**Rupam Biswas, Debajyoti Dutta
and Amit Kumar Das***Department of Biotechnology, Indian Institute of
Technology Kharagpur, Kharagpur 721 302,
IndiaCorrespondence e-mail:
amitk@hijli.iitkgp.ernet.in

Received 30 July 2013

Accepted 20 August 2013

Cloning, overexpression, purification, crystallization and preliminary X-ray diffraction analysis of Rv0241c (HtdX) from *Mycobacterium tuberculosis* H37Rv

Rv0241c (HtdX) is a putative (3*R*)-hydroxyacyl-CoA dehydratase of *Mycobacterium tuberculosis*. The *htdX* gene belongs to a conserved operon and is expressed in mycobacteria in the presence of several fatty-acid synthase II drugs. To elucidate the structure of HtdX, the protein was cloned, overexpressed, purified to homogeneity and crystallized. The protein was crystallized from two conditions: (i) 3 *M* sodium chloride, 0.1 *M* Na HEPES pH 8.0 and (ii) 2.5 *M* sodium chloride, 0.1 *M* Tris-HCl pH 8.5. A complete diffraction data set was collected from crystals from both conditions. The crystal from the first condition diffracted to 2.3 Å resolution and belonged to space group *I*₄₁, with unit-cell parameters *a* = *b* = 61.51, *c* = 143.81 Å. Crystals from the second condition diffracted to 3.1 Å resolution and belonged to space group *P*₄₃₂₁² or *P*₄₁₂₁², with unit-cell parameters *a* = *b* = 63.67, *c* = 140.88 Å. Both crystals contained one molecule in the asymmetric unit.

1. Introduction

Tuberculosis caused by *Mycobacterium tuberculosis* remains a major health problem. It infects millions of people every year and is the second leading cause of death worldwide (World Health Organization, 2012). The appearance of new multidrug-resistant TB (MDR-TB) and extremely multidrug-resistant TB (XDR-TB) has made this problem more complex (World Health Organization, 2010). Hence, new drugs are necessary in order to combat the disease.

M. tuberculosis possesses an extremely lipid-rich and highly impermeable cell wall. The key components of this outer envelope are mycolic acids. Mycolic acids are the product of two fatty-acid synthase systems present in mycobacteria: mammalian-type fatty-acid synthase I (FAS-I) (Smith *et al.*, 2003) and bacterial-type fatty-acid synthase II (FAS-II) (Takayama *et al.*, 2005; White *et al.*, 2005). FAS-I consists of a single enzyme containing all of the catalytic domains required for the fatty-acid elongation cycle (Boehringer *et al.*, 2013; Ciccarelli *et al.*, 2013). On the other hand, the FAS-II pathway employs the discrete enzymes mtFabD (Kremer *et al.*, 2001), KasA/KasB (Gao *et al.*, 2003; Schaeffer *et al.*, 2001), MabA (Marrakchi *et al.*, 2002), HadABC (Sacco *et al.*, 2007), InhA (Dessen *et al.*, 1995) and AcpM (Wong *et al.*, 2002).

Classically, the dehydration step of the FAS-II system is carried out by either a (3*R*)-hydroxyacyl-ACP dehydratase (FabZ) or a (3*R*)-hydroxyacyl-ACP dehydratase/isomerase (FabA). However, in mycobacteria there is no single FabZ or FabA protein in the FAS-II pathway; instead, a gene cluster HadABC consisting of HadA (Rv0635), HadB (Rv0636) and HadC (Rv0637) is required for the dehydratase step (Sacco *et al.*, 2007). Functional analysis based on BLAST results shows that in addition to HadABC the mycobacterial genome encodes genes which can perform a dehydration step. These include Rv0216, Rv0241c, Rv0504c, Rv2499c, Rv3389c, Rv3538 and Rv3542c (Castell *et al.*, 2005; Takayama *et al.*, 2005). Among these, Rv0241c or HtdX was first identified in the mycobacterial membrane-associated fraction (Gu *et al.*, 2003) and was eventually identified in the mRNA pool when cultured in the presence of several FAS-II drugs (Boshoff *et al.*, 2004). Recent work on heterologous expression of HtdX has shown that it can participate in the yeast mitochondrial fatty-acid synthetic pathway (Gurvitz *et al.*, 2009) but can only restore the dehydratase activity in *Escherichia coli* (Sacco *et al.*, 2010). This

