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

Chunying Niu, Jiang Yin,* Maia M. Cherney and Michael N. G. James*

Protein Structure and Function Group, Department of Biochemistry, School of Molecular and Systems Medicine, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Alberta T6G 2H7, Canada

Correspondence e-mail: jyin@ualberta.ca, michael.james@ualberta.ca

Received 26 June 2011
Accepted 19 September 2011

Expression, purification and preliminary crystallographic analysis of Rv2247, the β subunit of acyl-CoA carboxylase (ACCD6) from *Mycobacterium tuberculosis*

Mycobacterium tuberculosis (*Mtb*) acyl-CoA carboxylase is involved in the biosynthesis of mycolic acids, which are a key component of the bacillus cell wall. The *Mtb* genome encodes six acyl-CoA carboxylase β subunits (ACCD1–6), three of which (ACCD4–6) are essential for survival of the pathogen on minimal medium. *Mtb* ACCD6 has been expressed, purified and crystallized. The two forms of *Mtb* ACCD6 crystals belonged to space groups $P4_12_12$ and $P2_12_12_1$ and diffracted to 2.9 and 2.5 Å resolution, respectively, at a synchrotron-radiation source.

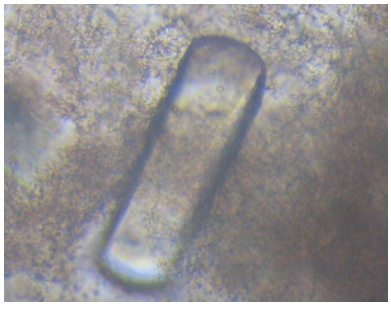
1. Introduction

Tuberculosis (TB), caused by *Mycobacterium tuberculosis* (*Mtb*), is one of the most prevalent bacterial infections in humans. More than one third of the world's population is currently infected by *Mtb* and close to two million TB-associated deaths are reported annually (World Health Organization, 2006). Because of the ability of *Mtb* to persist in infected hosts for prolonged periods of time, treatment of TB requires the administration of a multi-drug regimen for six to nine months. The lengthy course of anti-TB treatment sometimes leads to noncompliance in patients, who often relapse with TB caused by *Mtb* strains with enhanced drug resistance (Bradford & Daley, 1998).

In recent years, there has been a rising trend in the occurrence of multidrug-resistant TB (MDR-TB) and extensively drug-resistant TB (XDR-TB); these strains are characterized as highly virulent *Mtb* strains that are nonresponsive to most front-line anti-TB drugs and to many second-line therapeutic compounds. The death rate of patients contracting MDR-TB or XDR-TB is much higher than that of patients infected by regular *Mtb*, especially in patients with compromised immune systems. The appearance of these newly emerging *Mtb* strains has severely weakened the effectiveness of the present anti-TB therapy in developing countries, many of which are already facing mounting numbers of cases of HIV/AIDS patients in their populations (World Health Organization, 2006).

As a member of the TB structural genomics consortium, we are actively pursuing the structural determination of *Mtb* proteins with critical biological functions. Structural biology has a successful track record as a drug-development tool. It is our hope that the structures of these essential *Mtb* proteins will shed light on the biological processes in which these *Mtb* proteins participate and may one day lead to new anti-TB therapies.

In this paper, we report the molecular cloning, protein purification and preliminary X-ray crystallographic analysis of Rv2247, a β subunit of acyl-CoA carboxylase (ACCD6). Carboxylases are involved in the biosynthesis of fatty acids that are key components of the *Mtb* cell wall (Cronan & Waldrop, 2002; Takayama *et al.*, 2005; Tong, 2005). *Mtb* possesses six acyl-CoA carboxylase β subunits (ACCD1–6). Three of them, ACCD4–6, are essential genes for the survival of *Mtb* on minimal medium (Fig. 1; Sassetti *et al.*, 2001). The *Mtb* cell wall is heavily decorated by *Mtb*-specific fatty acids (*e.g.* mycolic acids) which are likely to contribute to the ability of *Mtb* to evade the host immune response, resist antibiotics and persist in immune cells such as macrophages. The β subunits carry out the only substrate-specific step during fatty-acid chain elongation catalyzed by acyl-CoA



