Crystal Structure of Reduced MsAcg, a Putative Nitroreductase from *Mycobacterium smegmatis* and a Close Homologue of *Mycobacterium tuberculosis* Acg

François-Xavier Chauviac[‡1]**, Martin Bommer**[§¶]**, Jun Yan**[‡||]**, Gary Parkin**[**]**, Tina Daviter**[‡‡]**, Philip Lowden**[‡]**, Emma L. Raven**[**]**, Konstantinos Thalassinos**[‡||]**, and Nicholas H. Keep**[‡2]

From the ‡*Crystallography, Institute for Structural and Molecular Biology, Department of Biological Sciences, Birkbeck, University of London, Malet Street, London, WC1E 7HX, United Kingdom,* §*Institut für Biologie, Strukturbiologie/Biochemie, Humboldt-Universität zu Berlin, 10115 Berlin, Germany,* ¶*Macromolecular Crystallography Group, Institute of Soft Matter and Functional Materials, Helmholtz-Zentrum Berlin für Materialien und Energie, Albert-Einstein-Strasse 15, 12489 Berlin, Germany,* ||*Institute of Structural and Molecular Biology, Division of Biosciences, Darwin Building, University College London, Gower Street, London, WC1E 6BT, United Kingdom,* **Department of Chemistry, University of Leicester, Leicester, LE1 7RH, United Kingdom, and* ‡‡*Biophysics Centre, Institute for Structural and Molecular Biology, Department of Biological Sciences, Birkbeck, University of London, Malet Street, London, WC1E 7HX, United Kingdom*

Background: Acg proteins are up-regulated during dormancy in tuberculosis.
Results: Acg proteins bind flavin mononucleotide like nitroreductases but with the active site closed by a lid. They are not reduced by NADPH or NADH.
Conclusion: Acg proteins may have evolved from active nitroreductases to sequester FMN instead.
Significance: Turning off a flavin-dependent pathway may be important in tuberculosis dormancy.

This paper presents the structure of MsAcg (MSMEG_5246), a *Mycobacterium smegmatis* homologue of *Mycobacterium tuberculosis* Acg (Rv2032) in its reduced form at 1.6 Å resolution using x-ray crystallography. Rv2032 is one of the most induced genes under the hypoxic model of tuberculosis dormancy. The Acg family turns out to be unusual flavin mononucleotide (FMN)-binding proteins that have probably arisen by gene duplication and fusion from a classical homodimeric nitroreductase such that the monomeric protein resembles a classical nitroreductase dimer but with one active site deleted and the other active site covered by a unique lid. The FMN cofactor is not reduced by either NADH or NADPH, but the chemically reduced enzyme is capable of reduction of nitro substrates, albeit at no kinetic advantage over free FMN. The reduced enzyme is rapidly oxidized by oxygen but without any evidence for a radical state commonly seen in oxygen-sensitive nitroreductases. The presence of the unique lid domain, the lack of reduction by NAD(P)H, and the slow rate of reaction of the chemically reduced protein raises a possible alternative function of Acg proteins in FMN storage or sequestration from other biochemical pathways as part of the bacteria's adaptation to a dormancy state.

In the 21st century, tuberculosis remains a major burden to the world, causing around 1.4 million deaths and about 8.8 million new cases worldwide in 2010. The disease is deeply associated with poverty, and thus the developing world bears the bulk (95%) of the victims (1). It is estimated that a third of the world population is infected by *Mycobacterium tuberculosis*, the causative agent of tuberculosis. Among the carriers of *M. tuberculosis*, 5–10% will develop active tuberculosis during their lifetime (2), whereas the rest will remain in the latent phase of the disease. Thus, understanding the physiological state of the bacteria during this latency is of critical interest for the fight against tuberculosis. A particular problem is that bacteria in the dormant phase are far less sensitive to antibiotics, which means a long course of antibiotic treatment is required (3).

The exact nature of the dormant state within the lungs is hard to probe. A generally accepted laboratory model of dormancy, the Wayne model, is induced by hypoxia (4). Expression studies show that the cells adapt to hypoxia due to a two-component regulatory system (2CR)[3] called *dosRS* (5, 6). The *dosRS* 2CR controls the transcription of a regulon formed of around 50 genes induced under hypoxia (5) but also under nitric oxide exposure (7). Other conditions that trigger the *dosRS* regulon *in vitro* include *S*-nitrosoglutathione (8) and carbon monoxide (9). These signals are thought to be present in macrophages after their infection by *M. tuberculosis*. Inactivation of the whole *dos* operon unexpectedly produced a hypervirulent mutant in mice and in macrophage models of tuberculosis (10).

Among the 50 genes under the control of dosRS 2CR, few are of known function, making it difficult to understand the phys-

---

[1] Sponsored by a Bloomsbury Colleges Studentship. Present address: Virology Dept., Structural Virology Unit, Institut Pasteur, 28 rue du docteur Roux, 75724 Paris cedex 15, France.
[2] To whom correspondence should be addressed: Crystallography, Institute for Structural and Molecular Biology, Department of Biological Sciences, Birkbeck, University of London, Malet Street, London WC1E 7HX, UK. Tel.: 4420-7631-6852; Fax: 4420-7631-6803; E-mail: n.keep@mail.cryst.bbk.ac.uk.

[3] The abbreviations used are: 2CR, two component regulator; *acg*, α-crystalline co-regulated gene (Rv2032); MS, mass spectrometry; MsAcg MSMEG_5246, a *M. smegmatis* homologue of acg, RvXXXX, gene number XXXX in the *M. tuberculosis* genome; r.m.s.d., root mean square deviation.

iological outcome of the induction of the regulon. The most highly induced protein by the dosRS 2CR is Acr1 (Rv2031), which was previously shown to be the most abundant protein in dormant cells and to be essential for growth of the bacteria in macrophages (11, 12) while being induced by hypoxia or nitric oxide exposure. Acr1 (α-crystallin 1) is a small heat shock protein that forms a tetrahedral dodecamer (13), one of two small heat shock proteins in *M. tuberculosis*. Small heat shock proteins bind to partially unfolded proteins *in vitro* and prevent aggregation (14). The best studied example of a small heat shock protein is α-crystallin, which is proposed to prevent formation of cataract in the eye lens, where the cells do not undergo further protein synthesis or division after birth (15). Deletion of Acr1 led to an attenuated phenotype in macrophage infection model (12), whereas another study reported that the deletion of Acr1 induces a faster growth of the *M. tuberculosis* in macrophages and in mice (16). The discrepancy is due to differences in the knock-out strategy.

A study of the promoter region of *acr1* by Purkayastha *et al.* (17) showed that the divergently transcribed genes *acr1* and *Rv2032* are co-regulated under hypoxic conditions but also inside macrophages and led to the renaming of *Rv2032* as *acg* (*acr* co-regulated gene). The authors tentatively annotated Acg (the product of *acg*) as an unusual member of the classical nitroreductase family (Pfam ID PF00881). Acg, which is 331 amino acids long, is somewhat larger than the classical nitroreductases, which are around 220 amino acids long. There are two additional paralogues of Acg in the *M. tuberculosis* genome Rv3127 and Rv3131, all of which are controlled by the dosRS 2CR (8, 18).

Nitroreductases are broad substrate range enzymes that utilize FMN as co-factor, are homodimers, and use NAD(P)H as the source of reducing power. They catalyze the reduction of nitro compounds to hydroxylamino or amino forms. They are found in all kingdoms and are present in higher eukaryotes. The physiological role of nitroreductases is unclear (19); in some cases they are involved in catabolism of xenobiotics such as trinitrotoluene or in activation or deactivation of drugs, but these compounds are of recent origin, and therefore, this is unlikely to be the role for which these proteins originally evolved. Two classes of enzyme can be identified depending on the reduction mechanism: type I, oxygen-insensitive nitroreductases, and type II, oxygen-sensitive nitroreductases. The oxygen-insensitive enzyme catalyze the reduction of the nitro group by a direct two electron transfer, whereas the oxygen-sensitive enzyme catalyze the reduction through a one-electron transfer and the formation of a nitro anion radical before a second one-electron transfer to yield the product. However, in the presence of oxygen the nitro anion radical can be reoxidized to the nitro compound in a futile cycle that utilizes reducing power, generates superoxide anions and does not yield reduced substrate. The most studied class, the type I is further divided in two groups, A and B, depending on their usage of only NADPH or either NADPH or NADH, respectively, but all follow a ping-pong bi-bi mechanism (19).

Homologues of Acg are largely restricted to *Actinomyces* but are also found in some *Burkholderia* species. Three of the six putative nitroreductases in *M. tuberculosis* belong to the Acg

family. The other three are a classical nitroreductase (Rv3368), Rv0306, a *bluB* gene involved in dimethybenzimidazole biosynthesis, and Rv3262, a γ-glutamyl ligase involved in F420 biosynthesis. In this paper we present the structure of the first member of the Acg family of proteins in its reduced state, which gives insight into the role of these proteins in dormant tuberculosis.