R-specific β -hydroxyacyl-CoA dehydratase is hypothesized to be involved in an alternative or intermediate fatty-acid synthetic pathway required for antibiotic resistance (Kikuchi & Kusaka, 1983; Boshoff *et al.*, 2004; Sacco *et al.*, 2010). Based on the genomic arrangement in *Mycobacteria* and STRING functional protein association networks, a putative fatty-acid metabolizing operon was identified showing two genes Rv0241c (HtdX) and Rv0242c (FabG4) could be involved in the pathway (Gurvitz, 2009; Dutta *et al.*, 2011). FabG4, which has been characterized recently, is an NADH-dependent β -ketoacyl CoA reductase (Dutta *et al.*, 2011, 2013). To elucidate the function of Rv0241c (HtdX), we attempted to determine the structure of HtdX. This communication reports the cloning strategy, purification, crystallization, data collection and initial diffraction data analysis of Rv0241c.

2. Materials and methods

2.1. Cloning

The sequence corresponding to the ORF of Rv0241c (UniProt ID O53664) was amplified by polymerase chain reaction (PCR) using *M. tuberculosis* H37Rv genomic DNA as template with the primers 5'-CAGGGATCCGTGACTCAACCCAGCGGCCTGAAG-3' (forward primer with *Bam*HI site) and 5'-GTGGAAGCTTCTATAGACCCCGCACGGTAGCGG-3' (reverse primer with *Hind*III site).

The amplified PCR product was purified by a Qiagen purification kit and cloned into the pQE30 expression vector (Qiagen, USA) which contains six consecutive His codons at the N-terminus of the multiple cloning site (MCS). The recombinant DNA was transformed into *E. coli* SG13009 (pREP4) cells and subsequently selected on ampicillin/kanamycin plates. The positive clones were verified by DNA sequencing. This construct however produced insoluble recombinant protein.

To produce soluble protein, we first predicted the secondary structure of HtdX using *PSIpred* (Jones, 1999) and found that a long N-terminal region (1–25) might not have any rigid secondary structure. Owing to the long unstructured region there is a chance of misfolding when a large amount of full-length protein is expressed in *E. coli*. For this, we designed an N-terminal 28 residues truncated construct using the primers 5'-CAGGGATCCAACCGGACGGT-GACCGTC-3' (forward primer with *Bam*HI site) and 5'-GTGGAAGCTTCTATAGACCCCGCACGGTAGCGG-3' (reverse primer with *Hind*III site) and followed the previous protocol for cloning and transformation. The resulting recombinant protein HtdX^{29–280} was expressed in the soluble fraction. The resulting protein contains Gly and Ser residues between the hexahistidine tag and HtdX^{29–280} owing to the *Bam*HI restriction site which was used for crystallization.

2.2. Overexpression and purification

The positive clone harbouring the desired construct of HtdX (HtdX^{29–280}) was grown in Luria broth supplemented with 100 mg ml^{−1} ampicillin and 25 mg ml^{−1} kanamycin at 310 K for 3 h, after which the OD₆₀₀ reached 0.6, induced with 100 mM IPTG and grown for a further 16 h at 288 K in order to maximize the protein yield. The cells from a 4 l culture were harvested and resuspended in buffer A (10 mM Tris–HCl pH 8.0, 300 mM NaCl, 10 mM imidazole, 5% glycerol) containing 0.1 mM each of leupeptin, pepstatin and aprotinin, and 0.02 mM phenylmethylsulfonyl fluoride (PMSF). The suspension was lysed by ultrasonication on ice and the lysate was centrifuged at 14 000 rev min^{−1} for 40 min. The supernatant was loaded onto Ni-Sepharose High Performance affinity matrix (GE Healthcare Biosciences) pre-equilibrated with buffer A. The column was then washed extensively with buffer A followed by buffer B (10 mM Tris–HCl pH 8.0, 300 mM NaCl, 50 mM imidazole, 5% glycerol) to remove nonspecifically bound contaminants. The protein was finally eluted with buffer C (10 mM Tris–HCl pH 8.0, 300 mM NaCl, 300 mM imidazole, 5% glycerol). The eluted protein was subjected to size-exclusion chromatography using Superdex 200 pre-pgrade matrix in a 16/70C column (GE Healthcare Biosciences) on an ÄKTAprime Plus system (GE Healthcare Biosciences) equilibrated with buffer D (10 mM Tris–HCl pH 8.0, 200 mM NaCl, 10 mM MgCl₂, 5% glycerol). Fractions of 2 ml were collected at a flow rate of 1 ml min^{−1}. The fractions containing the desired protein were pooled together. The protein concentration was estimated by the method of Bradford (1976) and the purity was verified by 12% SDS–PAGE.

