eventually, the best fit was discovered as a 50:50 combination of two complete full-length models, as described fully below.

**Results**

**FbiB Expression and Purification**—The full-length, N-terminal, and C-terminal constructs of FbiB were expressed as soluble proteins in both *E. coli* and *M. smegmatis* cells using an autoinduction protocol. All three constructs were purified using immobilized metal affinity chromatography and size exclusion chromatography steps.

**Functional Characterization**—The FbiB constructs expressed in *E. coli* were used for functional studies to alleviate the complications arising from the presence of co-purified F_{420} in the proteins expressed in *M. smegmatis* (Fig. 1B). The γ-glutamyl ligase activity of FbiB constructs was investigated by monitoring the addition of 1-glutamate residues to F_{420} using HPLC, as described previously for the CoFe protein (16). The F_{420}-0 substrate was prepared by enzymatic hydrolysis of the poly-γ-glutamate tail of F_{420} using carboxypeptidase G (17), and the resulting F_{420}-0 was used as a substrate for FbiB activity experiments and also for crystallographic binding studies.

The intrinsic fluorescence of F_{420} was used to monitor the addition of 1-glutamate residues to F_{420} during enzymatic reactions. A range of different conditions was optimized for the γ-glutamyl ligase activity of the full-length FbiB protein, including pH (6.0–9.0), monovalent (Na\(^+\) and K\(^+\)) and divalent cation (Mg\(^{2+}\) and Mn\(^{2+}\)) composition, nucleotides (GTP, dGTP, ATP, and dATP), and also various time points and temperatures. FbiB showed the highest activity at pH 8.5, the same pH dependence displayed by the CoFe protein (17). A combination of Na\(^+\) and Mn\(^{2+}\) produced the highest activity in FbiB, in contrast to the previously reported dependence on K\(^+\) and

### Table 3

**Crystal structure refinement statistics**

<table>
<thead>
<tr>
<th></th>
<th>FbiB-F_{420}</th>
<th>FbiB-FMN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein Data Bank code</td>
<td>4XOM</td>
<td>4XQO</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>23.3/26.7</td>
<td>21.2/24.5</td>
</tr>
<tr>
<td>R_{merge} (R_{all})</td>
<td>0.010</td>
<td>0.011</td>
</tr>
<tr>
<td>Root mean square deviation from ideality Bonds (Å)</td>
<td>1.35</td>
<td>1.46</td>
</tr>
<tr>
<td>Average B factors (Å(^2))</td>
<td>32.5</td>
<td>33.5</td>
</tr>
<tr>
<td>Sulfates (Å(^2))</td>
<td>53.7</td>
<td>59.2</td>
</tr>
<tr>
<td>Waters</td>
<td>31.3</td>
<td>34.7</td>
</tr>
<tr>
<td>Ramachandran statistics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Favored (%)</td>
<td>97.9</td>
<td>97.6</td>
</tr>
<tr>
<td>Allowed (%)</td>
<td>100.0</td>
<td>99.5</td>
</tr>
<tr>
<td>Outliers (%)</td>
<td>0.0</td>
<td>0.2</td>
</tr>
<tr>
<td>Molprobity score; percentile</td>
<td>1.30 100th</td>
<td>1.51 100th</td>
</tr>
</tbody>
</table>

### Table 4

**Small angle x-ray scattering parameters and statistics**

<table>
<thead>
<tr>
<th>Data collection parameters</th>
<th>FbiB-F_{420}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beamline</td>
<td>AS SAXS/WAX</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.0322</td>
</tr>
<tr>
<td>Detector</td>
<td>Decestra-Pilatus 1M</td>
</tr>
<tr>
<td>Camera length (mm)</td>
<td>1600</td>
</tr>
<tr>
<td>g range (Å(^{-1}))</td>
<td>0.006 – 0.6</td>
</tr>
<tr>
<td>Total sample volume (µl)</td>
<td>100</td>
</tr>
<tr>
<td>Exposure time/image (s)</td>
<td>1</td>
</tr>
<tr>
<td>No. of images/sample</td>
<td>20</td>
</tr>
<tr>
<td>Concentration range (mg/ml)</td>
<td>0.25–0.50</td>
</tr>
<tr>
<td>Temperature (K)</td>
<td>283</td>
</tr>
</tbody>
</table>

