Crystal Structures of $F_{420}$-dependent Glucose-6-phosphate Dehydrogenase FGD1 Involved in the Activation of the Anti-tuberculosis Drug Candidate PA-824 Reveal the Basis of Coenzyme and Substrate Binding*

Received for publication, March 7, 2008, and in revised form, April 8, 2008. Published, JBC Papers in Press, April 22, 2008. DOI 10.1074/jbc.M801854200

Ghader Bashiri¹, Christopher J. Squire, Nicole J. Moreland, and Edward N. Baker²

From the School of Biological Sciences and Centre of Molecular Biodiscovery, University of Auckland, Private Bag 92019, Auckland, New Zealand

The modified flavin coenzyme $F_{420}$ is found in a restricted number of microorganisms. It is widely distributed in mycobacteria, however, where it is important in energy metabolism, and in Mycobacterium tuberculosis (Mtb) it is implicated in redox processes related to non-replicating persistence. In Mtb, the $F_{420}$-dependent glucose-6-phosphate dehydrogenase FGD1 provides reduced $F_{420}$ for the in vivo activation of the nitroimidazopyran prodrug PA-824, currently being developed for anti-tuberculosis therapy against both replicating and persistent bacteria. The prodrug PA-824, currently being developed for anti-tuberculosis therapy, is a promising anti-TB compound currently undergoing clinical trial with the Global Alliance for TB Drug Development. The compound exhibits potent bactericidal activity against both growing and dormant Mtb and requires reductive, bacterial activation that has been shown to occur via a $F_{420}$-dependent mechanism. A comparison of Mtb strains susceptible or resistant to PA-824 revealed that mutations in the open reading frame Rv0407, encoding FGD1 ($F_{420}$-dependent glucose-6-phosphate dehydrogenase 1), result in PA-824 resistance. Recent studies have shown that FGD1 does not interact directly with the compound but, rather, provides reduced $F_{420}$ to an accessory protein (Rv3547), which in turn activates PA-824 (3) (Fig. 1).

The unusual coenzyme $F_{420}$ (a 7,8-dimethyl-8-hydroxy-5-deazariboflavin electron transfer agent) consists of an isoalloxazine chromophore with a side chain comprising ribitol, phosphate, and lactate residues and a poly-glutamate tail of variable length (Fig. 1). Originally isolated from methanogenic Archaea (4), $F_{420}$ was discovered in Mtb in the mid-1980s (5) and was subsequently found to be widely distributed in Mycobacterium species (6). $F_{420}$ is used in a variety of roles in different organisms; that is, energy generation in Archaea, antibiotic biosynthesis pathways in Streptomyces species, and DNA photoreactivation reactions in Scenedesmus and Synechocystis species (7). In Mycobacterium and Nocardia species, it is used by $F_{420}$-dependent glucose-6-phosphate dehydrogenases (FGDs) (6, 8), two of which (FGD1 and FGD2) are present in Mtb. Aside from the absolute requirement for $F_{420}$ and FGD1 in the activation of PA-824, the role of $F_{420}$ in Mtb is not well understood. However, the redox potential of $F_{420}$ (~380 mV) is lower than that of the classical hydrogen carrier NAD(P)$^+$ (~320 mV), and it has been hypothesized that reduced $F_{420}$ in Mtb may have an important role in low-redox potential reactions that would be associated with anaerobic survival and persistence (9).

Reduction of the $F_{420}$ coenzyme is achieved by hydride transfer from a substrate molecule to C5 of the 5’-deazaflavin (10). There is relatively little structural information, however, on the factors that facilitate this reaction. Although crystal structures are available for a small number of $F_{420}$-dependent enzymes, only for three proteins has the mode of $F_{420}$ binding been

---

Mycobacterium tuberculosis (Mtb),¹ the causative agent of tuberculosis (TB), is a devastating pathogen. The World Health Organization estimates that 1.6 million people die of the disease each year and one-third of the world population is currently infected with the TB bacillus. Current TB chemotherapy comprises a 6-month-long regimen, and no new drug has been introduced into this regimen for more than 30 years (1). PA-824, a nitroimidazopyran, is a promising anti-TB compound currently undergoing clinical trial with the Global Alliance for TB Drug Development. The compound exhibits potent bactericidal activity against both growing and dormant Mtb and requires reductive, bacterial activation that has been shown to occur via a $F_{420}$-dependent mechanism (2). A comparison of Mtb strains susceptible or resistant to PA-824 revealed that mutations in the open reading frame Rv0407, encoding FGD1 ($F_{420}$-dependent glucose-6-phosphate dehydrogenase 1), result in PA-824 resistance (2). Recent studies have shown that FGD1 does not interact directly with the compound but, rather, provides reduced $F_{420}$ to an accessory protein (Rv3547), which in turn activates PA-824 (3) (Fig. 1).

The unusual coenzyme $F_{420}$ (a 7,8-dimethyl-8-hydroxy-5-deazariboflavin electron transfer agent) consists of an isoalloxazine chromophore with a side chain comprising ribitol, phosphate, and lactate residues and a poly-glutamate tail of variable length (Fig. 1). Originally isolated from methanogenic Archaea (4), $F_{420}$ was discovered in Mtb in the mid-1980s (5) and was subsequently found to be widely distributed in Mycobacterium species (6). $F_{420}$ is used in a variety of roles in different organisms; that is, energy generation in Archaea, antibiotic biosynthesis pathways in Streptomyces species, and DNA photoreactivation reactions in Scenedesmus and Synechocystis species (7). In Mycobacterium and Nocardia species, it is used by $F_{420}$-dependent glucose-6-phosphate dehydrogenases (FGDs) (6, 8), two of which (FGD1 and FGD2) are present in Mtb. Aside from the absolute requirement for $F_{420}$ and FGD1 in the activation of PA-824, the role of $F_{420}$ in Mtb is not well understood. However, the redox potential of $F_{420}$ (~380 mV) is lower than that of the classical hydrogen carrier NAD(P)$^+$ (~320 mV), and it has been hypothesized that reduced $F_{420}$ in Mtb may have an important role in low-redox potential reactions that would be associated with anaerobic survival and persistence (9).

