Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallization and preliminary X-ray diffraction analysis of the high molecular weight ketoacyl reductase FabG4 complexed with NADH

Debajyoti Dutta, Sudipta Bhattacharyya and Amit Kumar Das*

Department of Biotechnology, Indian Institute of Technology, Kharagpur, Kharagpur 721 302, India

Correspondence e-mail: amitk@hijli.iitkgp.ernet.in

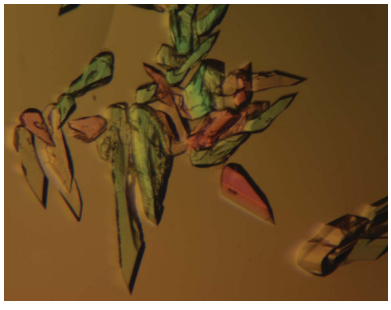
Received 12 January 2012
Accepted 5 May 2012

FabG4 from *Mycobacterium tuberculosis* belongs to the high molecular weight ketoacyl reductases (HMwFabGs). The enzyme requires NADH for β-ketoacyl reductase activity. The protein was overexpressed, purified to homogeneity and crystallized as a FabG4–NADH complex. A mountable FabG4:NADH complex crystal diffracted to 2.59 Å resolution and belonged to space group $P1$, with unit-cell parameters a = 63.07, b = 71.03, c = 92.92 Å, α = 105.02, β = 97.06, γ = 93.66°. The Matthews coefficient suggested the presence of four monomers in the unit cell. In addition, a self-rotation function revealed the presence of two twofold NCS axes and one fourfold NCS axis. At χ = 180° the highest peak corresponds to the twofold NCS between two monomers, whereas the second peak corresponds to the twofold NCS between two dimers.

1. Introduction

Complex circuits of lipid metabolism in *Mycobacterium tuberculosis* enable the organism to synthesize several unusual lipid derivatives for its virtually impenetrable cell envelope. Nonpolar lipid derivatives resist the entrance of deleterious molecules generated by the host into the bacterial cell. The bacteria also keep changing their cell-envelope composition in response to environmental conditions (Makinoshima & Glickman, 2005). Lipid metabolism, including fatty-acid metabolism, thus plays a significant role in this organism.

Fatty-acid synthesis in *M. tuberculosis* is classified into two types: FAS-I and FAS-II (Takayama *et al.*, 2005). The eukaryotic-like single enzyme multi-domain system FAS-I synthesizes fatty-acyl chains of up to 20 carbons in length. FAS-II, which is a set of enzymes, subsequently elongates the fatty-acyl chain up to 70–80 carbons in length. The elongation module of the FAS-II pathway consists of the enzymes ketoacyl synthase (KasA/KasB; Mdluli *et al.*, 1998; Schaeffer *et al.*, 2001; Gao *et al.*, 2003), ketoacyl reductase (FabG1 or MabA; Marrakchi *et al.*, 2002), hydroxyacyl dehydratase (HadAB/HadBC; Sacco *et al.*, 2007), enoyl reductase (InhA; Dessen *et al.*, 1995) and an acyl carrier protein (AcpM; Wong *et al.*, 2002). FabG1, which catalyses the second step in the FAS-II module, is also known as β-ketoacyl reductase. The *M. tuberculosis* genome encodes more than one copy of the β-ketoacyl reductase gene. The annotated FabG genes are *fabG1* (Rv1483), *fabG2* (Rv1350), *fabG3* (Rv2002), *fabG4* (Rv0242c), *fabG5* (Rv2766c) and *fabG6* (Rv3502c). Ketoacyl reductases belong to the short-chain dehydrogenase/reductase (SDR) family and mostly have a core structure of 250–350 amino acids in length (Oppermann *et al.*, 2003). All of the FabG gene products of *M. tuberculosis* are 247–317 amino acids in length, with the exception of FabG4. FabG4 is unusually long (454 amino acids) and contains two distinguishable domains (Dutta *et al.*, 2011). The N-terminal domain is a flavodoxin-type domain and the C-terminal is a ketoreductase domain. Thus, this protein, along with its sequence homologues, is categorized as a high molecular weight FabG (HMwFabG). Although the role of FabG4 in mycobacteria is yet to



be addressed, some reports have recently identified FabG4 as an essential gene in mycobacteria (Beste *et al.*, 2009; Gurvitz, 2009; Sharma *et al.*, 2010).