EXPERIMENTAL PROCEDURES

Cloning—The genomic DNA of *M. tuberculosis* and *M. smegmatis* were kindly provided by Prof. Neil Stoker at the Royal Veterinary College, London, UK and by Dr. Sanjib Bhakta at Birkbeck, University of London, London, UK, respectively. The genes were cloned in pNIC28-Bsa4 using the ligation-independent cloning strategy developed at the Structural Genomic Consortium, Oxford, UK by Dr. Opher Gileadi (20). The MSMEG_5246 gene was amplified by PCR using the following primers designed using the gene sequence from the *M. smegmatis* MC[2]-155 genome (GenBank[TM] accessions number ABK74735.1) flanked with the ligation-independent cloning forward and reverse sequences, respectively, 5′-tacttccaatcc-ATGTCCGACACCAGGCTCGACG-3′ and 5′-tatccacctttactgTCAACTCCTCGGGCGAACTTCG-3′. The Rv2032 (*Acg*) gene was amplified by PCR using the following primers designed using the gene sequence from the *M. tuberculosis* H37Rv genome (GenBank[TM] CAA17246.1) flanked with the ligation-independent cloning forward and reverse sequences respectively, 5′-tacttccaatccATGCCGGACACCATGGTG-3′ and 5′-tatccacctttactgCTACCGGTGATCCTTAGCCCGA-3′. The cloned sequences were verified by sequencing.

Protein Expression and Purification—The pNIC28-Bsa4-MSMEG_5246 plasmid was transformed into *Escherichia coli* Rosetta2(DE3) strain (Novagen), and the cells were grown in ZYM5052 auto-inducing medium (21) at 37 °C for 3 h followed by 24 h at 25 °C. The selenomethionine-labeled protein was produced in glucose-free selenomethionine medium (AthenaES) supplemented with 5 g/liter glycerol, 0.5 g/liter glucose, and 2 g/liter lactose to obtain an auto-inducing medium and 80 mg/liter selenomethionine for labeling. 6.4 mg/liter methionine and 100 nM of vitamin B12 were also added to the medium to help repress the biosynthesis of methionine (21).

The cells were harvested by centrifugation and were resuspended in 100 mM Tris-HCl, pH 8.0, 500 mM sodium chloride, 0.02% (w/v) sodium azide before 0.5 mg of DNase I (Roche Applied Science) and 1.5 mg of lysozyme (Sigma) per 50 ml of cell suspension were added along with 1 mM PMSF and 1 tablet of cOmplete EDTA free protease inhibitor mix (Roche Applied Science). The cells were broken by sonication on ice. The cell lysate was cleared by centrifugation at 48,000 $\times$ *g* at 4 °C for 1 h. The cleared lysate was then applied to a XK 16/20 column (GE Healthcare) filled with 25 ml of nickel immobilized metal ion affinity chromatography-Sepharose HP (GE Healthcare) equilibrated with 100 mM Tris-HCl, pH 8.0, 500 mM sodium chloride, 25 mM imidazole, 0.02% (w/v) sodium azide. After extensive washing the bound protein was eluted in 100 mM Tris-HCl, pH 8.0, 500 mM sodium chloride, 300 mM imidazole, 0.02% (w/v) sodium azide. 1 mM β-mercaptoethanol and 1 mg of tobacco etch virus protease per 5 ml of eluted fraction were

added and the sample dialyzed overnight at 4 °C against 25 mM Tris-HCl, pH 8.0, 100 mM sodium chloride, 0.02% (w/v) sodium azide. The untagged protein was then applied to an anion exchange HiLoad 26/10 Q Sepharose HP column (GE Healthcare) equilibrated with 25 mM Tris-HCl, pH 8.0, 100 mM sodium chloride, 0.02% (w/v) sodium azide. The protein was eluted with a linear gradient from 100 to 500 mM NaCl in 25 mM Tris-HCl, pH 8.0, 0.02% (w/v) sodium azide. The fractions were pooled and concentrated using a Vivaspin centrifugal concentrator with a cutoff of 10 kDa (Sartorius) and loaded on a gel filtration HiPrep 16/60 Sephacryl S-200 HR column (GE Healthcare) equilibrated with 25 mM Tris-HCl, pH 8.0, 100 mM sodium chloride, 0.02% (w/v) sodium azide. Finally, the purified protein was concentrated to 10 mg/ml based on the extinction coefficient at 450 nm of 12,700 liter/cm/mol before being aliquoted and flash-frozen in liquid nitrogen and store at −80 °C.

The quality and oligomeric state of the purified protein was assessed by a combination of analytical gel filtration, SDS-PAGE stained by Coomassie blue based stain, and analytical ultracentrifugation sedimentation velocity experiments as well as ^{1}H NMR to assess the proper folding of the sample.

Analytical Ultracentrifugation—Sedimentation velocity experiments were carried out on a ProteomeLab^{TM} XLI analytical ultracentrifuge (Beckman Coulter) with an An-60 Ti rotor using 2-channel centerpieces and sapphire windows at 60,000 r.p.m. Scans were recorded at 20 °C in absorbance mode at 280 and 450 nm as well as in interference mode. The protein samples were in 25 mM Tris, pH 8.0, 100 mM NaCl at 45, 23, and 4.3 μM. Protein partial specific volume and buffer density and viscosity were calculated using *SEDNTERP*. The experimental data were analyzed using *Sedfit* by fitting to the $c(s)$ with one discrete component model with MsAcg.

Mass Spectrometry—MsAcg was subjected to limited proteolysis by chymotrypsin for 4 h at 22 °C at 1.2 mg/ml protein and a protease ration of 1:2000 (w/w). The reaction was stopped by the addition of PMSF, and the sample was applied on an analytical gel filtration column for purification and analysis. Mass spectrometry experiments were carried out on a Synapt HDMS (Waters Ltd, Manchester, UK) mass spectrometer (22). The instrument was mass-calibrated using a 33 μM solution of caesium iodide in 250 mM ammonium acetate. 20 μM MsAcg sample was buffer-exchanged into 250 mM ammonium acetate, pH 7.5, using Biospin-6 columns (Bio-Rad) before being delivered to the mass spectrometer by means of nanoelectrospray ionization using gold-coated capillaries prepared in house. Typical instrumental parameters for native MS were as follows: source pressure 5 mbar, capillary voltage 1.0 kV, cone voltage 50 V, trap energy 8 V, transfer energy 6 V, bias voltage 5 V, and trap pressure 3.6×10^{-2} mbar. For tandem MS, trap energy was increased to 80 V, with all other parameters unchanged. Data acquisition and processing were carried out using MassLynx (Version 4.1) software (Waters Corp., Milford, MA). Mass assignment was achieved by the method described in Tito *et al.* (23).

Protein Crystallization—The crystallization trials were carried out with an Oryx Nano crystallization robot (Douglas Instruments) contained within in an anaerobic chamber (Coy) with a 95% nitrogen, 5% hydrogen atmosphere in the laboratory of Prof. Holger Dobbek at Humbolt-Universität zu Berlin. Sitting drop vapor diffusion trials were set up with 200 nl of protein + 200 nl of reservoir solution drops in MRC 2 sub-well plates using 80 μl of reservoir solution. The protein was kept reduced by the addition of 5 mM dithionite to the protein and the reservoir solution before setting up the crystallization trays.

The first crystals were obtained with the Morpheus screen (Molecular Dimension) (24) after 3–4 weeks in 0.1 M MES/imidazole, pH 6.5, 10% (w/v) PEG 20,000, 20% (v/v) PEG MME 550 with the halide, alcohol, ethylene glycol, amino acid, and carboxylic acid mixes. Final selenomethionine-labeled protein crystals were obtained after optimization in 0.1 M MES/imidazole, pH 6.5, 30% (v/v) PEG MME 550, 0.02 M sodium formate, 0.02 M ammonium acetate, 0.02 M trisodium citrate, 0.02 M sodium potassium DL-tartrate, and 0.02 M sodium oxamate reduced by 50 mM dithionite. No oxamate was present in the higher resolution dataset. The crystals were cryopreserved for data collection in the same buffer as the crystallization conditions with 20% (v/v) ethylene glycol as cryoprotectant.

Structure Determination and Refinement—A native protein data set at 2.2 Å and selenomethionine-labeled protein data sets at 2.4 and 1.6 Å resolution were collected on BL 14.1 or 14.2 at the BESSY II synchrotron operated by Helmholtz Zentrum Berlin für Materialen und Energie (25). The datasets were indexed and integrated using XDS (26) before being converted into CCP4 format by COMBAT (27). The datasets were scaled using POINTLESS/SCALA. The structure was solved by Se-SAD using the SAS protocol of Auto-Rickshaw, the EMBL-Hamburg automated crystal structure determination platform using the 2.4 Å dataset (28, 29). Anomalous scatterer structure factor amplitudes (FA) values were calculated using the program SHELXC (30). All of the four heavy atoms requested were found using the program SHELXD (31). The correct hand for the substructure was determined using the programs ABS (32) and SHELXE (33). Initial phases were calculated after density modification using the program SHELXE (33). 97.27% of the model was built using the program ARP/wARP (34, 35). The model was refined using REFMAC (36), and the model building was done using COOT (37). The native dataset to 2.2 Å was refined from this starting structure using REFMAC (36). A further 1.6 Å dataset was initially refined in REFMAC and then further refined with phenix.refine (38). There were no significant differences between the structures (r.m.s.d. 0.14 Å over 325 residues). Figures were drawn using CCP4MG (39). The Butterfly angle was determined using Chimera (40) axes/plane/centeroid function using 10 atoms per plane as defined in Lennon *et al.* (41) and other stereochemistry analyzed with Molprobity (42). The structure factors and the structure coordinates of the final high resolution selenomethionine structure of MsAcg have been deposited to the Protein Data Bank at PDBe under accession number 2ymv.