Reduction of the $F_{420}$ coenzyme is achieved by hydride transfer from a substrate molecule to C5 of the 5’-deazaflavin (10). There is relatively little structural information, however, on the factors that facilitate this reaction. Although crystal structures are available for a small number of $F_{420}$-dependent enzymes, only for three proteins has the mode of $F_{420}$ binding been

---

¹ This work was supported by the New Economy Research Fund of New Zealand, the Health Research Council of New Zealand, and Centres of Research Excellence funding to the Maurice Wilkins Centre of Molecular Biodiscovery. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

² To whom correspondence should be addressed: School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland, New Zealand.

¹ Recipient of a Doctoral Scholarship from the Iranian Ministry of Science, Research, and Technology.

² This atomic coordinates and structure factors (codes 3C8N and 3B4Y) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

³ The abbreviations used are: Mtb, Mycobacterium tuberculosis; TB, tuberculosis; FGD1, $F_{420}$-dependent glucose-6-phosphate dehydrogenase 1; G6P, glucose 6-phosphate; β-ME, β-mercaptoethanol; MALDI-TOF, matrix-assisted laser desorption-ionization reflection time-of-flight; SeMet, selenomethionine.
defined; they are a secondary alcohol dehydrogenase (Adf), a methylene-tetrahydromethanopterin reductase (Mer), and an F420H2:NADP+/H11001 oxidoreductase (Fno), all from methanogenic Archaea (11–13). Adf and Mer belong to the bacterial luciferase family of proteins (11), and sequence similarities between the archaeal Mer enzymes and mycobacterial FGD (23–28% identity) suggested that the latter might also belong to this family (14). Adf and Mer share a common mode of F420 binding (12) that might be expected also in FGD1. The substrates in these proteins vary widely, however, and the sections of polypeptide that enclose the active sites appear to vary significantly. Catalysis depends critically on the relationship between substrate and coenzyme binding, in distance, orientation, and conformation. Given its role in the development of resistance to PA-824 and its likely importance in mycobacterial physiology, we sought to determine the structural determinants of FGD1 function.

Here we describe the three-dimensional structure of Mtb FGD1 determined by x-ray crystallography at 1.9 Å resolution both with and without bound F420 and in the presence of a bound citrate ion, which is shown to be a competitive inhibitor of substrate binding. These structures together with kinetic and spectroscopic analyses and in silico modeling of substrate binding allow the development of a hypothesis for hydride transfer between the substrate and cofactor, the occurrence of which is crucial for PA-824 activation in Mtb.

MATERIALS AND METHODS

Cloning, Expression, Purification, and Crystallization of ApoFGD1—Native and selenomethionine-substituted FGD1 overexpressed in the Mycobacterium smegmatis strain mc24517 using the expression vector pYUB1049 and were purified by immobilized metal ion affinity chromatography and size exclusion chromatography (15). The N-terminal polyhistidine tag, added by the expression vector and used for affinity purification, was not cleaved and remained on the protein in subsequent crystallization and functional experiments. Selenomethionine-substituted FGD1 crystals were grown as described (15) in a solution containing 1.4 M trisodium citrate, pH 6.5, and 0.1% dioxane. The crystals are monoclinic, space group P21, with unit cell dimensions a = 80.95, b = 89.13, c = 90.86 Å, and β = 91.64°, and the asymmetric unit contains four FGD1 protomers.

Data Collection and Structure Determination of ApoFGD1—For data collection, the crystals were flash-frozen in liquid nitrogen after being placed in a cryoprotectant comprising 70% N-paratone, 30% mineral oil. Se-MAD data sets were collected at three wavelengths (0.97929, 0.91162, and 0.97941 Å) at the Stanford Synchrotron Radiation Laboratory (beamline 9–2 with a MarMosaic-325 CCD detector). A further dataset was collected from a selenomethionine-substituted crystal using a Rigaku rotating copper anode (λ = 1.5418 Å) and a MarResearch dtb345 image plate detector. Data were processed with HKL2000 (16) or MOSFLM and SCALA from the CCP4 program suite (17). Data collection and processing statistics are shown in Table 1.

The structure of FGD1 was solved by multiwavelength anomalous dispersion methods. Selenium sites were found using SHELXD (18), which located 28 of the expected 36 sites. Selenium position refinement and phase calculation were performed by SHARP (19), and density modification was performed by SOLOMON (20). The majority of the model was built by PHENIX (21), with the remainder built manually using COOT (22). Water molecules were identified by their spherical electron density (height greater than 3σ in Fo – Fc maps or 1σ in 2Fo – Fc maps) and appropriate hydrogen bond geometry with surrounding structure. Refinement was carried out using REFMAC5 (23, 24); statistics are shown in Table 1. The final model, refined at 1.9 Å resolution with R = 0.212 and Rfree = 0.255, contains residues 3–334 (of the expected 336) for three monomers and residues 1–334 plus 12 residues of the N-terminal His
tag for the fourth monomer (A). The Ramachandran plot produced by PROCHECK (25) shows that 91.3% of all residues fall in the preferred regions, and there are no outliers. In the analysis of the structure, hydrogen bonds were inferred where bond distances of 2.5–3.4 Å and bond angles greater than 90° at the donor H or acceptor O/N atoms were found.