**Structural parameters**

| I(0) (cm\(^{-1}\)) (from P(r)) | 0.557 ± 0.002 |
| I(0) (cm\(^{-1}\)) (from Guinier) | 0.561 ± 0.001 |
| R_{p} (Å) (from P(r))          | 39.82 ± 0.001 |
| R_{g} (Å) (from Guinier)       | 39.8 ± 0.2    |
| R_{exp} (Å)                    | 130.0         |
| Porod volume estimate (Å\(^3\)) | 56.501    |
| Dry volume calculated from sequence (Å\(^3\)) | 57.567 |

**Software**

<table>
<thead>
<tr>
<th>Primary data collection</th>
<th>scatterBrain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data processing</td>
<td>scatterBrain</td>
</tr>
<tr>
<td>Rigid body modelling</td>
<td>SASREF</td>
</tr>
<tr>
<td>Computation of model intensities</td>
<td>CRYSOl</td>
</tr>
</tbody>
</table>

* Full details of the beamline specifications are available at the Australian Synchrotron website.
* For a 5 mg/ml sample.
Mg$^{2+}$ for CoE (17). The enzyme was only active in the presence of GTP, with no observed activity for dGTP, ATP, and dATP nucleotides. The enzymatic assays for all FbiB constructs were subsequently conducted using the optimized condition (50 mM HEPES, pH 8.5, 100 mM NaCl, 5 mM MnCl$_2$, 10 mM l-glutamate, 5 mM GTP, and 2 μM F$_{420}$-0) and were incubated for periods of up to 1 week at 37 °C.

The functional assays show that full-length FbiB can convert F$_{420}$-0 to F$_{420}$ molecules with varying numbers of residues in the poly-L-glutamate tail (Fig. 2A). F$_{420}$-2 production could be detected within 1 h, with no higher order F$_{420}$ molecules appearing within 2 h. Incubation for 24 h showed species as large as F$_{420}$-5, and by 72 h (data not shown), the products resolved into two species with much earlier retention time than that of F$_{420}$-5; although attempted, mass spectrometry analysis of these two large species was unsuccessful. This result, however, is not achieved by either the N- or C-terminal domains alone (Fig. 2, B and C, respectively), nor by a mixture of these two separate domains (Fig. 2D). Our results show that the N-terminal domain of FbiB adds one L-glutamate residue (to produce F$_{420}$-1), albeit in a very slow manner compared with that of the full-length protein.

**Ligand Binding**—Expression of FbiB constructs in M. smegmatis host cells resulted in co-purification of F$_{420}$ with a distinct coloration of the full-length and N-terminal domain constructs. Mass spectrometry analysis confirmed that F$_{420}$ species containing up to 11 L-glutamate residues in the poly-γ-glutamate tail co-purify with the full-length protein (Fig. 1B). Given that E. coli does not produce F$_{420}$, the proteins expressed in E. coli cells were not expected to have any ligands co-purified. The C-terminal domain, however, showed a faint yellow color upon purification suggesting co-purification of flavin-like ligands from E. coli.

Fluorescence spectroscopy was used to investigate binding of F$_{420}$ and other flavins (e.g. FMN) to FbiB constructs and to determine their dissociation constant ($K_d$) (Fig. 3). The results show that the full-length (two-site model; $K_d$ = 0.2 ± 0.4 μM and $K_d$ = 3.0 ± 0.9 μM), N-terminal (one-site model; $K_d$ = 1.4 ± 0.1 μM), and C-terminal domains (one-site model; $K_d$ = 1.47 ± 0.07 μM) could each bind F$_{420}$ in solution. Surprisingly, the full-length and C-terminal constructs could also bind FMN although with a lower affinity (one-site model; $K_d$ = 14.7 ± 0.9 μM). Although binding with 10-fold lower affinity to the C-terminal domain of FbiB, FMN was subsequently included in the functional studies of the different domain constructs to assess a possible role in catalytic function or inhibition/regulation.