Anaerobic Titration—All titration experiments were carried using sealable quartz cuvette (Hellma, Germany) with a new septum for each experiment on a Cary 300 UV-visible spectrophotometer (Agilent). The cuvettes were filled and sealed in an anaerobic chamber (Coy) filled with nitrogen and around 3% hydrogen atmosphere. All buffers and solutions were made

anaerobic by repeated cycles of evacuation and flushing with nitrogen in the airlock of the anaerobic chamber and kept on ice in the anaerobic chamber throughout the experiment. The stoichiometric titration solution consisted of $40-45$ μM MsAcg in 1 ml of 20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 0.02% (w/v) NaN$_3$ for the dithionite titration and of 13–15 μM MsAcg in 1 ml of 20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 0.02% (w/v) NaN$_3$ for the reoxidation experiment by nitrofuran. The titrants were either dithionite (calibrated against oxidized FMN), NADPH, and NADH (both in 0.01 M NaOH), calibrated spectrophotometrically (FMN$_{ox}$: $\lambda_{max} = 445$, $\epsilon_{445} = 12{,}600$ M^{-1} cm^{-1}; NADH: $\lambda_{max} = 340$, $\epsilon_{340} = 6{,}220$ M^{-1} cm^{-1}; NADPH: $\lambda_{max} = 340$, $\epsilon_{340} = 6{,}220$ M^{-1} cm^{-1}). Small amounts of titrant (5 μl) were injected through the septum with a gas-tight syringe (Hamilton), and a UV-visible spectrum was recorded between 300 and 700 nm. The nitrofuran reoxidation experiment was done by sequential addition of 5 μl to a total of 30 μl of dithionite at 1 mM in the same buffer as the protein to achieve almost total reduction. 5-μl additions of 1 mM nitrofuran in *N,N*-dimethylformamide were injected, and a spectrum between 600 and 300 nm was acquired after each addition. For the kinetic experiment, the protein was reduced following the same protocol as the other experiments followed by a single injection of 10 μl of 1 mM nitrofuran in *N,N*-dimethylformamide, and the absorbance was measured at 450 nm for 90 min.

Determination of the Mid-point Potential of the FMN—Reduction potentials were determined according to published methods (43– 46). The spectrophotometer was blanked with pH 7 buffer (50 mM potassium phosphate, 50 mM potassium chloride) containing 300 μM xanthine, 5 mM glucose, 1 μg/ml catalase, and 10 μg/ml glucose oxidase in a sealed cuvette. The dye (Nile blue, reduction potential -116 mV (47)) and protein were then titrated up to an absorbance of 0.1 at their respective maxima. Xanthine oxidase was added to a concentration of 40 μg/ml (43, 48). All enzymes and dyes used were sourced from Sigma. Spectroscopic readings were recorded systematically as the reduction potential of the cuvette is lowered. Readings were taken every 15–30 s at the start and gradually reduced to once every 5 min as the reaction slowed. Once the spectra remained constant for at least 10 min sodium dithionite was added, after which readings were taken until the spectra remained constant for at least a further 5 min. Reduction potentials are reported versus the normal hydrogen electrode (NHE). The results are the average of three separate determinations.

RESULTS

Protein Characterization—MsAcg, the gene product of MSMEG_5246, is one of four Acg family members in the fast growing mycobacteria *M. smegmatis.* Purified MsAcg has a deep yellow color as expected for flavoproteins. The cofactor extracted by thermal denaturation co-migrates with FMN and not FAD or riboflavin on thin layer chromatography. The overall purity of the sample is >95% as estimated from the SDS-PAGE and migrates just below the 40-kDa marker corresponding to the expected 36-kDa molecular mass. The oligomeric state of the protein was characterized in solution by analytical gel filtration (38.4 kDa) and analytical ultracentrifugation and

shown to be monomeric. Chymotrypsin-proteolysed protein retains FMN on gel filtration. The limited proteolysis of MsAcg resulted in cleavage at two sites: after residues 187 and 222. This was shown by native and tandem mass spectrometry (MS). Analysis of the chymotrypsin-proteolysed MsAcg by native nanoelectrospray MS showed predominantly two species corresponding to the cleaved (residues1–187 plus 223–330) and full-length (residues 1–330) MsAcg, both with FMN bound (Fig. 1*A*). Further investigation of these two species was made by tandem MS whereby each species was isolated using the quadrupole and subjected to collision induced dissociation to disrupt non-covalent interactions. Tandem MS of the cleaved MsAcg resulted in three species, one containing residues 223–330 and two containing residues 1–187 with and without FMN bound (Fig. 1*B*). Tandem MS of the full-length MsAcg (Fig. 1*C*) resulted in the same species and two further species that contain residues 1–222, again with and without FMN bound. MS of the proteolysed MsAcg protein under denaturing conditions, however, showed no evidence for the presence of full-length protein indicating that the full-length (residues 1–330) MsAcg observed in the native spectrum comprises the cleaved products held together by non-covalent interactions.

Structure of MsAcg—The three-dimensional structure of full-length MsAcg in the reduced form was obtained by selenium-SAD phasing in space group *P321* with a 2.4 Å dataset and refined against a second 1.6 Å selenomethionine containing dataset (Table 1). The cloned protein is 330 residues long, as an extra serine is left after the cleavage of the His tag (residue 1); all residues could be seen in the electron density map with the exception of residues 1, 2, and 330. Electron density in the region of FMN is shown in Fig. 2*C*.

The structure of MsAcg consists of a monomeric protein of two domains centered on β sheets. The N-terminal domain is formed of four antiparallel β strands, $\beta1$, $\beta2$, $\beta4$, $\beta3$, plus a complementary fifth β strand, $\beta9$, parallel to the first strand. The sheet is surrounded by α helices $\alpha1$, $\alpha2$, and $\alpha3$. The second domain consists of a four-stranded antiparallel β sheet formed by β strands $\beta5$, $\beta6$, $\beta8$, and $\beta7$ and surrounded by helices $\alpha4$, $\alpha5$, $\alpha6$, $\alpha7$, $\alpha8$, and $\alpha9$ (Fig. 2, *A* and *B*). A single FMN binds to the second domain. The fold of the monomer resembles a dimer of a normal nitroreductase. Below we compare MsAcg with NfnB from the *M. smegmatis* (49). Structure comparison against the PDB was done using the Fold server at European Bioinformatics Institute (50), and this proved to be sensitive to the protocol used. There is no one clear top hit, so we have selected NfnB as being from the same species and a strong hit based on a balance of the various match criteria. Of 174 matches to nitroreductase chains in the PDB to the full-length MsAcg at normal accuracy (40% match limit), NfnB is first in terms of sequence identity (21%), 4th in alignment length (245) (top hit 246), and top for secondary structure element matches (83%). On the statistical measures Q (53rd, 0.25 max 0.32), P (12th, 14.3 max 15.4), and Z (12th, 12.4 max 13.3), NfnB is less prominent but not much lower in score than the higher hits.

Secondary structure matching superposition, carried out using *CCP4MG* (39), shows that one molecule of MsAcg superimposes on a homodimer of NfnB, 234 residues per chain, (PDB ID 2wzw) with an overall r.m.s.d. of 2.4 Å for a length of super-
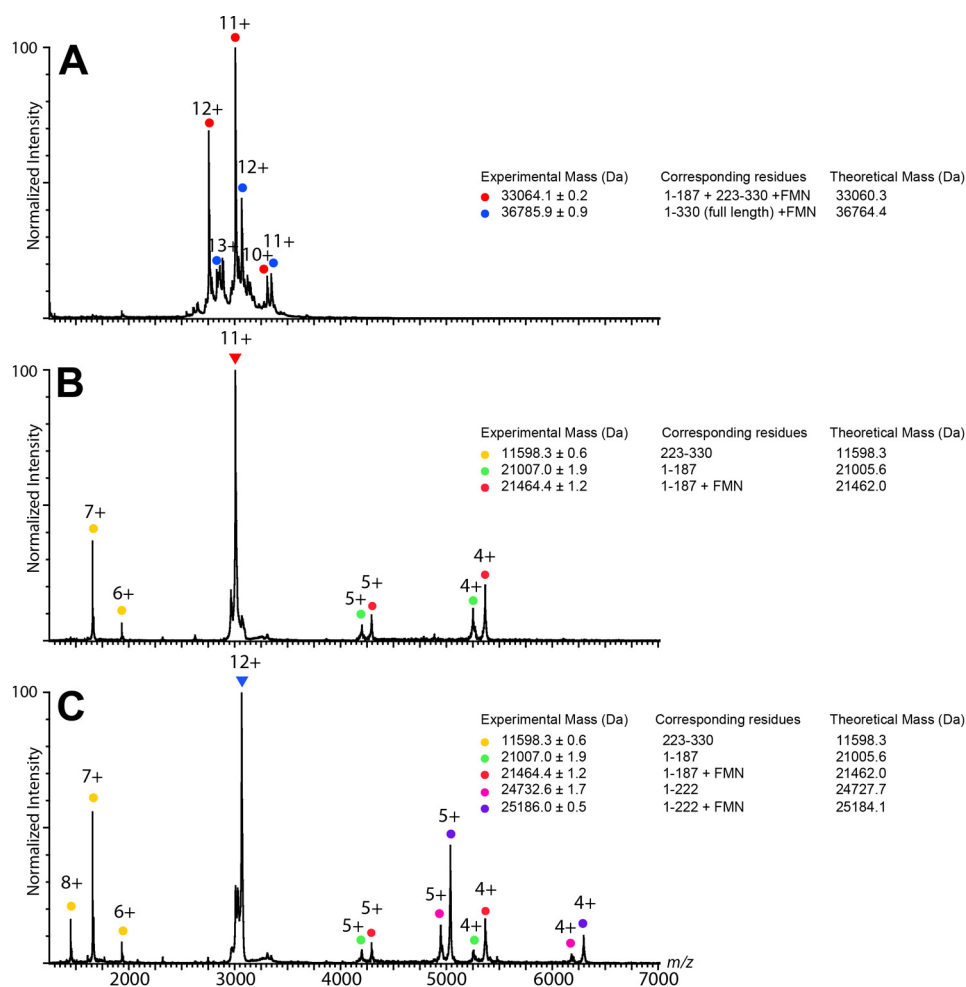
FIGURE 1. **Mass spectrometry analysis of proteolysed MsAcg.** *A*, shown is a native mass spectrum of the protein. *B*, shown is tandem mass spectrum of the 11+ ion of the lid free protein. *C*, shown is tandem mass spectrum of the 12+ ion of the full size protein.