Crystallographic Data Collection and Structure Determination of FGD1 Complexes—The complex structure with F420 was obtained by soaking native FGD1 crystals with F420 purified from M. smegmatis and a substrate analog. The substrate analog, 1,5-anhydro-D-glucitol 6-phosphate, was synthesized by phosphorylating 1,5-anhydro-D-glucitol (Toronto Research Chemicals) using hexokinase (Sigma) as previously reported (26). FGD1 crystals were transferred into sitting drops containing 1,2M tri-sodium citrate, pH 6.5, 1mM F420, and 50mM sodium citrate ions (Table 1).

To determine the optimal pH for F420 binding, FGD1 (5 μM) and F420 (10 μM) were incubated in buffer solutions with a pH range of 4–9 (0.5 steps) for 1 h at room temperature before fluorescence analysis. Reactions contained 50 mM buffer, 300 mM NaCl, 1 mM β-ME, and 1 mM EDTA and were corrected against a control lacking FGD1. The optimal pH for F420 binding was found to be 6.5–7.0, and pH 7.0 was used in subsequent fluorescence experiments. To determine the dissociation constant for F420 binding to FGD1, F420 (0–25 μM) and FGD1 (100 μM) were mixed in 50 mM Tris-HCl, pH 7.0, 300 mM NaCl, 1 mM β-ME, and 1 mM EDTA and incubated as above. Ligand binding data were fitted with SigmaPlot v10 (Systat Software Inc.) using a one-site saturation model, \( \Delta F = (\Delta F_{max}[\text{lignand}])/(K_D + [\text{lignand}]), \) where \( \Delta F \) is the normalized change in fluorescence compared with a solution of 100 μM protein alone, \( \Delta F_{max} \) is the maximum normalized change in fluorescence at saturation, \( K_D \) is the dissociation constant, and [lignand] is the concentration of ligand.

| TABLE 1 | Data collection and refinement statistics |
|---------|-----------------------------------------|
| Data collection | SeMet peak | SeMet remote | SeMet inflection | SeMet (apo) | F420-citrate complex |
| Wavelength (Å) | 0.97929 | 0.91162 | 0.97941 | 1.5418 | 1.5418 |
| Resolution (Å) | 2.4 (2.53–2.4) | 2.5 (2.64–2.5) | 2.7 (2.85–2.7) | 1.9 (1.95–1.9) | 1.95 (2.06–1.95) |
| Total reflections | 359,388 | 162,314 | 129,193 | 354,015 | 457,092 |
| Unique reflections | 49,000 | 43,839 | 35,007 | 101,311 | 47,109 |
| Mean I/σ(I) | 19.8 (3.2) | 13.8 (1.7) | 14.3 (1.9) | 13.5 (1.7) | 17.6 (1.2) |
| Rmerge | 0.088 (0.517) | 0.072 (0.560) | 0.071 (0.500) | 0.062 (0.552) | 0.114 (0.484) |
| Completeness (%) | 100 (100) | 100 (100) | 100 (100) | 99.6 (98.0) | 96.7 (92.6) |
| Multiplicity | 7.3 (7.0) | 3.7 (3.7) | 3.7 (3.7) | 3.5 (3.2) | 9.8 (2.0) |
| Wilson B factor (Å²) | 29.9 | 23.8 |
| Refinement | ApoFGD1 | F420-citrate complex |
| Resolution range (Å) | 40.1–1.90 | 24.7–1.95 |
| Number of reflections (working/test) | 96,205/5,077 | 44,527/2,368 |
| R factor/| 0.211/0.255 | 0.170/0.228 |
| Rfree (%) | 10.378 | 5.165 |
| Number of atoms (non-hydrogen) | 0.016 | 0.014 |
| Protein | 10.378 | 5.165 |
| Ligand | 0.016 | 0.014 |
| Solvent | 1.451 | 1.646 |
| Root mean square deviations from ideality | 371 | 196 |
| Bonds (Å) | 25.9 | 22.9 |
| Angles (degree) | 0.016 | 0.014 |
| Average B factors (Å²) | 1.451 | 1.646 |
| Protein atoms | 25.9 | 22.9 |
| Ligand | 27.7 | 24.2 |
| Water molecules | 27.7 | 24.2 |
| Residues in most favored region (%) | 91.3 | 93.6 |

* Values in parentheses are for the outermost resolution shells.
Absorbance-based Assays—FGD1 activity was monitored by UV-visible spectroscopy on a SpectraMax microplate spectrophotometer (Molecular Devices). The decrease of F₄₂₀ absorbance at 420 nm was monitored over 10 min in all activity assays. All reactions contained 100 μl of 50 mM Tris-HCl, pH 7.0, 300 mM NaCl, 1 mM β-ME, and 1 mM EDTA and were performed in 96-well format (Greiner bio-one, Germany). Kinetic studies with the substrate glucose 6-phosphate (G6P) were performed using 100 nM FGD1, 20 μM F₄₂₀, and varying concentrations of G6P (0.01–1 mM). The Kₘ was determined using the same one-site saturation model used for Kᵥ, where rate replaces ΔF, and Kₘ replaces Kᵥ. The inhibitory effect of citrate on FGD1 activity was assayed using 100 nM FGD1, 20 μM F₄₂₀, 100 μM G6P, and varying concentrations of citrate (100 μM–20 mM). The experimental data were fitted, and IC₅₀ values were calculated using SigmaPlot v10 and a one-site competition model.