The C-terminal domain (215–454) of FabG4 shares 32% sequence identity with *M. tuberculosis* FabG1 (Cohen-Gonsaud *et al.*, 2002), 35% with *Plasmodium falciparum* FabG (Wickramasinghe *et al.*, 2006), 36% with *Brassica napus* FabG (Fisher *et al.*, 2000), 38% with *Staphylococcus aureus* FabG1 (Dutta *et al.*, 2012) and *Rickettsia prowazekii* FabG (Subramanian *et al.*, 2011), 39% with *Bacillus anthracis* FabG (Zaccai *et al.*, 2008), 40% with *Aquifex aeolicus* FabG (Mao *et al.*, 2007) and 41% with *Escherichia coli* FabG (Price *et al.*, 2001). Unlike all of these FabGs, FabG4 utilizes NADH for catalysis. In the cases of *E. coli* FabG and *M. tuberculosis* FabG1, NADPH has been hypothesized to reorient the catalytic residues in position for catalysis (Price *et al.*, 2004; Cohen-Gonsaud *et al.*, 2005). These holo structures also maintain two loops near the catalytic residues in stable conformations. In some FabG structures the stability of these two loops does not depend on the coenzyme (Wickramasinghe *et al.*, 2006). It has been proposed that the amino-acid residue preceding the catalytic serine is responsible for holding the catalytic loop in its proper position (Poncet-Montange *et al.*, 2007). We have recently determined the crystal structure of apo FabG4$^{17–448}$, truncating the six C-terminal residues ^{449}QAMIGA^{454} (Dutta *et al.*, 2011). The C-terminal residues are involved in a hydrophobic interaction with the corresponding catalytic loop in FabG4, thereby stabilizing it. This clearly indicates the role of the conserved C-terminus in HMwFabGs. Intriguingly, a conserved C-terminus is also observed for FabG1s among mycobacterial species (Cohen-Gonsaud *et al.*, 2005).

Although much study has been carried out on FabGs, the nature of the HMwFabGs has remained obscure, including their coenzyme specificity. The present work reports the overexpression, purification, crystallization and preliminary X-ray diffraction studies of the FabG4$^{17–454}$:NADH complex.

2. Materials and methods

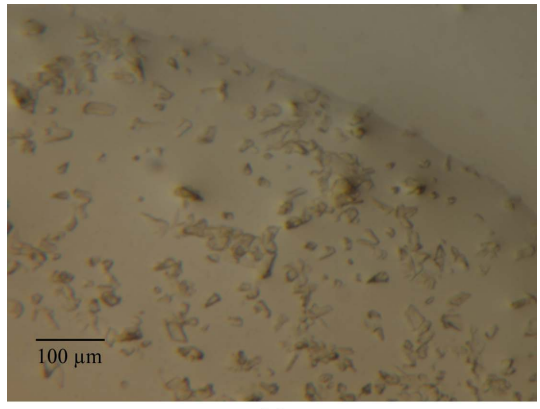
2.1. Purification of FabG4$^{17–454}$

The cloning of the FabG4 gene from the *M. tuberculosis* H37Rv genome in pQE30 expression vector and the overexpression of recombinant His$_6$-FabG4$^{17–454}$ in *E. coli* M15 (pREP4) cells have previously been reported (Dutta *et al.*, 2011). The entire FabG4 gene was not soluble. The protein became soluble upon truncation of the first 16 amino-acid residues.