2.3. Crystallization

The purified protein was concentrated to 15 mg ml^{−1} in buffer D using a 10 kDa cutoff Vivaspin 20 concentrator (GE Healthcare). Preliminary screening for initial crystallization conditions was performed by the sitting-drop vapour-diffusion method using Crystal Screen and Crystal Screen 2 (Hampton Research, USA) at 298 K by mixing 2 μ l droplets of concentrated protein solution with an equal volume of reservoir solution in a 96-well Corning treated flat-bottom plate (Corning, New York, USA) and equilibrated against 100 μ l

Figure 1
Crystals of HtdX. (a) Crystals obtained from 3 M sodium chloride, 0.1 M Na HEPES pH 8. (b) Crystals obtained from 2.5 M sodium chloride, 0.1 M Tris–HCl pH 8.5.

Table 1
Data collection and processing statistics.

Values in parentheses are for the outermost resolution shell.

	Crystal (i)	Crystal (ii)
Wavelength (Å)	1.5418	1.5418
Space group	<i>I</i> ₄	<i>P</i> ₄ ₁ ₂ ₁ ₂ / <i>P</i> ₄ ₃ ₂ ₁ ₂
Unit-cell parameters (Å)	<i>a</i> = <i>b</i> = 61.51, <i>c</i> = 143.81	<i>a</i> = <i>b</i> = 63.66, <i>c</i> = 140.87
Matthews coefficient (<i>V</i> _M) (Å ³ Da ^{−1})	2.43	2.55
Solvent content (%)	49.39	51.77
Monomers per asymmetric unit (<i>Z</i>)	1	1
Resolution range (Å)	19.49–2.3 (2.43–2.30)	19.36–3.10 (3.27–3.10)
Total No. of observations	179265 (25751)	156123 (22247)
No. of unique reflections	11816 (1729)	5714 (798)
Multiplicity	15.2 (14.9)	27.3 (27.9)
Completeness (%)	99.8 (99.3)	99.8 (99.7)
Average <i>I</i> /σ(<i>I</i>)	30.5 (5.3)	22.9 (8.0)
<i>R</i> _{int} † (%)	9.9 (72.1)	17.6 (61.9)
<i>R</i> _{merge} ‡ (%)	9.6 (69.7)	17.3 (60.8)
CC _{1/2}	99.9 (92.9)	99.9 (99.1)

† $R_{int} = \frac{\sum_{hkl} \{1/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$. ‡ $R_{merge} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$, where $I_i(hkl)$ is the *i*th observation of reflection *hkl*.

reservoir solution. Preliminary crystals were obtained from the condition 4.3 M sodium chloride, 0.1 M Na HEPES pH 7.5, which was further fine screened around the NaCl concentration and the buffer composition (pH range 7–9). Crystals of good diffraction quality appeared in about 4 d from two conditions: (i) 3 M sodium chloride, 0.1 M Na HEPES pH 8.0 (Fig. 1*a*) and (ii) 2.5 M sodium chloride, 0.1 M Tris–HCl pH 8.5 (Fig. 1*b*). A total of 3 mg protein was obtained from 4 l culture.