In Silico Docking—Both α- and β-anomers of G6P were constructed and energy-minimized in the molecular modeling package SYBYL7.3 (Tripos Inc., St. Louis, MO). Both α- and β-D-G6P were docked without constraints into the FGD1 active site using GOLD (31). The top 10 solutions were fitted, and IC₅₀ values were calculated using SYBYL7.3.

RESULTS

FGD1 Expression, Crystallization, and Cofactor Preparation—M. tuberculosis FGD1 was overexpressed in M. smegmatis and crystallized in citrate buffer supplemented with dioxane. Subsequent preparations in minimal media supplemented with selenomethionine gave crystals of the SeMet-substituted protein that were used for structure determination by multiwavelength anomalous diffraction methods. The F₄₂₀ used in functional assays and crystallographic binding experiments was purified from M. smegmatis cell extracts, with a typical yield of 100 μg of F₄₂₀ per liter of culture. Analysis by MALDI-TOF mass spectrometry shows that the F₄₂₀ isolated from M. smegmatis contains up to nine glutamate residues, although the predominant species has six (Fig. 2), in line with previously published analyses which indicated five or six glutamates (32).

FGD1 Overall Structure—The crystal structure of apoFGD1, determined at 1.90 Å resolution, comprises four molecules in the asymmetric unit arranged as two independent homodimers. The liganded structure, containing F₄₂₀ and citrate and refined at 1.95 Å resolution, contains two molecules in each asymmetric unit which form the same biological dimer. Superposition of the monomer structures shows little difference between the apo and liganded forms, with an overall root mean square difference in Cα atomic positions of 0.78 Å over all 332 residues. The largest local differences are in solvated loop regions and areas involved in crystal packing.

Each monomer of FGD1, comprising residues 3–334, forms an (α/β)₃ TIM-barrel, in which the active site is located as usual at the C-terminal end of each barrel (Fig. 3A). One molecule each of F₄₂₀ and citrate bind deeply in the active site of the liganded structure (Fig. 3B), occupying a cavity measuring ~950 Å³, as calculated by the Computed Atlas of Surface Topography of Proteins, CASTp (33). The two monomers that comprise the FGD1 dimer are related by an approximate 2-fold axis parallel with the central β-strands and associate by the interaction of three α-helices, α₁, α₂, and α₃, and associated loops. The dimer interface, as analyzed by the program PISA (34), buries 2022 Å² or 15% of each monomer surface and contains 29 hydrogen bonds and 22 salt bridges.

A search of the Protein Data Bank with SSM (35) shows that FGD1 most closely matches proteins of the bacterial luciferase family (11, 36), as expected from its 20–30% sequence identity with other family members. This family includes FMN- and F₄₂₀-dependent oxidoreductases which act on a diverse range of substrates. The SSM analysis shows that the F₄₂₀-dependent enzymes form a distinct subgroup within the family, comprising FGD1, Adf, and Mer. The closest structural homologue to FGD1 is Adf (11), which shares not only the same monomer fold but also the same dimer association; the root mean square difference is 1.68 Å for 639 Cα atoms of the dimer. As in the other bacterial luciferase family members, three insertion segments (IS1, IS2, and IS3) “cap” the C-terminal end of the barrel, folding over the active site to form the substrate binding pocket. Multiple sequence alignments (Fig. 4) highlight sequence vari-
Structure and Function of $F_{420}$-dependent G6P Dehydrogenase

A

FIGURE 3. Structure of M. tuberculosis FGD1. A, ribbon diagram of the FGD1 monomer, shown as a stereo-view. The (α/β)$_8$ barrel is shown in blue (α-helices) and yellow (β-strands). Insertion sequences (ISs) in the core barrel form structures that cap the active site and are shown in red (IS1), green (IS2), and cyan (IS3). Secondary structures are labeled. Bound $F_{420}$ and citrate are shown in pink and red, respectively, in ball and stick representation. B, stereo view of the FGD1 active site. Citrate and $F_{420}$ are shown in their $2F_o-F_c$ omit density, contoured at 1.0 σ, and drawn as stick models with carbon atoms colored yellow and pink, respectively. Protein is drawn as a ribbon model, colored blue, with selected side chains drawn as line models. Water molecules are shown as red spheres, and hydrogen bond contacts are shown as broken lines. This and other structural figures were drawn with Pymol (DeLano Scientific, Palo Alto, CA).

ations in these insertion segments and structural differences in the capping region correlate with different substrate specificity. This is most apparent in FGD1 for the phosphate binding pocket described later. A large deviation between FGD1 and Adf is found in the loop region Thr$^{195}$–Glu$^{201}$, which includes two key residues that contact citrate in our liganded structure or the G6P phosphate in our modeled structure. In the area around the $F_{420}$ binding site, particularly the isoalloxazine system, differences are small and indicate a strong conservation of structure associated with $F_{420}$ binding and in the barrel core.