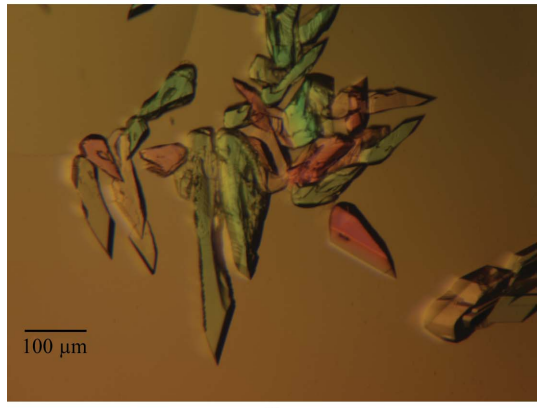
The cells from 2 l culture were resuspended in buffer *A* (10 m*M* Tris–HCl pH 7.9, 300 m*M* NaCl, 10 m*M* imidazole, 5% glycerol) containing 0.1 m*M* each of leupeptin, aprotinin and pepstatin and 0.02 m*M* phenylmethylsulfonyl fluoride. The following steps were carried out at 277 K. The suspension was lysed by ultrasonication and the resulting lysate was centrifuged at 22 000*g* for 30 min. The supernatant was passed through Ni–NTA Sepharose high-performance affinity matrix (GE Healthcare Biosciences) pre-equilibrated with buffer *A*. The column was then extensively washed with buffer *A*. An intermediate buffer *B* (10 m*M* Tris–HCl pH 7.9, 300 m*M* NaCl, 100 m*M* imidazole, 5% glycerol) was passed through the column to remove any nonspecifically bound contaminant. Finally, the protein was eluted with buffer *C* (10 m*M* Tris–HCl pH 7.9, 300 m*M* NaCl, 300 m*M* imidazole, 5% glycerol). The eluted protein was subjected to gel-filtration chromatography using Superdex 200 prep-grade matrix in a 16/70C column (GE Healthcare Biosciences) pre-equilibrated with buffer *D* (10 m*M* Tris–HCl pH 7.9, 200 m*M* NaCl, 5% glycerol) on an ÄKTAprime Plus system. The flow rate was set to 1 ml min^{-1}. The fractions containing the desired protein were pooled together and kept for further use. The purity of the protein sample was assessed by SDS–PAGE and the protein concentration was measured using both the Abs$_{280}$ and the Bradford method (Bradford, 1976). The absence of NAD$^+$ cofactor was confirmed by the OD$_{280}$/OD$_{260}$ ratio as described previously (Krimsky & Racker, 1963).

2.2. Crystallization of the FabG4$^{17–454}$:NADH binary complex

Purified FabG4 was concentrated to 20 mg ml^{-1} using a 10 kDa cutoff Vivaspin 20 concentrator (GE Healthcare). Concentrated protein was mixed with NADH (Sigma) in a 1:5 molar ratio followed by an incubation period of 30 min at 300 K. The mixture was subjected to crystallization using the previously reported crystallization condition for the C-terminally truncated mutant of apo FabG4 (Dutta *et al.*, 2011). Despite several trials, no crystals appeared from this condition. The mixture was thus subjected to preliminary crystallization trials using Crystal Screen, Crystal Screen 2 and Index from Hampton Research by mixing 2 µl protein/NADH mixture with an equal volume of reservoir solution using the sitting-drop vapour-diffusion method. Initial hits were only obtained using a condition consisting of 0.1 *M* bis-Tris pH 6.5, 45%(*v*/*v*) polypropylene glycol P400. The single crystals that appeared from the initial crystallization condition were too tiny to mount (Fig. 1*a*). The initial crystallization condition was further fine-screened using the hanging-drop vapour-diffusion method to optimize the crystal size and quality. After a number of screening trials with different buffering reagents and



(*a*)



(*b*)

Figure 1
Crystals of the FabG4$^{17–454}$:NADH complex. (*a*) Initial crystals obtained from the preliminary crystallization screening. (*b*) Mountable crystals obtained from 0.1 *M* MES pH 6.5, 45%(*v*/*v*) polypropylene glycol P400.

precipitating reagents, mountable crystals were obtained from the condition 0.1 *M* MES pH 6.5, 45%(*v*/*v*) polypropylene glycol P400 (Fig. 1*b*).