position of 233 residues. The C-terminal region of MsAcg, *i.e.* between residues 109 and 328, aligns with the B chain of NfnB (r.m.s.d. of 2.5 Å over 147 residues), whereas the N-terminal region of MsAcg (residues 4–103) shows a superposition of 86 residues with a r.m.s.d. of 2.2 Å with chain A of NfnB. This shows that MsAcg could be the product of a gene duplication and fusion. The match is, however, not perfect as a portion of the chain B of NfnB, *i.e.* from residues 84 to 137, does not align with MsAcg, and a fragment of MsAcg between residues 170 and 231 does not align with NfnB despite being similar lengths (Fig. 3). The N-terminal region of MsAcg aligns with chain A of NfnB preserving the core β sheet and the loop that contributes to FMN binding to chain B but omitting most of the helical regions that contribute to the FMN binding site in chain A and the final β strand that extends the β sheet of chain B (Fig. 3). This shows that the N-terminal domain is reminiscent of a nitroreductase fold but has suffered from drastic evolution leading to a reduction to almost half the size of the C-terminal region and loss of the second FMN binding site. Although FMN binds to the C-terminal domain of MsAcg, residues identified as interacting with FMN (see below) are located in a region of the nitroreductase fold that is deleted in the N-terminal domain, thus making impossible the binding of FMN to the

N-terminal β sheet and excluding the presence of a second active site.

There is one other monomeric nitroreductase structure known, that of Tm1586 from *Thermotoga maritima* (PDB ID 1vkw), which is only 205 residues.[4] Again this contains two β sheets in a single chain similar to MsAcg. Tm1586 has no FMN bound but does have a sulfate from the crystallization condition at the expected position of the FMN phosphate in the N-terminal β sheet, so probably binds FMN but weakly enough to have been displaced at some point in the crystallization or purification. The C-terminal β sheet of Tm1586 does not conserve the FMN binding residues.

Upon examination of the surface representation of *M. smegmatis* NfnB (Fig. 4*A*) and indeed almost all other nitroreductase structures, the FMN is easily accessible to both solvent and substrates, whereas in the full-length MsAcg it is almost impossible to access the FMN except via a very narrow cavity, much smaller than most substrates or reducing agents (Fig. 4*B*). The region of the C terminus 170–231 that does not overlap with the B chain of NfnB blocks access, and we term this the lid (shown in *green in the figure*). However, when the lid is

---

[4] Joint Center for Structural Genomics, unpublished information.

TABLE 1

Data collection and refinement statistics

Numbers between parentheses are for the highest resolution shell.

Data set	Phasing set	Final deposited dataset
Beamline	14.1 @ BESSY II	14.2 @ BESSY II
Wavelength (Å)	0.91841	0.91841
Space group	*P321*	*P321*
Unit cell parameters (Å,°)	$a = b = 167.32, c = 43.93, \alpha = \beta = 90.00, \gamma = 120.00$	$a = b = 167.8, c = 43.98, \alpha = \beta = 90.00, \gamma = 120.00$
Resolution range (Å)	48.3-2.4 (2.49-2.4)	26.57-1.60 (1.63-1.6)
Total number of reflections	619,086 (64,268)	255,495 (12,681)
Number of unique reflections	27,778 (2,879)	92,796 (4,640)
Completeness (%)	100 (100)	99.4 (100)
Redundancy	22.3 (22.3)	2.8 (2.7)
$\langle I/\sigma(I)\rangle$	18.0 (5.5)	9.4 (2.5)
R_{merge}[a] (%)	16.0 (70.2)	6.1 (36.1)
R_{pim}[b] (%)	4.9 (21.8)	4.0 (24.6)
Solvent content (%)	74.3	75.7
Molecule/asymmetric unit	1	1
Wilson *B*-factor (Å^2)	34.4	15.4
Number of selenium sites	4	
SHELXD Patterson Correlation Coefficient	24.4 (all) 15.1 (weak)	
SHELXE Phasing Figure of Merit	0.68	
Resolution range (Å)		26.6-1.6
R_{work}[c] (%)		0.1188
R_{free}[c] (%)		0.1470
Reflection, working		88,140
Reflection, free		4,626
No of non-H atoms		3,216
Protein		2,625
FMN		31
Water		518
Other compounds[d]		42
B-factor (Å^2)		
Protein		19.1
FMN		9.8
Water		35.4
Other compounds[d]		46.2
r.m.s.d. bond length (Å)		0.019
r.m.s.d. bond angle (°)		1.721
Favored	Ramachandran plot (%)	97.9
Allowed	Ramachandran plot (%)	1.8
Disallowed	Ramachandran plot (%)	0.3

[a] $R_{merge} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the average intensity of the reflection hkl.

[b] $R_{pim} = \Sigma_{hkl}\sqrt{(1/(n-1))}(\Sigma_i^n|I_i(hkl) - \langle I(hkl)\rangle|)/\Sigma_{hkl}\Sigma_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the average intensity of the reflection hkl.

[c] $R_{work} = \Sigma_{hkl}||F_{obs}| - |F_{calc}||/\Sigma_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively. R_{free} was calculated analogously for the test reflections, which were randomly selected and excluded from the refinement.

[d] Compounds found in the crystallization condition, *i.e.* imidazole, formate, acetate, PEG

removed from the model, the FMN is nicely exposed (Fig. 4*C*). The only other structure of a PF00881 protein with a concealed FMN is BluB, which catalyzes the cleavage of FMNH2 to dimethylbenzimidazole (51). Although the analogous region of BluB contributes part of its lid, there is no conservation of sequence or backbone position. Mycobacteria contain more direct homologues of BluB than the Acg family, so Acg is unlikely to be carrying out dimethylbenzimidazole biosynthesis.

Oxygen can penetrate the lid in the crystal and reoxidize FMN to the yellow state, with concomitant loss of diffraction of the crystals. Therefore, very small molecules can access FMN in the closed form but larger substrates will require conformational change. The proteolytic sensitivity of the protein at the beginning and end of the lid region indicates that there are likely to be conformationally flexible hinges for the lid region allowing substrate access.

The structure of the co-factor was obtained in its reduced form (Fig. 2*C*). The isoalloxazine ring was refined without planarity constraints to allow the FMN to adopt a bent conformation. Indeed the isoalloxazine ring is bent along the N5–N10 axis toward *si* face, with a butterfly angle of 23.3°. The range of this angle is quite wide in flavins in all oxidation states (41);

however photoelectron reduction by X-rays of the flavin (52) may mean that the reported oxidation state may not be that in the crystal. Nevertheless the butterfly angle is larger than that of *M. smegmatis* NfnB (49) (PDB ID 1wzw, 13.4°) or *Enterobacter cloacae* NfsB in its oxidized state (PDB ID 1kqc, 15.5°) and similar to *E. cloacae* NfsB in the reduced state (53) (PDB ID 1kqd, 22.3°), consistent with the reduced form. This along with crystal color shows that we have a reduced FMN in our crystals.