Active Site-$F_{420}$ Binding—The $F_{420}$ cofactor binds in the active site, with its three-ring isoalloxazine system innermost and its poly-glutamate tail extending into the solvent. As in Adf (11) and Mer (12), the isoalloxazine ring binds adjacent to a bulge in strand β3 containing a non-prolyl cis-peptide bond that joins Ser$^{73}$ and Val$^{74}$ in FGD1. Analogous non-prolyl cis-peptide bonds are found in some but not all other proteins of the bacterial luciferase family (11). Recognition of the isoalloxazine ring depends on shape through non-bonded contacts described below together with a number of specific hydrogen bonds that primarily involve the pyrimidine ring (Fig. 5A). The latter makes highly favorable hydrogen bonds with the main chain C=O and NH groups of Asp$^{39}$, the side chains of Thr$^{76}$ and Asn$^{112}$, and two buried water molecules. The ribitol and phosphate groups make a few hydrogen bonds with protein atoms, primarily from the polypeptide backbone, but are less tightly enclosed, and the poly-glutamate tail exits the active site cavity at the first glutamate, with no further residues being visible.

A striking feature of $F_{420}$ binding in FGD1 is the pronounced butterfly conformation of the isoalloxazine ring system (Fig. 5C), which is caused by steric contacts from the surrounding structure. As in Adf and Mer, the non-prolyl cis-peptide $β$-bulge packs against the chromophore from behind (the $Re$-face), at the center of the ring system; contacts are made by the carbonyl oxygen of Ser$^{73}$ and the side chain of Val$^{74}$. Additional interactions with the other face (the $Si$-face) at its two extremities then cause bending about an axis between the C5 and N3 atoms. At one end of the isoalloxazine system, the pyrimidine ring occupies a tight pocket formed by His$^{40}$, Glu$^{109}$, the protein backbone, and two ordered waters. At the other end the primary contact of the hydroxybenzyl ring is with the citrate ion, which occupies the substrate binding pocket, as described below. The C5 atom of the central pyridine ring faces into this substrate binding pocket and to proposed catalytic residues (see later), consistent with experimental data showing that the catalytic mechanism has $Si$-face stereospecificity at carbon C5 (37). Overall, the butterfly bend of 162° is similar to that seen for $F_{420}$ bound to Adf but differs in detail, with a smaller angle between pyrimidine and pyridine rings and a larger angle between pyridine and hydroxybenzyl, the inverse of what is seen for Adf. The major difference is that no hydrophobic residues restrain the hydroxybenzyl ring, as occurs in Adf; instead, this ring is restrained in FGD1 by steric interactions with the citrate ion.

Active Site-Citrate Binding—A bound citrate ion, presumably derived from the crystallization medium, is located immediately adjacent to the $F_{420}$ molecule, occupying a cavity measuring ~460 Å$^3$, much larger than the equivalent pocket in Adf (128 Å$^3$), which binds the smaller isopropanol substrate. The cavity is an appropriate size to fit a G6P molecule, which has a calculated molecular volume of ~200 Å$^3$. The C5 atom of $F_{420}$ is clearly presented to this substrate cavity for catalysis. Unlike $F_{420}$, the citrate ion hydrogen bonds exclusively to protein side chain atoms or water molecules, with no backbone interactions.
The optimum pH for F420 binding is between pH 6.5 and 7 (Fig. 5A).

His40, both putative catalytic residues, as discussed later. Five of six hydrogen bonds made by the y-carboxylate of citrate are to lysine or arginine side chains: Lys198, Lys259, and Arg44. The location of these positively charged side chains at the peripheral rings (180° being planar) and is 162° for FGD1 and 160° for Adf. The hydroxybenzyl ring of F420 is constrained by hydrophobic interactions in Adf but by steric interactions with citrate in FGD1.

The presence of a citrate ion in the substrate binding pocket suggested that citrate might be an inhibitor of FGD1 catalysis. FGD1 activity was, therefore, assayed while keeping protein, F420, and G6P concentrations constant but varying the citrate concentration and monitoring F420 reduction as above. This showed that citrate is indeed a competitive inhibitor of the enzyme with an IC50 of 43 μM (Fig. 6D).

Substrate Modeling—A substrate molecule was modeled into the active site of the protein by docking an idealized D-G6P molecule using the program GOLD (31). Both the α- and β-anomers were docked without introducing any constraints. In docking the α-sugar we observe two out of ten docking solutions with an appropriate placement of the anomeric C1 atom for hydride transfer to C5 of F420. In the case of the β-sugar, five out of ten docking solutions are in an appropriate location and conformation for hydride transfer. These solutions agree very closely (root mean square difference in atomic positions 0.56 Å), providing validation of the modeling result. Three water molecules from the experimental structure and function of F420-dependent G6P dehydrogenase.

In Vitro Binding of F420, Glucose 6-Phosphate, and Citrate—The optimum pH for F420 binding is between pH 6.5 and 7 (Fig. 6A), as measured by the fluorescence assay. In contrast to M. smegmatis FGD, which shows two maxima, at pH 5.5 and 8 (6) only one pH optimum is seen for M. tuberculosis FGD1. A pH of 7.0 was, therefore, chosen to quantify F420 binding and determine its dissociation constant (Kd). By varying the F420 concentration while keeping the protein concentration constant and measuring fluorescence emission, we determined the Kd for F420 binding to FGD1 to be 4.5 μM (Fig. 6B). The same value was obtained for two different protein concentrations, 100 and 5 μM, and is very similar to the value for M. smegmatis FGD, which was determined as 4 μM (6).