2.3. Data collection and processing

For data collection, an FabG4$^{17–454}$:NADH crystal was directly mounted from the mother liquor and flash-cooled in a nitrogen stream at 100 K. X-ray diffraction data were collected using an in-house X-ray diffraction facility. The facility was equipped with a Rigaku MicroMax-007 HF rotating-anode generator as a Cu $K\alpha$ X-ray source and a Rigaku R-AXIS IV^{++} image-plate detector. The crystal-to-detector distance was maintained at 200 mm and the crystal was rotated 360° with 1° rotation per frame. The crystal diffracted to a resolution of 2.59 Å. The images were processed with *XDS* (Kabsch, 2010) in space group *P*1 and scaled with *SCALA* (Evans, 1993) from the *CCP*4 suite (Winn *et al.*, 2011). The final statistics of data collection and processing are tabulated in Table 1.

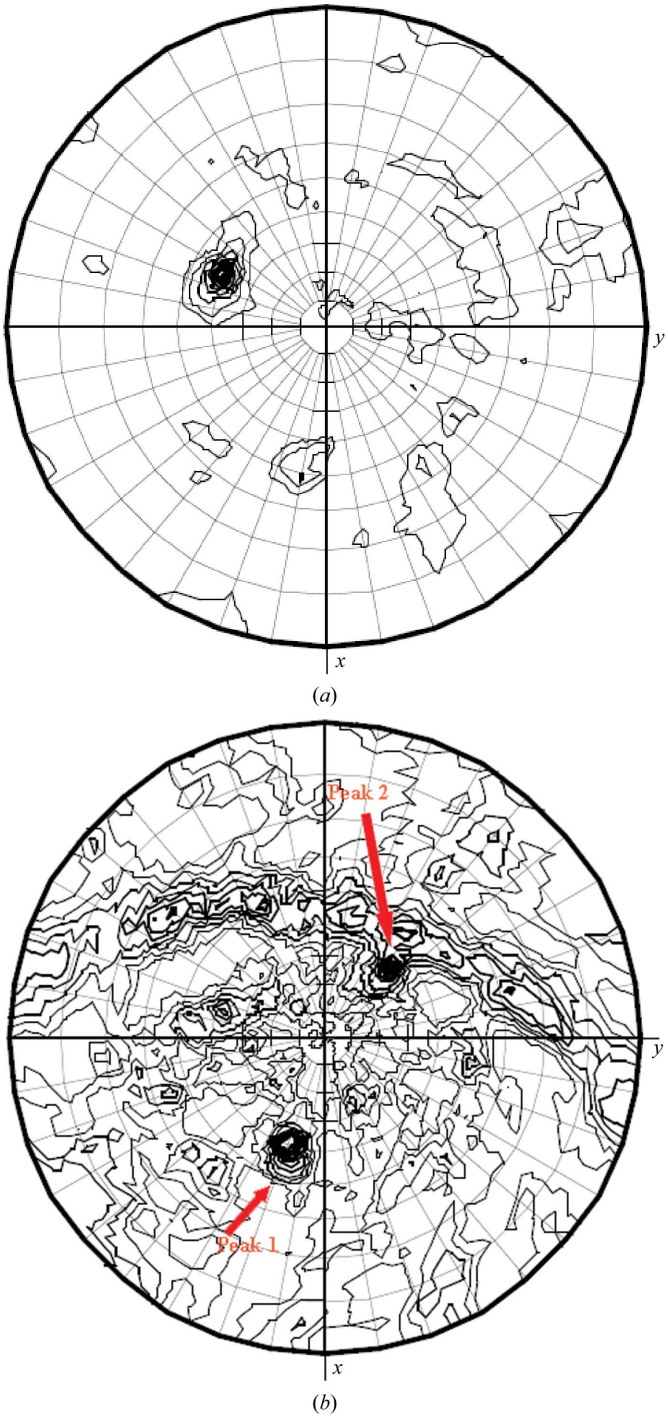


Figure 2
Self-rotation function plot obtained from the FabG4$^{17–454}$:NADH complex data for (*a*) χ = 90° and (*b*) χ = 180°. Two peaks (peak 1 and peak 2) are indicated by red arrows.