carboxylases. Therefore, the presence of multiple ACCD genes in the *Mtb* genome is thought to be linked to the wide variety of lipids found in the cell wall of *Mtb* (Cole *et al.*, 1998; Holton *et al.*, 2006; Lin *et al.*, 2006; Takayama *et al.*, 2005). *Mtb* ACCD6 has been shown to be able to use both propionyl-CoA and acetyl-CoA as substrates, with a threefold higher efficiency in carboxylation of the latter (Daniel *et al.*, 2007). Based on ACCD6 being part of the fatty-acid synthesis (FAS) II locus, it was hypothesized that *Mtb* ACCD6 may produce the malonyl-CoA molecules that are used to synthesize meromycolic acid by the catalytic components of FAS II. Kurth and coworkers reported that ACCD6 is essential in *M. smegmatis*, a common laboratory relative of *Mtb* (Kurth *et al.*, 2009). Consequently, ACCDs are attractive targets for the design of novel anti-TB therapeutics. An inhibitor, NCI-172033, has been identified to be effective against ACCD6 at 8 µ*M* *in vitro* and has been shown to exhibit rather wide-spectrum anti-mycobacterial activity (Kurth *et al.*, 2009). At low concentrations (∼25 µ*M*) this compound is bacteriostatic, whereas at 100 µ*M* it is clearly bacteriocidal. While the crystal structure of *Mtb* ACCD5 is known (Lin *et al.*, 2006), there is currently no deposition in the Protein Data Bank (PDB) for the molecular structure of *Mtb* ACCD6, which is a promising novel anti-TB target.

2. Materials and methods

2.1. Cloning, expression and purification of recombinant *Mtb* ACCD6

The Rv2247 gene was amplified by the polymerase chain reaction (PCR) using a bacterial artificial chromosome (BAC) genomic library received from L'Institut Pasteur (Brosch *et al.*, 1998; Gordon *et al.*, 1999). The sequences of the PCR primers (Integrated DNA Technologies) used in directional cloning of the amplified product into the