The interaction between the reduced FMN co-factor and the polypeptide chain of MsAcg were plotted using LIGPLOT at the PDBsum server and summarized in Fig. 4*D*. The interactions are of two types, hydrogen bonds and hydrophobic interactions. A major cluster of hydrogen bonds is found around the phosphate group of the FMN. The phosphate group contacts two arginines (Arg-121 and Arg-317) that will form salt bridges as well as hydrogen bonds. Additionally two threonines interact with the phosphate (Thr-123 and Thr-315). Another cluster is found around the isoalloxazine ring, with O4 contacting Arg-215 and also the amino group of Thr-274. Similarly, nitrogen N5 from FMN contacts Thr-274 through its side chain hydroxyl. Finally, O2 interacts with the guanidinium group of Arg-125, whereas the oxygen O2* of the
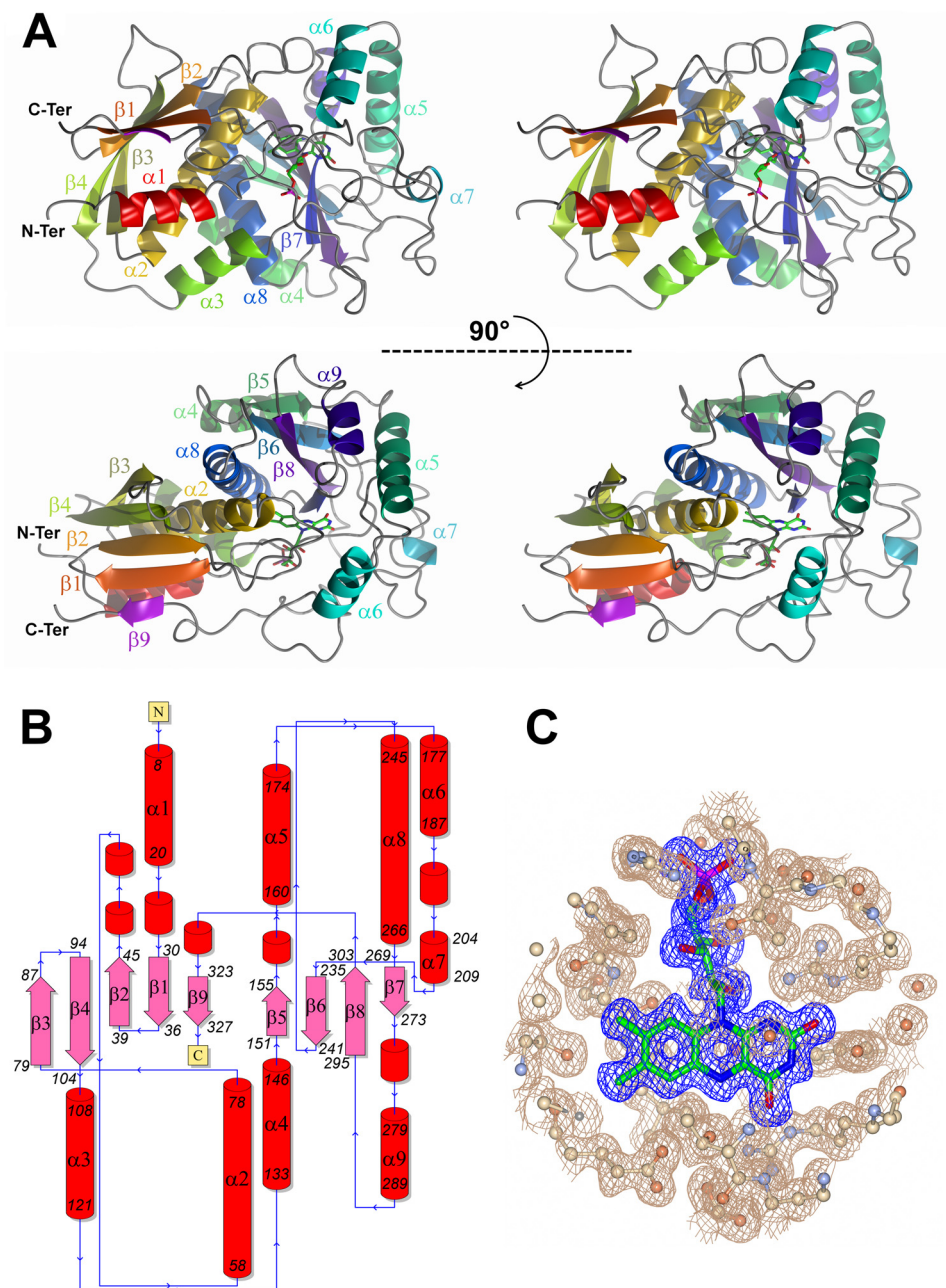
FIGURE 2. **Structure of MsAcg.** *A*, shows a stereo view of the structure of MsAcg in ribbon representation with two views 90° degrees apart around the horizontal. *B*, shows a schematic of the secondary structure. The protein is formed of two domains, one around each β sheet. The two β sheets are formed of β strands β1, -2, -4, -3, and -9 and β5, -6, -8, and -7, respectively, and surrounded, respectively, by α helices α1, -2, and -3 and α4, -5, -6, -7, -8, and -9. There is a single FMN co-factor associated with the second β sheet. *C*, shown is a final mF_o − dF_c electron density map in the region of the FMN. 10 Å extent of map, clipped to an 8 Å slice in z at 1.0 σ is shown in pale *brown*. The same map clipped to the FMN is shown in *blue*. The FMN is shown as *green ball-and-stick carbons*, and the adjacent protein residues are in *coral carbons*.

ribitol also interacts with the same residue. The hydrophobic interactions are spread over the whole FMN molecule (Fig. 4D). Arg-215 is part of the lid domain and is the least conserved interaction compared with FMN binding to *E. cloacae* NfsB or *M. smegmatis* NfnB.

Enzymological Studies—Neither NADPH nor NADH reduced the FMN of MsAcg, MTbAcg, or proteolytically treated MsAcg that partially removes the lid under aerobic or anaerobic conditions either by monitoring the nicotinamide at 340 nm or by monitoring the flavin at 445 nm. A control of *E. coli* NfnB behaved as expected (data not shown). Both

MsAcg and MTbAcg were reduced under anaerobic conditions by dithionite and rapidly reoxidized if oxygen was introduced into the cuvette. A single isosbestic point at 332 nm and a lack of any absorbance at 600 nm indicates that the semiquinone state is not formed and the reduction is a direct 2-electron reduction (Fig. 5A).

The mid-point reduction potentials were determined to be −165 ± 3 and −164 ± 3 mV (data not shown) for MTbAcg and MsAcg, respectively. The reduction potentials were calculated from experiments using Nile blue dye, which has a mid-point potential of −116 mV (47). These mid-point potentials of the
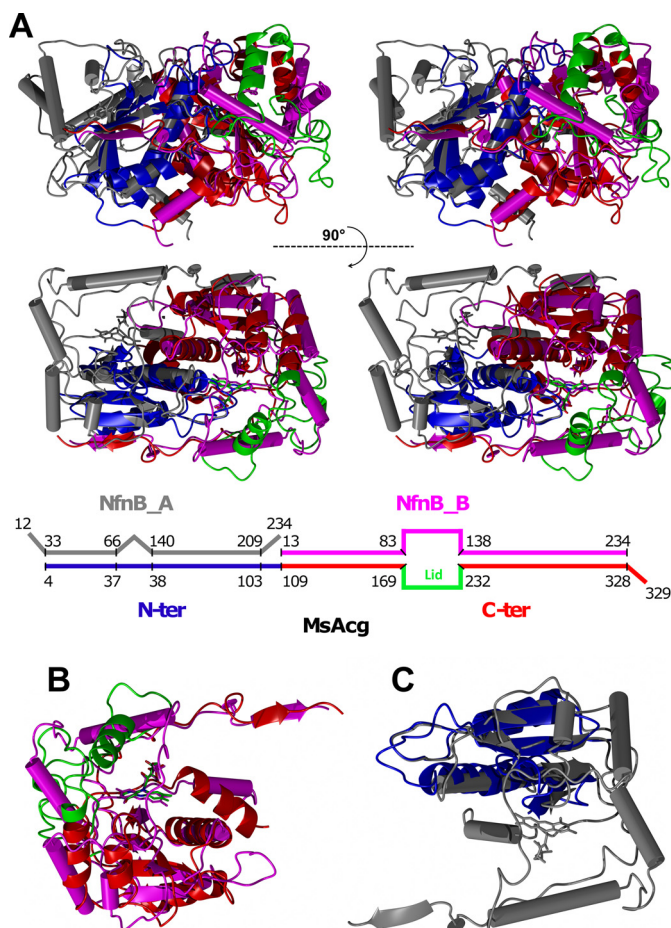
FIGURE 3. **Comparison of MsAcg with MsNfnB.** *A,* shown is a stereo view of a schematic representation of MsAcg colored by region (*blue* for the N-terminal region, *red* for the C-terminal region, and *green* for the lid) superimposed on MsNfnB (chain A in *gray* and B in *magenta*). Two views are shown 90° apart. A schematic of the corresponding regions is *below* shows gaps in the alignment bigger than 15 residues. *B,* shown is a close up of chain B of MsNfnB with the C-terminal region and lid of MsAcg. *C,* shown is a close up of chain A of NfsB with the N-terminal region of MsAcg.

Acgs are somewhat higher and, therefore, more weakly reducing than FMN alone, which has a 2-electron reduction potential of −207 mV (54) and than *E. cloacae* nitroreductase, an archetype of the FMN-dependent dimeric nitroreductases and 88% sequence identical to *E. coli* NfsB, which has a midpoint reduction potential of −196 mV (55).

The addition of the nitroreductase substrates nitrofuran, nitrofurantoin, and menadione to the anaerobic dithionite-reduced enzyme resulted in reoxidation of the FMN, implying turnover of the enzyme. The amount of dithionite was titrated to just reduce the FMN without excess dithionite. However, reduced FMN free of protein showed the same reaction, and indeed the rate of reoxidation of nitrofurantoin was significantly faster with free FMN compared with the enzyme bound FMN (Fig. 5, *B* and *C*).