Table 1
Data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

Wavelength (Å)	1.54
Space group	*P*1
Unit-cell parameters (Å, °)	a = 63.07, b = 71.03, c = 92.92, α = 105.02, β = 97.06, γ = 93.66
Resolution (Å)	19.78–2.59 (2.73–2.59)
Total No. of observations	179404 (23917)
No. of unique reflections	45759 (6394)
Completeness (%)	95.0 (90.7)
Multiplicity	3.9 (3.7)
Average $I/\sigma(I)$	7.7 (1.7)
R_{merge}† (%)	15.0 (84.2)
Molecules per asymmetric unit (Z)	4
Matthews coefficient (Å^3 Da^{-1})	2.21

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith observed intensity of reflection hkl and $\langle I(hkl)\rangle$ is the mean intensity over all i measurements.

3. Results and discussion

Recombinant FabG4$^{17–454}$ was successfully purified to homogeneity. The N-terminal residues (1–16) were found to hinder solubility and hence were truncated without altering the activity (Dutta *et al.*, 2011). FabG4$^{17–454}$:NADH complex crystals were obtained from 0.1 *M* MES pH 6.5, 45%(*v*/*v*) polypropylene glycol P400 and diffracted to 2.59 Å resolution. The crystals belonged to space group *P*1, with unit-cell parameters a = 63.07, b = 71.03, c = 92.92 Å, α = 105.02, β = 97.06, γ = 93.66°, which differed from those of the FabG4$^{17–448}$ crystals. The Matthews coefficient (2.21 Å^3 Da^{-1}) confirmed that there were four monomers in the unit cell (Matthews, 1968). The presence of a single fourfold NCS axis was also indicated by the self-rotation function plot at χ = 90° (Fig. 2*a*). The NCS axis makes an ∼30° angle with the crystallographic *y* axis. The self-rotation function at χ = 180° (Fig. 2*b*) also indicated two twofold NCS axes. The highest peak (peak 1; Rf = 508.2) is attributed to the dimeric axis between two monomers, whereas the second peak (peak 2; Rf = 338.7) is attributed to the dimeric axis between two dimers.

The structure was solved by the molecular-replacement method with *MOLREP* (Vagin & Teplyakov, 2010), using a monomer of the apo FabG4 structure (PDB entry 3lls; Seattle Structural Genomics Center for Infectious Disease, unpublished work) as a search model. A promising result with a homotetrameric assembly has been found with a final score of 0.71; the resulting *R* factor was 37.4%. The model was subsequently subjected to rigid-body refinement followed by restrained refinement using *REFMAC*5 (Murshudov *et al.*, 2011). Structural analysis is currently in progress.

This work was carried out using the protein X-ray crystallography facility funded by the Department of Biotechnology, Government of India and housed at the Central Research Facility (CRF) of the Indian Institute of Technology (IIT), Kharagpur. DD acknowledges the Department of Science and Technology, Government of India for

a fellowship. SB thanks IIT Kharagpur for an individual institute fellowship.

References

Beste, D. J., Espasa, M., Bonde, B., Kierzek, A. M., Stewart, G. R. & McFadden, J. (2009). *PLoS One*, **4**, e5349.

Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.

Cohen-Gonsaud, M., Ducasse, S., Hoh, F., Zerbib, D., Labesse, G. & Quemard, A. (2002). *J. Mol. Biol.* **320**, 249–261.

Cohen-Gonsaud, M., Ducasse-Cabanot, S., Quemard, A. & Labesse, G. (2005). *Proteins*, **60**, 392–400.

Dessen, A., Quémard, A., Blanchard, J. S., Jacobs, W. R. & Sacchettini, J. C. (1995). *Science*, **267**, 1638–1641.

Dutta, D., Bhattacharyya, S. & Das, A. K. (2012). *Proteins*, **80**, 1250–1257.