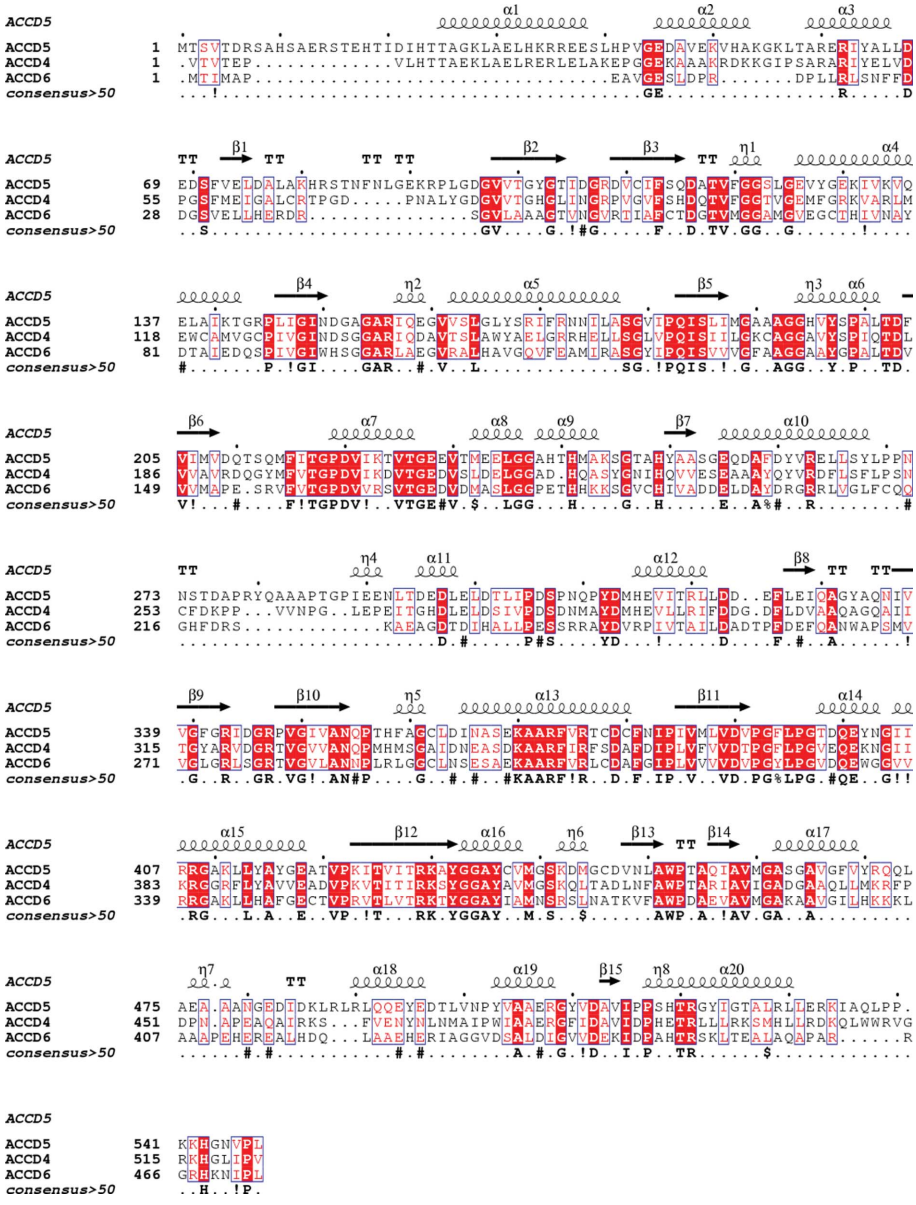


Figure 1
Sequence alignment of the *Mtb* ACCD4–6 β-carboxylases. The secondary structures corresponding to *Mtb* ACCD5 (PDB entry 2a7s; Lin *et al.*, 2006) are drawn at the top of the sequence alignment. Helices, β-strands and turns are represented by coils, arrows and 'T's. Conserved residues are highlighted with a red background, whereas residues with similar chemical properties are coloured red. Sequence alignment was performed using the program *ClustalW* (Thompson *et al.*, 1994) and the figure was generated with the program *ESPript* (Gouet *et al.*, 1999).

Gateway cloning system (Invitrogen) are 2247f, ggg aca agt ttg tac aaa aaa gca ggc tcg ggc agc **gaa aat ctg tat ttt cag ggc agc ggc** ATG ACA ATC ATG GCC CCC GAG G (where bold letters indicate the recombinant tobaco etch virus protease cleavage site and the upper case letters represent the Rv2247-coding region), and 2247r, gg gac cac ttt gta caa gaa agc tgg gtc ctc gag TTA CAG CGG GAT GTT CTT GTG GCG.

The full open reading frame encoding residues 1–473 of *Mtb* Rv2247 was gel-purified and inserted into the pDONR221 plasmid (Invitrogen) to generate an entry clone plasmid pDONR2247. Subsequently, the coding region of Rv2247 in pDONR2247 was cloned into the expression plasmid pDEST-15 (Invitrogen) with an amino-terminal glutathione *S*-transferase (GST) fusion partner. The coding sequence of Rv2247 in this expression plasmid (pGST-2247) was verified by DNA-sequencing analysis (DNA Core Facility, Department of Biochemistry, University of Alberta, Canada).

BL21 arabinose-inducible *Escherichia coli* cells transformed with pGST-2247 were incubated at 310 K in a shaker until the OD reached 0.6. 0.2% D-arabinose was then added to the culture to induce the production of recombinant Rv2247.