DISCUSSION

The Acg family of proteins has evolved from the classic nitroreductase homodimer by gene duplication and fusion and then loss of one of the two active sites. This has resulted in a monomeric protein with a single active site but with an overall

fold resembling the nitroreductase homodimer. The proteins have also evolved to have a lid over the FMN rather than an open binding pocket as seen in the other nitroreductase proteins. The region between the main protein and the lid is sensitive to protease, indicating these might be flexible hinges allowing the lid to open, although some of the cleaved lid remains bound even after gel filtration, indicating tight noncovalent binding of the lid to the rest of the protein, so the rate of the lid opening is probably slow.

The lid does seem to be able to open, as we were able to show reoxidation of chemically reduced enzyme by larger nitroreductase substrates that would not have direct access to the active site in the conformation trapped in the crystal. The rate is actually slightly slower than that for the equivalent reaction carried out by reduced FMN in solution, so arguably we are merely measuring the rate of dissociation of the cofactor. However the fact that stoichiometric amounts of FMN are retained on the protein throughout purification would argue that cofactor binding is extremely tight and that we are seeing substrate interacting with protein bound cofactor. The lid domain interacts directly with the FMN through Arg-215, which interacts both with the ribitol and the O2 of the nicotinamide. Whereas the other interactions with the FMN are homologous to that seen in *E. cloacae,* which has a much lower midpoint potential, this interaction is unique to MsAcg and, therefore, may modify the midpoint potential as well as contributing to the fairly tight binding of the lid.

It is unclear what the physiological reducing agent is within *Mycobacteria* for Acg proteins. NAD(P)H is the normal reducing agent for nitroreductases; some proteins accept either as a substrate, but others are specific to NADPH. The only structure of a nitroreductase with NADPH bound is *M. smegmatis* NfnB (PDB ID 2wzw). The specific binding is to the adenosine region, but this is by a region of the fold that is not common to all nitroreductases. This region, which forms the rim of the pocket in *M. smegmatis* NfnB, comes from the same chain as the β sheet. However, the analogous region in the *E. coli* and *E. cloacae* structure is provided by the other chain, and the channel to the FMN is formed across the dimer interface, and it is not clear how NADPH binds in this case. As there is lack of conservation of the NADPH binding mode in other nitroreductases, it is impossible to say that NAD(P)H binding has definitely been lost.

A Δ*acg* strain of *M. tuberculosis* was not able to grow or survive in either resting or activated macrophages *in vitro* (56). The knock-out strain was not able to kill immunosuppressed (SCID) mice, indicating that Acg has roles in growth within macrophages and virulence. The Δ*acg* mutant is more sensitive to the nitroreductase-activated antibiotics nitrofuran and, to a lesser extent, nitrofurantoin than the wild type or the complemented strain. These antibiotics are activated by nitroreductases to what is thought to be to the hydroxylamine form of the drug (57), which is a DNA-damaging agent. As mutations of type 1 oxygen-insensitive nitroreductases can confer resistance to nitrofurantoin in *E. coli* (58), it suggests that the futile cycling of an oxygen sensitive nitroreductase is not the cause of toxicity. This implies that nitroreductase activity increases in the Δ*acg* mutant of *M. tuberculosis.* Therefore, Acg acts as an inhibitor of
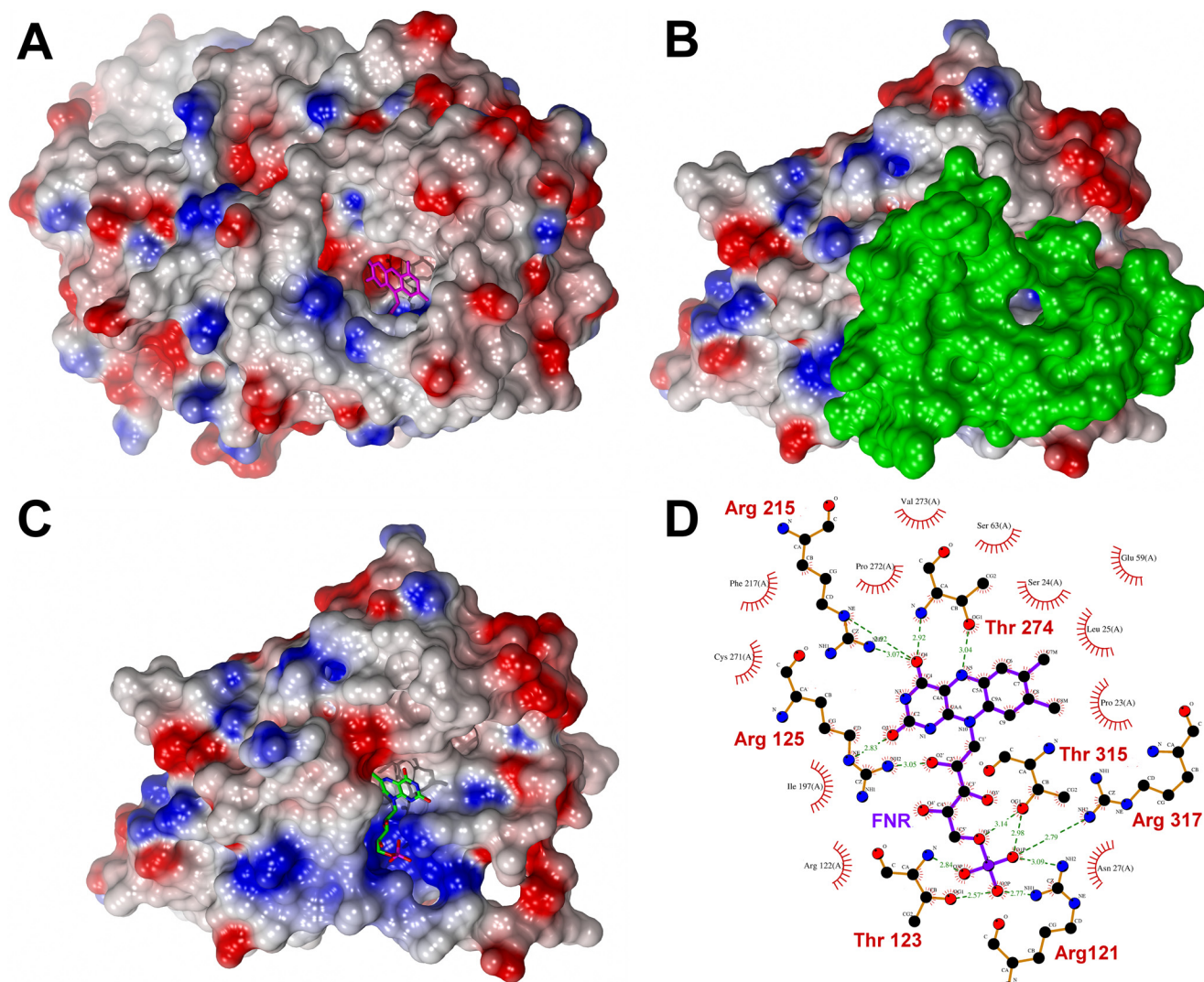
FIGURE 4. **Binding of FMN to MsAcg.** *A*, the electrostatic surface of MsNfnB shows the direct access to the FMN. *B*, shown is the electrostatic surface of MsAcg with the lid domain in *green*. *C*, the same view with the lid domain (169–232) removed shows FMN is now accessible. *D*, shown is a Ligplot drawing of the FMN binding pocket in MsAcg.

another nitroreductase, presumably the single classical nitrore-ductase Rv3368, rather than being a nitroreductase itself. This then produces an alternative model for the function of the Acg family as sequestering flavin. Acg would then bind FMN under its lid, depleting the classical nitroreductase of flavin and thus reducing the susceptibility to classical nitroreductase-activated drugs. The lack of a physiological reducing agent would not then be an issue, as the Acg proteins are not designed to function as enzymes. However, flavins are normally tightly bound in enzymes (K_d of 10^{-7}–10^{-10} M^{-1}) (59) so the inhibition is more likely to come from competition among newly synthesized apoenzymes than active extraction from holoenzymes. The inhibition of nitroreductase activity by Acg may well be a side effect rather than a cause, with another pathway being the primary target.

Flavin biosynthesis is essential for *M. tuberculosis*, as it has no flavin uptake pathway, and transposon-mediated disruption of several of the genes are fatal. Riboflavin biosynthesis enzymes are conserved in *Mycobacterium leprae*, regarded as a minimal mycobacterial genome. However, regulation of flavin synthesis

in *M. tuberculosis* is not understood, as there does not appear to be a riboswitch as found in *Bacillus subtilis* (60). More than 3% of *M. tuberculosis* genes are flavoproteins, the highest in a survey of 22 genomes (61), and is particularly rich in FAD-dependent acyl-CoA dehydrogenases required for lipid degradation. Inhibition of these enzymes by lack of free FMN to be converted to FAD may aid the switch to lipid body formation seen in persister-like bacteria (62). It may be relevant that the gene primarily implicated in this switch tgs1 Rv3130 is close to two of the Acg family members (Rv3127 and Rv3131) in the *M. tuberculosis* genome and also under control of the DosR regulon.