**Structure Determination**—The requirement of the C-terminal domain for full activity and its binding to both F$_{420}$ and FMN prompted us to pursue x-ray crystal structures of three FbiB constructs. Neither full-length FbiB nor its N-terminal domain could be crystallized. The C-terminal domain of FbiB (FbiB(249–448)), expressed in M. smegmatis cells, was successfully crystallized (26), however, and its structure was determined at 1.9 Å resolution by a multiwavelength anomalous dispersion method using crystals soaked in KBr solution (42) (Table 1). The unliganded FbiB C-terminal domain was crystallized in tetragonal space group P4$_1$2$_1$2 and contains four molecules in the asymmetric unit. F$_{420}$-0 and FMN-bound structures were produced by soaking ligand into preformed crystals and were solved using the unliganded structure for molecular replacement (Table 3).
Overall Structure—The four molecules in the asymmetric unit are organized as two sets of dimers that are related by a 2-fold non-crystallographic symmetry axis. The more extensive chain A-chain B dimer interface buries an average solvent-accessible surface area of 3247 Å²/monomer, representing 27% of each monomer’s surface, as assessed by the PDBePISA server (43), and contains 33 hydrogen bonds, 15 salt bridges, and 204 non-bonded contacts. The protein displays a fold that is typical of the family of FMN-dependent nitroreductases (Pfam PF00881). This fold is based around a central five-stranded β-sheet, made up of four antiparallel β-strands together with a fifth one (parallel to the first) that is contributed by the C-terminal residues of the opposing monomer (Fig. 4A). The β-sheet is flanked by two helices on the internal side, involved in dimerization, and three helices on the external (surface-exposed) side.

FMN Binding—Although FMN is not a previously known ligand for the reaction catalyzed by FbiB, our results show FMN binding to FbiB in solution. Four molecules of FMN are located unambiguously from the electron density maps (Fig. 4B, top), one in each protomer and binding in a conserved mode. Only two F₄₂₀-0 molecules (bound to chains A and B), however, have well-defined electron density beyond the phosphate group, enabling two atoms of a lactyl moiety to be modeled. Superimposition of the unliganded and F₄₂₀-bound structures of the C-terminal domain shows little change in the protein on ligand binding, with an average root mean square deviation of 0.31 Å over 782 aligned Cα atomic positions.

FIGURE 2. Analysis of FbiB reaction products in vitro. The HPLC traces of the reactions performed using different FbiB constructs (fluorescence versus elution time). In the presence of the full-length construct (A), F₄₂₀ molecules with various lengths of poly-γ-glutamate tail appear over time, whereas the N-terminal domain alone (B) shows only the production of F₄₂₀-1. The C-terminal domain alone does not catalyze the reaction (C). The reactions using a combination of separate N- and C-terminal domains produce only F₄₂₀-1 (D). The HPLC traces are shown for a time course up to 24 h as indicated. Selected peaks are labeled with the corresponding m/z values and F₄₂₀ species.

M. tuberculosis γ-Glutamyl Ligase FbiB
The F420-0 molecule binds to the C-terminal domain in a pocket near the dimer interface (Fig. 4A) with the 8-hydroxy-5-deazaalloxazine ring system buried deeply in the protein, whereas the tail of the molecule binds in a solvent-exposed channel (Fig. 5, D and E). The 8-hydroxy-5-deazaalloxazine chromophore is sandwiched between hydrophobic elements, at the back primarily by the side chain of Phe403 and at the front by a π-stacking interaction with the side chain of Trp317 (as illustrated in Fig. 5F). Intriguingly, only a single hydrogen bond is formed directly to the protein with most other polar/charged elements of the ligand binding indirectly to the protein through water-mediated hydrogen bonding.

Structural Comparisons—A search using the PDBeFold server (44) for structural homologs using one FbiB C-terminal dimer produces numerous matches, two of which are shown in multiple sequence alignment in Fig. 4C. The putative nitroreductase from Clostridium difficile in complex with FMN (Protein Data Bank entry 3GFA) shows only 23% sequence identity but a root mean square deviation of 2.1 Å over 338 Ca atoms and a nitroreductase family protein from Agrobacterium tumefaciens (Protein Data Bank entry 3K6H), also in complex with FMN, that displays no detectable sequence similarity by BLAST but has a root mean square deviation of 2.4 Å over 291 Ca atoms. Intriguingly, 3GFA shows a very similar dimer interface to that of the FbiB C-terminal domain with an average solvent-accessible surface area of 3538 Å² (31% of each monomer’s surface). The most obvious difference between the FbiB and the homologous structures is in regions illustrated in Fig. 4D. Movement of the β2-α5 loop and α4 helix produces the F420 binding site; the β5-β6 loop shift opens up the bottom of the FbiB structure near the FMN binding site.