For recombinant Rv2247 protein purification, the cell pellet from a 2 l culture was collected *via* centrifugation at ∼12 000*g* (max.) for 30 min and resuspended in 40 ml lysis buffer consisting of 20 m*M* Tris pH 7.5, Complete protease-inhibitor cocktail (Roche), 100 m*M* NaCl, 50 m*M* KCl, 1 m*M* DTT, 1 m*M* PMSF. The cells were lysed by one round of freeze–thawing followed by ultrasonication. The cell lysate was cleared by centrifugation at 30 000*g* (max.) for 30 min and the supernatant was loaded onto a 5 ml GSTrap glutathione Sepharose column (GE Healthcare) pre-equilibrated with buffer *A* (50 m*M* Tris–HCl pH 7.5, 150 m*M* NaCl, 5 m*M* *β*-mercaptoethanol). Recombinant GST-Rv2247 fusion protein was eluted using buffer *B* [50 m*M* Tris–HCl pH 8.5, 10 m*M* glutathione (GSH), 0.02% NaN_3]. Recombinant tobacco etch virus protease (AcTEV, Invitrogen) was used to proteolytically release the GST tag from the GST-Rv2247 fusion protein. After overnight dialysis to remove GSH, the digested protein mixture was again loaded onto a GSTrap column. The flowthrough fractions that contained liberated Rv2247 were concentrated and exchanged into buffer *C* (10 m*M* Tris–HCl pH 7.5, 50 m*M* NaCl, 50 m*M* KCl, 5 m*M* $MgCl_2$, 5 m*M* *β*-mercaptoethanol, 0.5 m*M* DTT, 0.02% NaN_3) before being loaded onto a size-exclusion column (Sephadex G-75). The eluted fractions were sampled for the presence of purified Rv2247 and the peak fractions were pooled and concentrated to ∼5 mg ml^{-1} using an Amicon Ultra filtration unit (30 kDa cutoff, Millipore); the purity of the protein (>95%) was confirmed by 15% SDS–PAGE (data not shown).

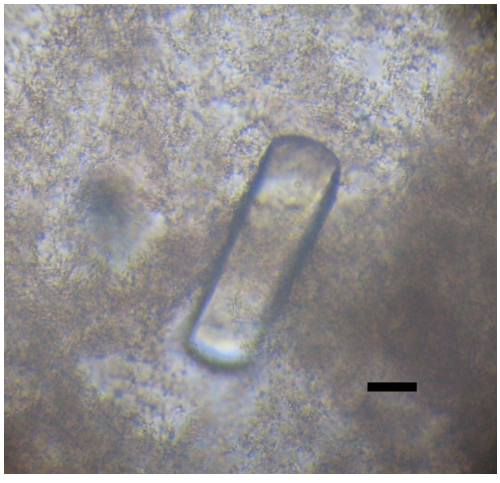


Figure 2
Morphology of form I Rv2247 crystals. The approximate dimensions of Rv2247 crystals of diffraction quality grown in 2.8 *M* sodium formate pH 7.5 over 2–3 d were 50 × 100 × 200 µm. The black scale bar indicates 50 µm.

Table 1
Crystallographic statistics of native Rv2247 data sets.

Values in parentheses are for the highest resolution shell.

Crystal	Form I	Form II
Space group	$P4_12_12$	$P2_12_12_1$
Unit-cell parameters (Å)		
a	82.19	75.68
b	82.19	88.24
c	159.56	154.78
No. of molecules in the asymmetric unit	1	2
Matthews coefficient (Å^3 Da^{-1})	2.69	2.57
Solvent content (%)	54.24	52.23
Data collection		
Temperature (K)	100	100
Detector	MAR 325 CCD	MAR 325 CCD
Wavelength (Å)	0.98	0.98
Resolution (Å)	40.0–2.9 (3.0–2.9)	40.0–2.51 (2.58–2.51)
Unique reflections	20014	35619
Multiplicity	3.4 (2.4)	6.7 (3.0)
$\langle I/\sigma(I)\rangle$	8.1 (2.6)	14.0 (2.1)
Completeness (%)	81.0 (26.0)	90.4 (62.9)
$R_{\rm merge}$† (%)	8.5 (33.8)	9.3 (80.7)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the *i*th observation of reflection *hkl* and $\langle I(hkl)\rangle$ is the weighted average intensity for all observations *i* of reflection *hkl*.