FGD1 activity was assayed by monitoring the decrease in F420 absorbance at 420 nm upon reduction. The protein and F420 concentrations were kept constant, whereas substrate concentration was varied, giving a calculated Kd value for G6P of 100 μM. The reaction kinetics follows a Michaelis-Menten model (Fig. 6C). The equivalent Kd for G6P for M. smegmatis FGD is higher, at 1.6 mm, which is possibly correlated with the high abundance of G6P in this organism in vivo (6).
mental citrate-bound structure are replaced by sugar hydroxyl groups in this model.

Two key features of the substrate binding model are that the sugar ring is sandwiched between F420 and Leu252, stacking against the F420 isoalloxazine ring system and that the phosphate moiety occupies the positively charged pocket formed by Lys198, Lys259, and Arg283 (Fig. 7). In all docking experiments, this phosphate group localization was strongly conserved. The sugar ring position was slightly less restricted, but all five top solutions share very similar hydrogen bonding patterns of the sugar hydroxyl groups, and in the top solution 13 hydrogen bonds link the sugar to protein side chains (Fig. 7).

**DISCUSSION**

The cofactor F420 is a relatively rare flavin variant found in a restricted range of microorganisms, among them mycobacteria. It is not made by *Escherichia coli*, and it was this that prompted the change to *M. smegmatis* expression of FGD1 after failure to achieve expression of soluble protein from *E. coli* (15). In fact, soluble, expressed FGD1 obtained from *M. smegmatis* did not contain bound F420, which had to be independently extracted from bacterial cultures and soaked into FGD1 crystals. This is consistent with our measured $K_D$ of 4.5 $\mu$M, which implies only a moderate affinity of F420 for FGD1, consistent with its role in making reduced F420 available for other redox processes *in vivo*, including the activation of the pro-drug PA-824; it was reduced F420, released from FGD1, that was utilized in the reduction of PA-824 (3).

Few structures are available for F420-dependent enzymes, and although F420 complexes of Adf and Mer have been defined structurally, there is little direct information on the relationship between substrate and F420 in these enzymes (11). The fortuitous discovery that citrate, used in the crystallization of FGD1, is a competitive inhibitor enables us to model the binding mode for the G6P substrate with a high degree of confidence, identify likely catalytic residues, and propose a catalytic mechanism for FGD1.

**Substrate and Inhibitor Binding to FGD1**—Attempts were made to visualize substrate binding directly in the presence of F420 by synthesizing a substrate analog (1,5-anhydro-D-glucitol 6-phosphate) and soaking it into pre-formed crystals. Crystallographic analysis of these soaked crystals showed no density...
for the substrate analog (data not shown), however, implying that it is unable to displace the citrate molecule from the active site. This suggested that citrate might be a competitive inhibitor of FGD1. The measured IC50 of 43 \( \mu \text{M} \) for citrate confirms this and is lower than the \( K_m \) for G6P (100 \( \mu \text{M} \)), explaining why repeated attempts to displace citrate with substrate or substrate analogs were unsuccessful.

The positively charged pocket formed by Lys198, Lys259, and Arg283 appears critical to the binding of both citrate and the G6P substrate. These residues stabilize the negative charge on the G6P phosphate group, providing five hydrogen bond interactions with the phosphate oxygens, all of highly favorable geometry. In the citrate structure two carboxylate groups bind on either side of the phosphate location. This shared binding mode between the citrate ion and the phosphate group of G6P is consistent with observations that citrate can act as a competitive inhibitor of phosphate in phosphate-binding proteins (39).

The modeled G6P substrate is positioned such that it makes key contacts both with protein residues and with F420 in the region of its anomic carbon C1, where the catalytic reaction occurs (Fig. 7). This strongly supports the validity of the modeling and leads to a proposed reaction mechanism, described below. In particular, the C1 hydrogen is only 2.1 Å from C5 of F420, the recipient of the hydride ion transferred in the redox reaction, and the O1 hydroxyl is in hydrogen-bond distance to NE1 of Trp44, NE2 of His40, and OE2 of Glu13. Trp44 can only act as a hydrogen bond donor, whereas His40 is expected to act as a proton acceptor through NE2, as its ND1 atom is hydrogen-bonded to the main chain C=O of Gln42 and is, thus, protonated. In contrast, Glu13 could act as a proton donor or acceptor, depending on its protonation state; theoretical calculation of its pK\(_a\) with the PROPKA server (38) gives a value of \( \sim 6 \).

**Reaction Mechanism**—We propose a mechanism for FGD1 catalysis (Fig. 7D) based on these crystallographic and modeling results and theoretical pK\(_a\) calculations. The mechanism is similar to that proposed for Adf (11) where a comparable arrangement of Glu, His, and Trp residues occurs. The hydrogen-bonding pattern around the O1 hydroxyl of G6P suggests His40 as the catalytic base in the initial hydrogen abstraction from O1 of G6P. The negatively charged reaction intermediate would then be stabilized by the hydrogen bond from Trp44 and possibly also Glu13, depending on its protonation state.