2.4. Data collection and processing

Prior to data collection, the crystals from both conditions were quick-soaked in a cryoprotectant solution containing 5%(v/v) glycerol in corresponding well solutions and flash-cooled in a nitrogen stream at 93 K. X-ray diffraction data were collected on an in-house Rigaku R-AXIS IV++ image-plate detector using Cu Kα X-rays generated by a Rigaku MicroMax HF007 rotating-anode generator. The crystal-to-detector distance was kept at 140 mm and the crystal was rotated through a total of 360° with 1° rotation per frame. The crystal from condition (i) diffracted to a resolution of 2.3 Å. A total of 360 images were collected and data were processed using *XDS* (Kabsch, 1993) and scaled with *SCALA* (Evans, 2006) in space group *I*4 with a resulting overall *R*_{int} value of 9.9%. Similarly, images were collected for the crystal from condition (ii) which diffracted to a resolution of 3.1 Å. Images were processed as above in space group *P*422 with a resulting overall *R*_{int} value of 17.6%. Screw axes were determined using *POINTLESS* (Evans, 2006) from the *CCP4* interface (Winn *et al.*, 2011). The final statistics of data collection and processing are given in Table 1.

3. Results and discussion

HtdX, a putative (3*R*)-hydroxyacyl-CoA dehydratase of *M. tuberculosis*, was successfully cloned and overexpressed in *E. coli* SG13009 cells. The N-terminally truncated protein (HtdX^{29–280}) was purified to homogeneity with an N-terminal His₆ tag. The molecular weight of HtdX was predicted to be 28 kDa by analysing its sequence and was confirmed by 12% SDS–PAGE. The protein was crystallized in two conditions: (i) 3 M sodium chloride, 0.1 M Na HEPES pH 8.0 and (ii) 2.5 M sodium chloride, 0.1 M Tris–HCl pH 8.5. A crystal from the first condition diffracted to a resolution of 2.3 Å (Fig. 2*a*) and diffraction

data were processed in space group *I*4₁, whereas a crystal from the second condition diffracted to a resolution of 3.1 Å (Fig. 2*b*) and data were processed in space group *P*422 with the most probable screw axes suggested by *POINTLESS* to be *P*₄₁₂₁₂/*P*₄₃₂₁₂. Matthews coefficients for the *I*4₁ and *P*₄₁₂₁₂/*P*₄₃₂₁₂ crystals were determined to be 2.43 and 2.55 Å³ Da^{−1}, respectively, suggesting a single molecule in the asymmetric unit (Matthews, 1968). Although HtdX and FabZs catalyze similar reactions, they show less than 10% sequence identity. Sequence prediction indicates that HtdX has a double hot-dog fold unlike FabZ (Leesong *et al.*, 1996). HtdX shares highest sequence identity (52%) with the dehydratase domain of mycobacterial

(a)

(b)

Figure 2
Diffraction images collected from crystals from (a) condition (i) and (b) condition (ii). For each frame, the oscillation width and the exposure time were 1° and 3 min, respectively.

Table 2

Homologous sequences used for model generation.

PDB entry	Identity (%)	Query cover (%)	Reference
1pn2	47	10	Koski <i>et al.</i> (2004)
1s9c	24	60	Koski <i>et al.</i> (2005)
3kh8	21	55	Wang <i>et al.</i> (2010)
3khp	36	19	Seattle Structural Genomics Center for Infectious Disease (unpublished work)
3oml	33	37	Haataja <i>et al.</i> (2011)

fatty-acid synthase (PDB entry 3zen; Boehringer *et al.*, 2013) with low query coverage (15% sequence overlap). HtdX also shows homology with MaoC-like dehydratases and eukaryotic peroxisomal multifunctional dehydrogenase and hydratase 2 (Table 2). Our initial attempt to solve the structure using molecular replacement employed the generation of several HtdX models using *MODELLER* (Sali & Blundell, 1993). Protein structures used as templates are listed in Table 2. However, molecular replacement failed, yielding solutions with low correlation coefficients and high *R* factors (60%). We assume that the molecular replacement did not work because either the sequence identity or the query coverage is compromised. Currently, preparations of heavy-atom derivatives are under way to obtain phase information either by SIRAS/MIRAS or SAD/MAD.

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. RB thanks CSIR for an individual fellowship. The authors acknowledge Mr Amlan Roychowdhury for his kind assistance in data processing. The genomic DNA of H37Rv is a kind gift from Professor Y. Zhang (Johns Hopkins University, Baltimore, Maryland, USA).