Building a Full-length Model Using SAXS—The full-length FbiB protein was subjected to SAXS at the Australian Synchrotron to characterize its structure in solution. Analysis of the SAXS data (Table 4 and Fig. 6A), including the calcula-
tion of various analytical curves, $R_g$ and $(0)/c$ values from a concentration series, shows that the scattering profile is not concentration-dependent and, with Guinier plots displaying linearity at low angle (Fig. 6A, inset), is indicative of no inter-particle interference or aggregation. A $P(r)$ function (not shown) suggests a maximum scattering particle dimension of 130 Å, and Kratky analysis shows a well folded protein.

A full-length model of FbiB was produced via rigid body modeling by combining our structure of the C-terminal domain of FbiB with an N-terminal homology model and using the
SAXS data for minimization. The homology model is based on the structure of an F420-γ-glutamyl ligase from Archaeoglobus fulgidus that shares 40% identity and 55% similarity with the FbiB N-terminal domain over 215 residues.

The full-length model shows the two domains as separated and linked by an α-helical segment (Figs. 6B and 7). The SAXS profile (Fig. 6A) is best fit by a combination of two models of the full-length protein, one “straight” (Fig. 6B, top) and one “bent” (Fig. 6B, bottom) and in equal proportions (50:50). A surface representation of the straight model (Fig. 7) shows the relative locations of the putative N-terminal active site cavity and the C-terminal FMN/F420 ligand binding sites. Directly adjacent to the N-terminal active site cavity is a surface groove lined with numerous basic residues that extends toward the C-terminal domain and into solvent space. The implications of this surface groove and the other features of this full two-domain model are discussed below.

Discussion

FbiB Function—The reaction catalyzed by CoE adds two L-glutamate residues to F420 to produce F420-2 in archaea (17). The presence of an additional domain in the mycobacterial FbiB protein suggested that this domain could assist the N-terminal domain in producing F420 molecules with longer poly-γ-glutamate tails in mycobacterial species. Our experimental results indeed demonstrate that the full-length FbiB...
protein does produce F_{420} molecules with multiple L-glutamate residues in the poly-L-glutamate tail (Fig. 1B). The addition of further L-glutamate residues to F_{420}-2 is evidently the rate-limiting step of the process because it takes a much longer time for the F_{420} species with longer poly-L-glutamate tails to appear. When F_{420} is purified from mycobacterial cells, the major species found contain between 5 and 7 L-glutamate residues in their poly-L-glutamate tails (3, 19). In addition, the full-length FbiB protein expressed in M. smegmatis shows co-purification of F_{420} molecules with up to 11 L-glutamate residues incorporated, with the major species having 6–9 residues. After 72 h in our in vitro experiments, we observe two major products presumably containing a larger number of L-glutamate residues in the poly-L-glutamate tail than F_{420}-5, but our inability to characterize these species by mass spectrometry does not allow us to comment further on the identity of these reaction products.

Investigation of the N-terminal and C-terminal domains of FbiB shows that these domains, by themselves, are unable to add L-glutamate residues to F_{420}-0, although the N-terminal domain shows production of F_{420}-1 in a very slow manner. This is a surprising observation because the N-terminal domain is homologous with the CofE protein of archaea (17), and we would have expected it to show L-glutamyl ligase activity on its own. Our observation that both F_{420}-0 and FMN bind to the C-terminal domain suggests a regulatory role for this domain in the reaction catalyzed by FbiB, perhaps both as a requirement for normal catalysis and as a sensor for the relative concentrations of F_{420} and FMN, in a feedback regulatory role. We could not, however, observe any effect of FMN on the activity of full-length FbiB over a wide range of concentrations, up to 500 times that of F_{420}-0 in the reaction mixture. It has been suggested that the intracellular concentration of FMN in M. smegmatis is ~10-fold higher than that of F_{420} (4).
M. tuberculosis γ-Glutamyl Ligase FbiB

Comparison of F$_{420}$ and FMN Binding—Given our observed binding of F$_{420}$ and FMN to FbiB in solution, we speculated that they would share the same binding site in the C-terminal domain. Surprisingly, this was not the case, as apparent in our crystal structure, although the two binding sites are closely adjacent in the protein structure.