The reaction continues with the hydride transfer step from C1 of G6P to C5 of F420. The hydrogen on C1 extends toward C5 of F420 at an appropriate distance (2.1 Å) and geometry for facile hydride transfer (Fig. 7C). The bending of the F420 chromophore by the bulge formed by the cis-peptide between residues Ser173 and Val174 helps make the C5 atom more reactive as it assumes a more sp\(^3\)-like geometry and electron arrangement. In FGD1, the bound substrate evidently also plays a significant role in promoting the butterfly bend of F420 as shown by the experimentally determined citrate binding mode and the modeled G6P. Unlike Adf, in which hydrophobic residues pack against the hydroxybenzyl ring of F420, in FGD1 it is the sub-

---

**FIGURE 7. Modeled substrate binding mode and proposed catalytic mechanism.**

A, stereo view showing G6P modeled into the active site of FGD1. G6P (stick model with carbon atoms colored yellow) is sandwiched between Trp44/Leu252 and F420. F420 is shown as a pink stick model. Probable hydrogen bond interactions with G6P are shown with broken lines. B, schematic representation of hydrogen bond contacts with G6P. C, interaction of G6P with the F420 isalloxazine system. The G6P-C1 hydrogen, to be transferred to F420, colored black, is 2.1 Å from F420-C5. D, likely reaction mechanism for FGD1. Proposed catalytic residues are shown with protonation states as calculated by PROPKA (38).
strate (or citrate inhibitor) that provides the main packing against this ring. This may also account for the more pronounced pyridine/hydroxybenzyl bend seen in FGD1 compared with Adf. Substrate binding has been shown to assist catalysis by promoting bending of the oxidized isoalloxazine moiety in another flavin-dependent enzyme, cholesterol oxidase (40).

Finally, the electronic rearrangements in the isoalloxazine system that accompany reduction also require protonation at N2 in the pyrimidine ring. This atom is 3.4 Å from OE1 of Glu109. Both the citrate structure and the theoretical calculations of pK_a strongly predict that this residue is protonated. Glu109 is in a highly hydrophobic environment, and the pK_a calculations predict Glu109 as protonated whether the structure submitted to PROPKA is apo, citrate-bound, or a model containing G6P or the phosphogluconolactone product. This suggests that Glu109 plays an important role in assisting F_420 reduction in FGD1.

**Role of FGD1 and F_420 in Mtb**—FGD1 first attracted attention because of its role in providing reduced F_420 for the subsequent reduction and activation of the prodrug PA-824, which is currently being developed as a new-generation anti-TB drug. This is clearly not the true physiological function of FGD1 in Mtb but points to an important role for F_420 in mycobacterial physiology. PA-824 has potent activity not only against replicating bacteria but also against persistent and multiply drug-resistant Mtb (41), implying that F_420, which is essential for PA-824 activation, is relevant to each of these states.

Whole-genome transposon mutagenesis screens (42) and other genetic studies (3) have failed to demonstrate essentiality of FGD1 in replicating Mtb or of any known F_420-dependent gene products or F_420 biosynthetic enzymes. On the other hand, there is growing evidence that F_420 plays an important part in defense against the host immune system and in non-replicating persistence, and a double knock-out of the fgd1 and fgd2 genes would be necessary to properly test their essentiality. Mtb mutants deficient in fbdC, encoding an F_420 biosynthetic enzyme, show hypersusceptibility to nitric oxide and other reactive nitrogen intermediates (43), known to be important in the host immune response. Transition to the persistent state involves major changes in energy metabolism (44), and it has been suggested that F_420, with its low redox potential, may have a role in reactions associated with anaerobic survival (9). Indeed, three putative nitroreductases, Rv2032, Rv3127, and Rv3131, which we have shown to be flavin-dependent, are among the cohort of genes that are up-regulated by hypoxia, modeling the entry to dormancy (45).

What might the specific role of FGD1 be? Mtb is unusual in having fewer than four open reading frames annotated as encoding G6P dehydrogenase enzymes, two of which require NADPH as an electron donor and two of which (FGD1 and FGD2) are F_420-dependent (9). The two F_420-dependent enzymes share 36% sequence identity over 363 residues, although FGD2 has an additional N-terminal extension (Fig. 4). The F_420 binding residues identified in FGD1 are also present in FGD2, but there are differences in the substrate binding site, including the loss of the phosphate binding residues Lys_198, Lys_259, and Arg_283. This suggests that whereas both FGD1 and FGD2 could reduce F_420, they may use different substrates.

The reaction catalyzed by FGD1 is equivalent to the first step in the pentose phosphate pathway, which normally provides NADPH for reductive biosynthetic reactions and maintenance of the cellular reductant state. The maintenance of both F_420 and NADPH-dependent G6P dehydrogenases in Mtb may be essential for providing the redox flexibility that enables the bacterium to switch to the persistent state following phagocytosis.

** Acknowledgments**—We thank Tom Caradoc-Davies, Richard Buneker, and Haojoo Kang for help with data collection and processing. Clark Ehlers and Brian Palmer for help with F_420 extraction and purification, and Christina Buchanan for help with mass spectrometry.