The Acg family of proteins has almost certainly evolved from the classical homodimeric nitroreductase by gene duplication, fusion, and reductive evolution to give proteins with a single FMN binding site protected by a lid domain. Lack of reduction by the normal physiological reducing agent NADPH raises the possibility that nitroreductase activity is no longer the function, although the protein will carry out these reactions when chemically reduced. It may be that the lid over the active site has evolved to store flavin or inhibit certain flavin-dependent met-
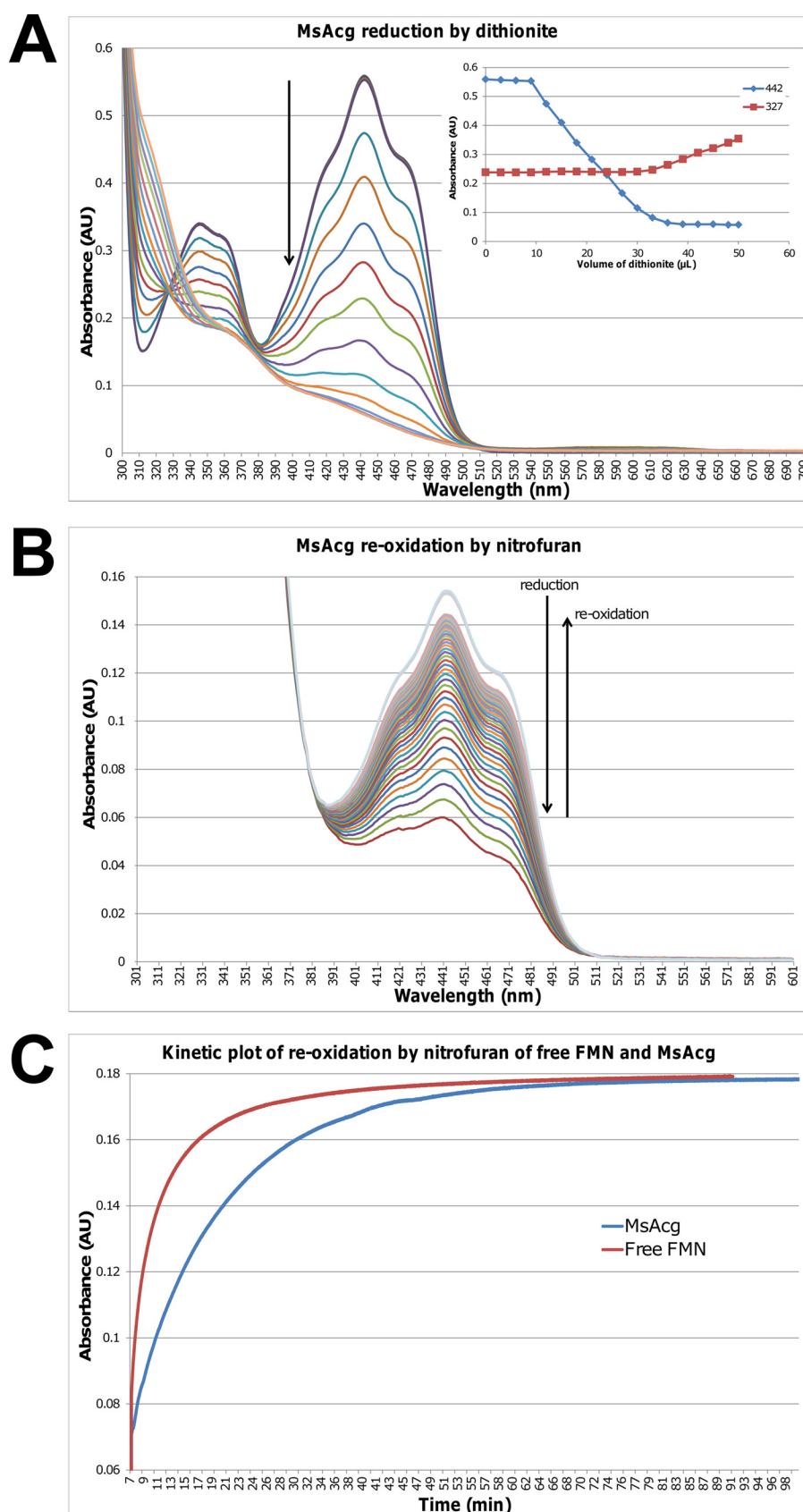
FIGURE 5. **Reduction oxidation of MsAcg.** *A*, shown is reduction of MsAcg by sequential addition of dithionite. *AU*, absorbance units. *B*, shown is re-oxidation of dithionite reduced MsAcg by sequential addition of nitrofuran. *C*, shown is the rate of oxidation of reduced free FMN and MsAcg bound FMN by nitrofuran.

abolic pathways during dormancy by sequestration of flavin cofactors.

REFERENCES

1. World Health Organization (2011) 2011/2012 Tuberculosis Global Facts
2. Parrish, N. M., Dick, J. D., and Bishai, W. R. (1998) Mechanisms of latency in *Mycobacterium tuberculosis*. *Trends Microbiol.* **6,** 107–112
3. Hu, Y., and Coates, A. R. M. (2003) in *Dormancy and Low Growth States in Microbial Disease* (Coates, A. R. M., ed) pp. 181–207, Cambridge University Press, Cambridge, UK
4. Wayne, L. G., and Sohaskey, C. D. (2001) Nonreplicating persistence of *Mycobacterium tuberculosis*. *Annu. Rev. Microbiol.* **55,** 139–163
5. Sherman, D. R., Voskuil, M., Schnappinger, D., Liao, R., Harrell, M. I., and Schoolnik, G. K. (2001) Regulation of the *Mycobacterium tuberculosis* hypoxic response gene encoding α-crystallin. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 7534–7539
6. Park, H. D., Guinn, K. M., Harrell, M. I., Liao, R., Voskuil, M. I., Tompa, M., Schoolnik, G. K., and Sherman, D. R. (2003) Rv3133c/dosR is a transcription factor that mediates the hypoxic response of *Mycobacterium tuberculosis*. *Mol. Microbiol.* **48,** 833–843
7. Voskuil, M. I., Schnappinger, D., Visconti, K. C., Harrell, M. I., Dolganov, G. M., Sherman, D. R., and Schoolnik, G. K. (2003) Inhibition of respiration by nitric oxide induces a *Mycobacterium tuberculosis* dormancy program. *J. Exp. Med.* **198,** 705–713
8. Kendall, S. L., Movahedzadeh, F., Rison, S. C., Wernisch, L., Parish, T., Duncan, K., Betts, J. C., and Stoker, N. G. (2004) The *Mycobacterium tuberculosis* dosRS two-component system is induced by multiple stresses. *Tuberculosis* **84,** 247–255
9. Kumar, A., Deshane, J. S., Crossman, D. K., Bolisetty, S., Yan, B. S., Kramnik, I., Agarwal, A., and Steyn, A. J. (2008) Heme oxygenase-1-derived carbon monoxide induces the *Mycobacterium tuberculosis* dormancy regulon. *J. Biol. Chem.* **283,** 18032–18039
10. Parish, T., Smith, D. A., Kendall, S., Casali, N., Bancroft, G. J., and Stoker, N. G. (2003) Deletion of two-component regulatory systems increases the virulence of *Mycobacterium tuberculosis*. *Infect. Immun.* **71,** 1134–1140
11. Yuan, Y., Crane, D. D., and Barry, C. E., 3rd. (1996) Stationary phase-associated protein expression in *Mycobacterium tuberculosis*. Function of the mycobacterial α-crystallin homolog. *J. Bacteriol.* **178,** 4484–4492
12. Yuan, Y., Crane, D. D., Simpson, R. M., Zhu, Y. Q., Hickey, M. J., Sherman, D. R., and Barry, C. E., 3rd. (1998) The 16-kDa α-crystallin (Acr) protein of *Mycobacterium tuberculosis* is required for growth in macrophages. *Proc. Natl. Acad. Sci. U.S.A.* **95,** 9578–9583
13. Kennaway, C. K., Benesch, J. L., Gohlke, U., Wang, L., Robinson, C. V., Orlova, E. V., Saibil, H. R., Saibi, H. R., and Keep, N. H. (2005) Dodecameric structure of the small heat shock protein Acr1 from *Mycobacterium tuberculosis*. *J. Biol. Chem.* **280,** 33419–33425
14. Basha, E., O'Neill, H., and Vierling, E. (2012) Small heat shock proteins and α-crystallins. Dynamic proteins with flexible functions. *Trends Biochem. Sci.* **37,** 106–117
15. Horwitz, J. (2003) α-Crystallin. *Exp. Eye Res.* **76,** 145–153
16. Hu, Y., Movahedzadeh, F., Stoker, N. G., and Coates, A. R. (2006) Deletion of the *Mycobacterium tuberculosis* α-crystallin-like hspX gene causes increased bacterial growth *in vivo*. *Infect. Immun.* **74,** 861–868
17. Purkayastha, A., McCue, L. A., and McDonough, K. A. (2002) Identification of a *Mycobacterium tuberculosis* putative classical nitroreductase gene whose expression is coregulated with that of the acr gene within macrophages, in standing *versus* shaking cultures, and under low oxygen conditions. *Infect. Immun.* **70,** 1518–1529
18. Chauhan, S., Sharma, D., Singh, A., Surolia, A., and Tyagi, J. S. (2011) Comprehensive insights into *Mycobacterium tuberculosis* DevR (DosR) regulon activation switch. *Nucleic Acids Res.* **39,** 7400–7414
19. Roldán, M. D., Pérez-Reinado, E., Castillo, F., and Moreno-Vivián, C. (2008) Reduction of polynitroaromatic compounds. The bacterial nitroreductases. *FEMS Microbiol Rev.* **32,** 474–500
20. Savitsky, P., Bray, J., Cooper, C. D., Marsden, B. D., Mahajan, P., Burgess-Brown, N. A., and Gileadi, O. (2010) High-throughput production of human proteins for crystallization. The SGC experience. *J. Struct. Biol.* **172,** 3–13
21. Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41,** 207–234
22. Pringle, S. D., Giles, K., Wildgoose, J. L., Williams, J. P., Slade, S. E., Thalassinos, K., Bateman, R. H., Bowers, M. T., and Scrivens, J. H. (2007) An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *Int. J. Mass Spectrom.* **261,** 1–12
23. Tito, M. A., Tars, K., Valegard, K., Hajdu, J., and Robinson, C. V. (2000) Electrospray time-of-flight mass spectrometry of the intact MS2 virus capsid. *J. Am. Chem. Soc.* **122,** 3550–3551
24. Gorrec, F. (2009) The MORPHEUS protein crystallization screen. *J. Appl. Crystallogr.* **42,** 1035–1042
25. Mueller, U., Darowski, N., Fuchs, M. R., Förster, R., Hellmig, M., Paithankar, K. S., Pühringer, S., Steffien, M., Zocher, G., and Weiss, M. S. (2012) Facilities for macromolecular crystallography at the Helmholtz-Zentrum Berlin. *J. Synchrotron Radiat.* **19,** 442–449
26. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 125–132
27. Collaborative Computational Project, Number 4 (1994) The CCP4 suite. Programs for protein crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* **50,** 760–763
28. Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S., and Tucker, P. A. (2005) Auto-Rickshaw. An automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. *Acta Crystallogr. D. Biol. Crystallogr.* **61,** 449–457
29. Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S., and Tucker, P. A. (2009) On the combination of molecular replacement and single-wavelength anomalous diffraction phasing for automated structure determination. *Acta Crystallogr. D. Biol. Crystallogr.* **65,** 1089–1097
30. Sheldrick, G. M., Hauptman, H. A., Weeks, C. M., Miller, R., and Uson, I. (2001) in *International Tables for Crystallography, Volume F: Crystallography of Biological Macromolecules* (Rossmann, M. G., and Arnold, E., eds) pp. 333–345, Kluwer Academic Publishers, Dordrecht, The Netherlands
31. Schneider, T. R., and Sheldrick, G. M. (2002) Substructure solution with SHELXD. *Acta Crystallogr. D. Biol. Crystallogr.* **58,** 1772–1779
32. Hao, Q. (2004) ABS. A program to determine absolute configuration and evaluate anomalous scatterer substructure. *J. Appl. Crystallogr.* **37,** 498–499.
33. Sheldrick, G. M. (2002) Macromolecular phasing with SHELXE. *Zeitschrift für Kristallographie* **217,** 644–650
34. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.* **6,** 458–463
35. Morris, R. J., Zwart, P. H., Cohen, S., Fernandez, F. J., Kakaris, M., Kirillova, O., Vonrhein, C., Perrakis, A., and Lamzin, V. S. (2004) Breaking good resolutions with ARP/wARP. *J. Synchrotron Radiat* **11,** 56–59
36. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D. Biol. Crystallogr.* **53,** 240–255
37. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66,** 486–501
38. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX. A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66,** 213–221
39. Potterton, L., McNicholas, S., Krissinel, E., Gruber, J., Cowtan, K., Emsley, P., Murshudov, G. N., Cohen, S., Perrakis, A., and Noble, M. (2004) De-