A comparison of the two binding sites and the chemical structures of F$_{420}$ and FMN explains why “crossover” is not seen in the ligand binding. Although both binding sites have similar electrostatic surface potentials, the F$_{420}$ chromophore extends a hydroxyl group toward Asp$^{320}$, forming a hydrogen bond (Fig. 5E). If FMN were placed in this site, the bulk of its dimethyl ring substituents could not be accommodated in the cavity available and additionally would not hydrogen-bond with Asp$^{320}$ as seen for F$_{420}$. In contrast, in its own site, FMN projects this same nonpolar dimethyl structure into a larger nonpolar pocket (Fig. 5B). F$_{420}$ would, if placed in the FMN pocket, extend its hydroxyl group into the nonpolar pocket with no hydrogen bond partner available, effectively producing a buried and unsatisfied hydrogen bonding potential.

The binding of both F$_{420}$ and FMN in the FbiB C-terminal domain is an interesting feature but not a unique occurrence. A 2012 study of F$_{420}$-dependent reductases from M. smegmatis demonstrated cofactor promiscuity in three F$_{420}$-dependent reductases, which can utilize both F$_{420}$ and FMN to catalyze different chemistries (oxidation and reduction) of the same substrate (4). The F$_{420}$-dependent reductases catalyze the reduction of aflatoxins and plant-derived furanocoumarins in the presence of F$_{420}$H$_2$. When FMN replaces F$_{420}$ three of the F$_{420}$-dependent reductases tested were found to catalyze FMN-mediated oxidation of two major aflatoxins, AFG1 and AFG2, via dehydrogenation. In another context, an example of competitive FMN/F$_{420}$ binding is provided by FprA, a di-iron flavo-protein F$_{420}$/H$_2$ oxidase found in methanogenic archaea, which catalyzes the four-electron reduction of O$_2$ to 2H$_2$O with 2 molar equivalents of reduced F$_{420}$ (45). A recent systematic study of the mycobacteria oxido-reductase superfamily (flavin/deazaflavin oxido-reductases) shows some degree of promiscuity within single proteins, binding F$_{420}$, FMN, and FAD with micromolar affinity (8). However, whereas the proteins appear promiscuous, they also show selectivity for a single cofactor (~5–10-fold). Similarly, F$_{420}$ binding in the C-terminal (oxido-reductase-derived) domain of FbiB shows a 10-fold selectivity over FMN, and both bind at low micromolar affinities (F$_{420}$ $K_d$ = 1.47 versus FMN $K_d$ = 14.7).

Our crystal structure analysis of FMN and F$_{420}$ binding, illustrated in Fig. 5, C and F, suggests that although these ligands do not occupy the same site, they cannot bind simultaneously because steric clashes would result; competitive binding in the closely adjacent site cannot, however, be precluded, and we have not examined this. Although it is not entirely unexpected that the C-terminal domain binds F$_{420}$ potentially in a feedback regulation role, there is no suggestion of why FMN also binds to this domain; could FMN binding in the FbiB C-terminal domain merely represent an evolutionary relic from an ancestral FMN-dependent nitroreductase domain? In this hypothesis, F$_{420}$ binding evolved in the C-terminal domain after being acquired through a recombination event; the FMN binding site has been retained, although we might have expected it to mutate over an evolutionary time scale to abolish binding unless FMN plays some role in the poly-$\gamma$-glutamyl reaction. A more compelling explanation is that the ancestral nitroreductase appropriated by FbiB through recombination was already a promiscuous enzyme, binding both F$_{420}$ and FMN for oxidoreductase activity potentially toward multiple substrates. Again, this hypothesis does not explain the retention of distinct F$_{420}$ and FMN binding sites in the C-terminal domain of FbiB nor the potential roles these redox molecules play in the poly-$\gamma$-glutamyl reaction.

Reaction Mechanism and the Role of the C-terminal Domain—Our biochemical evidence implies that the catalytic machinery of the FbiB protein is located in the N-terminal domain and functions similarly to the single domain homologs that produce F$_{420}$-2 as the largest product. The homolog structure from A. fulgidus (Protein Data Bank entry 2PHN), equivalent to the FbiB N-terminal domain, displays metal and GDP binding; glutamate and F$_{420}$ binding is not observed, but the locations of their binding sites and the catalytic residues involved in glutamate addition have been postulated (46). In relation to the C-terminal domain of FbiB, our biochemical analysis suggests that this domain only provides support for the catalytic activity of the N-terminal domain; both domains are required for catalysis, and either domain alone does not show catalytic activity. We have produced a composite model of the whole FbiB protein by combining our crystal structure of the C-terminal domain with a homology model of the N-terminal domain and by rigid body refinement of these two half-structures against SAXS data (Fig. 6). The two domains are linked via an $\alpha$-helical segment that we infer is contiguous between the domains; the C terminus of the homolog structure CoE is $\alpha$-helical, as is the N terminus of our crystal structure. Two models of full-length protein in combination show best agreement with the SAXS data. The first model can be referred to as “straight” with the C-terminal domain located directly over the N-terminal domain with both domains together showing 2-fold symmetry; the two domains are separated by ~14 Å and, outside of the linker segment, do not interact. Our second model, best referred to as “bent,” shows the C-terminal domain bending ~50°, toward the N-terminal domain (Fig. 6B). This model brings the putative N-terminal catalytic site and the FMN/F$_{420}$ sites of the C-terminal domain to within ~40 Å of each other, still relatively distant.