**REFERENCES**

1. Duncan, K., and Barry, C. E., III (2004) *Curr. Opin. Microbiol.* 7, 460–465
2. Stover, C. K., Wrennere, P., VanDevanter, D. R., Sherman, D. R., Arain, T. M., Langhorne, M. H., Anderson, S. W., Towell, J. A., Yuan, Y., McMurray, D. N., Kreiswirth, B. N., Barry, C. E., and Baker, W. R. (2000) *Nature* 405, 962–966
3. Manjunatha, U. H., Boshoff, H., Dowd, C. S., Zhang, L., Albert, T. J., Norton, T. J., Daniels, L., Dick, T., Pang, S. S., and Barry, C. E., III (2003) *Proc. Natl. Acad. Sci. U. S. A.* 100, 43–46
4. Cheeseman, P., Toms-Wood, A., and Wolfe, R. S. (1972) *J. Bacteriol.* 112, 527–531
5. Daniels, L., Bakhiet, N., and Harmon, K. (1985) *Syst. Appl. Microbiol.* 6, 12–17
6. Purwantini, E., and Daniels, L. (1997) *J. Bacteriol.* 178, 2861–2866
7. Isabelle, D., Simpson, D. R., and Daniels, L. (2002) *Appl. Environ. Microbiol.* 68, 5750–5755
8. Purwantini, E., Gillis, T., and Daniels, L. (1997) *FEMS Microbiol. Lett.* 146, 129–134
9. Boshoff, H. I., and Barry, C. E., III (2005) *Nat. Rev. Microbiol.* 3, 70–80
10. Walsh, C. (1985) *Acc. Chem. Res.* 19, 216–221
11. Aufhammer, S. W., Warkentin, E., Berk, H., Shima, S., Thauer, R. K., and Ermier, U. (2004) *Structure* 12, 361–370
12. Aufhammer, S. W., Warkentin, E., Ermier, U., Hague, C. H., Thauer, R. K., and Shima, S. (2005) *Protein Sci.* 14, 1840–1849
13. Warkentin, E., Mamat, B., Sorel-Blippert, M., Wicke, M., Thauer, R. K., Iwata, M., Iwata, S., Ermier, U., and Shima, S. (2001) *EMBO J.* 20, 6561–6569
14. Purwantini, E., and Daniels, L. (1998) *J. Bacteriol.* 180, 2212–2219
15. Bashiri, G., Squire, C. J., Baker, E. N., and Moreland, N. J. (2007) *Protein Expression Purif.* 54, 38–44
16. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* 276, 307–326
17. Collaborative Computational Project No. 4 (1994) *Acta Crystallogr. Sect. D* 50, 760–763
18. Schneider, T. R., and Sheldrick, G. M. (2002) *Acta Crystallogr. Sect. D* 58, 1772–1779
19. de La Fortelle, E., and Bricogne, G. (1997) *Methods Enzymol.* 276, 472–494
20. Abrahams, J. P., and Leslie, A. G. W. (1996) *Acta Crystallogr. Sect. D* 52, 30–42
21. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Icoger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Scaccetti, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) *Acta Crystallogr. Sect. D* 58, 1948–1954
22. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. Sect. D* 60, 2126–2132
23. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. Sect. D* 53, 240–255
24. Winn, M., Isupov, M., and Murshudov, G. N. (2000) *Acta Crystallogr. Sect.*
25. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26**, 283–291
26. Ferrari, R. A., Mandelstam, P., and Crane, R. K. (1959) *Arch. Biochem. Biophys.* **80**, 372–377
27. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2005) *Acta Crystallogr. Sect. D* **61**, 458–464
28. Jacobson, F. S., Daniels, L., Fox, J. A., Walsh, C. T., and Orme-Johnson, W. H. (1982) *J. Biol. Chem.* **257**, 3385–3388
29. Eirich, L. D., Vogels, G. D., and Wolfe, R. S. (1978) *Biochemistry* **17**, 4583–4593
30. Eirich, L. D., Vogels, G. D., and Wolfe, R. S. (1979) *J. Bacteriol.* **40**, 20–27
31. Verdonk, M. L., Cole, J. C., Hartshorn, M. J., Murray, C. W., and Taylor, R. D. (2003) *Proteins* **52**, 609–623
32. Bair, T. B., Isabelle, D. W., and Daniels, L. (2001) *Arch. Microbiol.* **176**, 37–43
33. Dundas, I., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., and Liang, J. (2006) *Nucleic Acids Res.* **34**, 116–118
34. Krissinel, E., and Henrick, K. (2007) *J. Mol. Biol.* **372**, 774–797
35. Krissinel, E., and Henrick, K. (2004) *Acta Crystallogr. Sect. D* **60**, 2256–2268
36. Fisher, A. J., Thompson, T. B., Thoden, J. B., Baldwin, T. O., and Raynelt, I. (1996) *J. Biol. Chem.* **271**, 21956–21968
37. Klein, A. R., Berk, H., Purwantini, E., Daniels, L., and Thauer, R. K. (1996) *Eur. J. Biochem.* **239**, 93–97
38. Hui, L., Robertson, A. D., and Jensen, J. H. (2005) *Proteins* **61**, 704–721
39. Harrison, D. H., Bohen, K. M., Ringe, D., Petsko, G. A., and Gabbay, K. H. (1994) *Biochemistry* **33**, 2011–2020
40. Lyubimov, A. Y., Heard, K., Tang, H., Sampson, N. S., and Vrielink, A. (2007) *Protein Sci.* **16**, 2647–2656
41. Lenaerts, A. J., Gruppo, V., Marietta, K. S., Johnson, C. M., Driscoll, D. K., Tompkins, N. M., Rose, J. D., Reynolds, R. C., and Orme, I. M. (2005) *Antimicrob. Agents Chemother.* **49**, 2294–2301
42. Sassetti, C. M., Boyd, D. H., and Rubin, E. J. (2003) *Mol. Microbiol.* **48**, 77–84
43. Darwin, K. H., Eht, S., Gutierrez-Ramos, J.-C., Weich, N., and Nathan, C. F. (2003) *Science* **302**, 1963–1966
44. Shi, L., Sohaskey, C. D., Kana, B. D., Dawes, S., North, R. J., Mizrahi, V., and Gennaro, M. L. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15629–15634
45. Sherman, D. R., Yusupov, M., Schnappinger, D., Liao, R., Harrell, M. I., and Schoolnik, G. K. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98**, 7534–7539
46. Gouet, P., Robert, X., and Courcelle, E. (2003) *Nucleic Acids Res.* **31**, 3320–3323