Dutta, D., Bhattacharyya, S., Mukherjee, S., Saha, B. & Das, A. K. (2011). *J. Struct. Biol.* **174**, 147–155.

Evans, P. R. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 114–122. Warrington: Daresbury Laboratory.

Fisher, M., Kroon, J. T., Martindale, W., Stuitje, A. R., Slabas, A. R. & Rafferty, J. B. (2000). *Structure*, **8**, 339–347.

Gao, L.-Y., Laval, F., Lawson, E. H., Groger, R. K., Woodruff, A., Morisaki, J. H., Cox, J. S., Daffe, M. & Brown, E. J. (2003). *Mol. Microbiol.* **49**, 1547–1563.

Gurvitz, A. (2009). *Mol. Genet. Genomics*, **282**, 407–416.

Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.

Krimsky, I. & Racker, E. (1963). *Biochemistry*, **2**, 512–518.

Makinoshima, H. & Glickman, M. S. (2005). *Nature (London)*, **436**, 406–409.

Mao, Q., Duax, W. L. & Umland, T. C. (2007). *Acta Cryst.* F**63**, 106–109.

Marrakchi, H., Ducasse, S., Labesse, G., Montrozier, H., Margeat, E., Emorine, L., Charpentier, X., Daffé, M. & Quémard, A. (2002). *Microbiology*, **148**, 951–960.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Mdluli, K., Slayden, R. A., Zhu, Y., Ramaswamy, S., Pan, X., Mead, D., Crane, D. D., Musser, J. M. & Barry, C. E. III (1998). *Science*, **280**, 1607–1610.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Oppermann, U., Filling, C., Hult, M., Shafqat, N., Wu, X., Lindh, M., Shafqat, J., Nordling, E., Kallberg, Y., Persson, B. & Jörnvall, H. (2003). *Chem. Biol. Interact.* **143–144**, 247–253.

Poncet-Montange, G., Ducasse-Cabanot, S., Quemard, A., Labesse, G. & Cohen-Gonsaud, M. (2007). *Acta Cryst.* D**63**, 923–925.

Price, A. C., Zhang, Y.-M., Rock, C. O. & White, S. W. (2001). *Biochemistry*, **40**, 12772–12781.

Price, A. C., Zhang, Y.-M., Rock, C. O. & White, S. W. (2004). *Structure*, **12**, 417–428.

Sacco, E., Covarrubias, A. S., O'Hare, H. M., Carroll, P., Eynard, N., Jones, T. A., Parish, T., Daffé, M., Bäckbro, K. & Quémard, A. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 14628–14633.

Schaeffer, M. L., Agnihotri, G., Volker, C., Kallender, H., Brennan, P. J. & Lonsdale, J. T. (2001). *J. Biol. Chem.* **276**, 47029–47037.

Sharma, P., Kumar, B., Singhal, N., Katoch, V. M., Venkatesan, K., Chauhan, D. S. & Bisht, D. (2010). *Indian J. Med. Res.* **132**, 400–408.

Subramanian, S., Abendroth, J., Phan, I. Q. H., Olsen, C., Staker, B. L., Napuli, A., Van Voorhis, W. C., Stacy, R. & Myler, P. J. (2011). *Acta Cryst.* F**67**, 1118–1122.

Takayama, K., Wang, C. & Besra, G. S. (2005). *Clin. Microbiol. Rev.* **18**, 81–101.

Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.

Wickramasinghe, S. R., Inglis, K. A., Urch, J. E., Müller, S., van Aalten, D. M. & Fairlamb, A. H. (2006). *Biochem. J.* **393**, 447–457.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

Wong, H. C., Liu, G., Zhang, Y.-M., Rock, C. O. & Zheng, J. (2002). *J. Biol. Chem.* **277**, 15874–15880.

Zaccai, N. R., Carter, L. G., Berrow, N. S., Sainsbury, S., Nettleship, J. E., Walter, T. S., Harlos, K., Owens, R. J., Wilson, K. S., Stuart, D. I. & Esnouf, R. M. (2008). *Proteins*, **70**, 562–567.