Either in a catalytic enhancement or a regulatory role, the C-terminal domain may produce an allosteric signal from one or the other of the ligand binding sites to influence catalysis in the N-terminal domain (although we have not observed any effect). We can also hypothesize that the C-terminal domain and perhaps its dynamic behavior in solution, implied by our two-structure model of the SAXS data, apply some mechanical force through the connecting linker sequence to promote or regulate the N-terminal domain function; this is a hypothesis that we have previously explored for another M. tuberculosis enzyme, 2-isopropylmalate synthase, where its C-terminal domain is required for both feedback regulation and catalytic activity in the separate and distant N-terminal domain (47).
One very interesting feature of our N-terminal homology model is the presence of a positively charged surface groove lined by 11 arginine and lysine residues; this is not a feature of the homologous A. fulgidus structure from which the model was produced. If, as we contend, the N-terminal domain contains all of the catalytic machinery for F420 elongation, then this basic surface groove connected to the postulated active site cavity (Fig. 7), could bind to the growing poly-γ-glutamate tail with its negatively charged carboxylate groups and direct the growing chain toward solvent space or toward the adjacent C-terminal domain.

Elongation Mechanism; Insertion or Extension?—Poly-γ-glutamate tails are found on a limited number of biomolecules, including cofactor F420 and folate derivatives. Poly-γ-glutamic acid is also a polymer that is produced by a number of microorganisms, with roles from virulence to promising potential for medical and industrial applications (48). FbiB (CoE in archaea), folypolyglutamate synthase, and the poly-γ-glutamic acid synthetase complex, catalyze the poly-γ-glutamylation reactions to produce these molecules. The chemical mechanism has been generally assumed to be similar for all of these enzymes, involving activation of the carboxylic acid on the elongated substrate in a nucleotide-dependent manner (GTP in F420 and ATP in folates and poly-γ-glutamic acid), formation of an acyl phosphate intermediate, and finally nucleophilic attack by the incoming l-glutamate (17, 46).

Our functional characterization described in the present work demonstrates the addition of multiple l-glutamate residues to the growing poly-γ-glutamate tail of F420 carried out by FbiB. It is not clear, however, whether each l-glutamate is added to the terminal residue of the growing chain (an extension mechanism) or inserted somewhere into the middle of the chain, perhaps between the phospholactate moiety and the first l-glutamate residue (an insertion mechanism). Elucidation of the full mechanistic details of the elongation mechanism will also have implications for further understanding of the function and mechanism of the folypolyglutamate synthase and poly-γ-glutamic acid synthetase enzymes that also carry out poly-γ-glutamylation reactions.

Author Contributions—G. B., E. N. B., and C. J. S. designed the study. G. B. and A. M. R. performed the majority of experiments, including protein production and characterization. A. M. R. carried out protein crystallization. G. B. and S. S. performed the functional experiments. Final structure refinement and deposition was carried out by H. M. B. and C. J. S. G. B., E. N. B., and C. J. S. wrote the paper, and all authors analyzed the results and approved the final version of the manuscript.

Acknowledgments—We thank Ehab Jirgis for technical assistance and Martin Middleditch for mass spectrometry, performed at the Centre for Genomics, Proteomics, and Metabolomics at the University of Auckland. This research was undertaken on the MX1, MX2, and SAXS/WAXS beamlines at the Australian Synchrotron (Victoria, Australia). We thank Drs. Neil Patterson, Jeremy Keown, David Goldstone, and Shaun Lott for synchrotron data collection.

References

**M. tuberculosis γ-Glutamyl Ligase FbiB**


6894 JOURNAL OF BIOLOGICAL CHEMISTRY VOLUME 291 • NUMBER 13 • MARCH 25, 2016