References

- Boehringer, D., Ban, N. & Leibundgut, M. (2013). *J. Mol. Biol.* **425**, 841–849.
- Boshoff, H. I., Myers, T. G., Copp, B. R., McNeil, M. R., Wilson, M. A. & Barry, C. E. (2004). *J. Biol. Chem.* **279**, 40174–40184.
- Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
- Castell, A., Johansson, P., Unge, T., Jones, T. A. & Bäckbro, K. (2005). *Protein Sci.* **14**, 1850–1862.
- Ciccarelli, L., Connell, S. R., Enderle, M., Mills, D. J., Vonck, J. & Grininger, M. (2013). *Structure*, **21**, 1251–1257.
- Dessen, A., Quémard, A., Blanchard, J. S., Jacobs, W. R. & Sacchettini, J. C. (1995). *Science*, **267**, 1638–1641.
- Dutta, D., Bhattacharyya, S., Mukherjee, S., Saha, B. & Das, A. K. (2011). *J. Struct. Biol.* **174**, 147–155.
- Dutta, D., Bhattacharyya, S., Roychowdhury, A., Biswas, R. & Das, A. K. (2013). *Biochem. J.* **450**, 127–139.
- Evans, P. (2006). *Acta Cryst. D* **62**, 72–82.
- 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.
- Gu, S., Chen, J., Dobos, K. M., Bradbury, E. M., Belisle, J. T. & Chen, X. (2003). *Mol. Cell. Proteomics*, **2**, 1284–1296.
- Gurvitz, A. (2009). *Mol. Genet. Genomics*, **282**, 407–416.
- Gurvitz, A., Hiltunen, J. K. & Kastaniotis, A. J. (2009). *J. Bacteriol.* **191**, 2683–2690.
- Haataja, T. J., Koski, M. K., Hiltunen, J. K. & Glumoff, T. (2011). *Biochem. J.* **435**, 771–781.
- Jones, D. T. (1999). *J. Mol. Biol.* **292**, 195–202.
- Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
- Kikuchi, S. & Kusaka, T. (1983). *J. Biochem.* **94**, 1045–1051.
- Koski, M. K., Haapalainen, A. M., Hiltunen, J. K. & Glumoff, T. (2004). *J. Biol. Chem.* **279**, 24666–24672.
- Koski, M. K., Haapalainen, A. M., Hiltunen, J. K. & Glumoff, T. (2005). *J. Mol. Biol.* **345**, 1157–1169.
- Kremer, L., Nampoothiri, K. M., Lesjean, S., Dover, L. G., Graham, S., Betts, J., Brennan, P. J., Minnikin, D. E., Loch, C. & Besra, G. S. (2001). *J. Biol. Chem.* **276**, 27967–27974.
- Leesong, M., Henderson, B. S., Gillig, J. R., Schwab, J. M. & Smith, J. L. (1996). *Structure*, **4**, 253–264.
- 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.
- 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.
- Sacco, E., Slama, N., Bäckbro, K., Parish, T., Laval, F., Daffé, M., Eynard, N. & Quémard, A. (2010). *J. Bacteriol.* **192**, 4037–4044.
- Sali, A. & Blundell, T. L. (1993). *J. Mol. Biol.* **234**, 779–815.
- Schaeffer, M. L., Agnihotri, G., Volker, C., Kallender, H., Brennan, P. J. & Lonsdale, J. T. (2001). *J. Biol. Chem.* **276**, 47029–47037.
- Smith, S., Witkowski, A. & Joshi, A. K. (2003). *Prog. Lipid Res.* **42**, 289–317.
- Takayama, K., Wang, C. & Besra, G. S. (2005). *Clin. Microbiol. Rev.* **18**, 81–101.
- Wang, H., Guo, J., Pang, H. & Zhang, X. (2010). *Acta Cryst. F* **66**, 272–274.
- White, S. W., Zheng, J., Zhang, Y.-M. & Rock, (2005). *Annu. Rev. Biochem.* **74**, 791–831.
- 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.
- World Health Organization (2010). *Multidrug and Extensively Drug-Resistant TB (M/XDR-TB): 2010 Global Report on Surveillance and Response*. Geneva: World Health Organization. <http://www.who.int/tb/publications/2010/978924599191/en/>.
- World Health Organization (2012). *Global Tuberculosis Report 2012*. Geneva: World Health Organization. http://www.who.int/tb/publications/global_report/en/.