velopments in the CCP4 molecular-graphics project. *Acta Crystallogr. D. Biol. Crystallogr.* **60,** 2288–2294

40. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera. A visualization system for exploratory research and analysis. *J. Comput Chem.* **25,** 1605–1612

41. Lennon, B. W., Williams, C. H., Jr., and Ludwig, M. L. (1999) Crystal structure of reduced thioredoxin reductase from *Escherichia coli.* Structural flexibility in the isoalloxazine ring of the flavin adenine dinucleotide cofactor. *Protein Sci.* **8,** 2366–2379

42. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity. All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 12–21

43. Massey, V. (1991) in *Flavins and Flavoproteins* (Curti, B., Ronchi, S., and Zanetti, G., eds) pp. 59–66, Walter de Gruyter, Berlin

44. Efimov, I., and McIntire, W. S. (2004) A study of the spectral and redox properties and covalent flavinylation of the flavoprotein component of *p*-cresol methylhydroxylase reconstituted with FAD analogues. *Biochemistry* **43,** 10532–10546

45. Basran, J., Rafice, S. A., Chauhan, N., Efimov, I., Cheesman, M. R., Ghamsari, L., and Raven, E. L. (2008) A kinetic, spectroscopic, and redox study of human tryptophan 2,3-dioxygenase. *Biochemistry* **47,** 4752–4760

46. Papadopoulou, N. D., Mewies, M., McLean, K. J., Seward, H. E., Svistunenko, D. A., Munro, A. W., and Raven, E. L. (2005) Redox and spectroscopic properties of human indoleamine 2,3-dioxygenase and a H303A variant. Implications for catalysis. *Biochemistry* **44,** 14318–14328

47. Clark, W. M. (1972) *Oxidation-Reduction Potentials of Organic Systems*, Robert E. Kreiger Publishing Co., New York

48. Efimov, I., Papadopoulou, N. D., McLean, K. J., Badyal, S. K., Macdonald, I. K., Munro, A. W., Moody, P. C., and Raven, E. L. (2007) The redox properties of ascorbate peroxidase. *Biochemistry* **46,** 8017–8023

49. Manina, G., Bellinzoni, M., Pasca, M. R., Neres, J., Milano, A., Ribeiro, A. L., Buroni, S., Skovierová, H., Dianišková, P., Mikušová, K., Marák, J., Makarov, V., Giganti, D., Haouz, A., Lucarelli, A. P., Degiacomi, G., Piazza, A., Chiarelli, L. R., De Rossi, E., Salina, E., Cole, S. T., Alzari, P. M., and Riccardi, G. (2010) Biological and structural characterization of the *Mycobacterium smegmatis* nitroreductase NfnB and its role in benzothiazinone resistance. *Mol. Microbiol.* **77,** 1172–1185

50. Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D. Biol. Crystallogr.* **60,** 2256–2268

51. Taga, M. E., Larsen, N. A., Howard-Jones, A. R., Walsh, C. T., and Walker, G. C. (2007) BluB cannibalizes flavin to form the lower ligand of vitamin B12. *Nature* **446,** 449–453

52. Røhr, A. K., Hersleth, H. P., and Andersson, K. K. (2010) Tracking flavin conformations in protein crystal structures with Raman spectroscopy and QM/MM calculations. *Angew Chem. Int. Ed Engl* **49,** 2324–2327

53. Haynes, C. A., Koder, R. L., Miller, A. F., and Rodgers, D. W. (2002) Structures of nitroreductase in three states. Effects of inhibitor binding and reduction. *J. Biol. Chem.* **277,** 11513–11520

54. Mayhew, S. G. (1999) The effects of pH and semiquinone formation on the oxidation-reduction potentials of flavin mononucleotide. A reappraisal. *Eur. J. Biochem.* **265,** 698–702

55. Nivinskas, H., Koder, R. L., Anusevicius, Z., Sarlauskas, J., Miller, A. F., and Cenas, N. (2001) Quantitative structure-activity relationships in two-electron reduction of nitroaromatic compounds by *Enterobacter cloacae* NAD(P)H:nitroreductase. *Arch. Biochem. Biophys.* **385,** 170–178

56. Hu, Y., and Coates, A. R. (2011) *Mycobacterium tuberculosis* acg gene is required for growth and virulence *in Vivo. PLoS One* **6,** e20958

57. McCalla, D. R., Reuvers, A., and Kaiser, C. (1970) Mode of action of nitrofurazone. *J. Bacteriol.* **104,** 1126–1134

58. Sandegren, L., Lindqvist, A., Kahlmeter, G., and Andersson, D. I. (2008) Nitrofurantoin resistance mechanism and fitness cost in *Escherichia coli. J. Antimicrob. Chemother.* **62,** 495–503

59. Lederer, F., Ruterjans, H., and Fleischmann, G., Chapman, S. K., and Reid, G. A. (1999) in *Flavoprotein Protocols*, pp. 149–155, Humana, Totowa, NJ

60. Long, Q., Ji, L., Wang, H., and Xie, J. (2010) Riboflavin biosynthetic and regulatory factors as potential novel anti-infective drug targets. *Chem. Biol. Drug Des.* **75,** 339–347

61. Macheroux, P., Kappes, B., and Ealick, S. E. (2011) Flavogenomics-a genomic and structural view of flavin-dependent proteins. *FEBS J.* **278,** 2625–2634

62. Garton, N. J., Waddell, S. J., Sherratt, A. L., Lee, S. M., Smith, R. J., Senner, C., Hinds, J., Rajakumar, K., Adegbola, R. A., Besra, G. S., Butcher, P. D., and Barer, M. R. (2008) Cytological and transcript analyses reveal fat and lazy persister-like bacilli in tuberculous sputum. *PLoS Med.* **5,** e75

Crystal Structure of Reduced MsAcg, a Putative Nitroreductase from *Mycobacterium smegmatis* and a Close Homologue of *Mycobacterium tuberculosis* Acg

François-Xavier Chauviac, Martin Bommer, Jun Yan, Gary Parkin, Tina Daviter, Philip Lowden, Emma L. Raven, Konstantinos Thalassinos and Nicholas H. Keep