2.2. Crystallization

Crystallization conditions were manually screened using the hanging-drop method and commercial sparse-matrix-based solution sets (Index HT, Crystal Screen and Crystal Screen 2; Hampton Research). Two different crystal forms of *Mtb* ACCD6 of similar diffraction quality were obtained by mixing equal volumes (0.5–1 µl) of ACCD6 solution (20 mg ml^{-1}) and reservoir solution and letting the mixture equilibrate over 1 ml reservoir solution at 297 K. Form I crystals were obtained using the reservoir solution 2.8 *M* sodium formate pH 7.5 and form II crystals were obtained using 0.1 *M* Tris pH 8.5, 0.2 *M* ammonium sulfate, 23% PEG 3350. Both of these crystals grew to dimensions of ∼50 × 100 × 200 µm in 2–3 d (Fig. 2).

2.3. Data collection

As sodium formate at concentrations of 2.8 *M* and higher is a proven cryoprotectant, form I Rv2247 crystals were flash-cooled by directly submerging looped crystals in liquid nitrogen, whereas form II crystals were cooled in liquid nitrogen after passing the crystals through a drop of reservoir solution supplemented with 30% glycerol as a cryoprotectant. Native X-ray diffraction data sets were collected on a Rigaku R-AXIS IV^{++} rotating-anode X-ray generator and at the Canadian Light Source (beamline CMCF1). Raw data were indexed, integrated and scaled using the *HKL*-2000 program suite (Otwinowski & Minor, 1997). The cutoff resolution was chosen on the basis of the signal-to-noise ratio (>2.0) and the value of $R_{\rm merge}$ (<1.0).

3. Results

We have cloned, expressed, purified and crystallized *Mtb* Rv2247, a *β* subunit of acyl-CoA carboxyltransferase (ACCD6). X-ray diffraction data were collected both at the home source and on a synchrotron beamline. Form I and II crystals diffracted to 2.9 and 2.5 Å resolution, respectively, at the synchrotron-radiation source (Fig. 3).

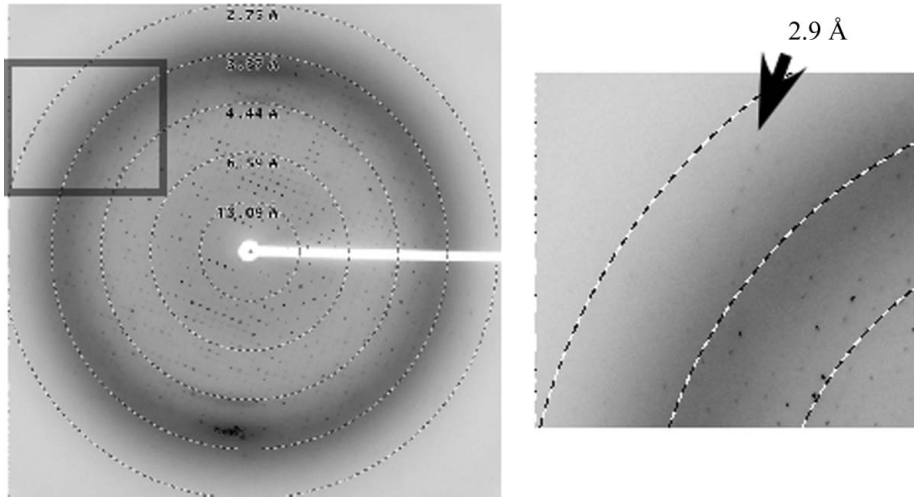


Figure 3
A representative frame of the form I Rv2247 data set showing the diffraction quality. The numbers indicate resolution in Å.

Preliminary X-ray crystallographic analysis (Table 1) indicated that the form I crystals belonged to space group $P4_12_12$, with unit-cell parameters a = 82.19, b = 82.19, c = 159.56 Å. As the calculated molecular weight of ACCD6 is 50.1 kDa, we concluded that there is one molecule of ACCD6 in the asymmetric unit of the crystal lattice. This gives a Matthews coefficient of 2.69 $Å^3\,Da^{-1}$ and a corresponding solvent content of 54.24%. The form II crystals belonged to space group $P2_12_12_1$, with unit-cell parameters a = 75.68, b = 88.24, c = 154.78 Å and two molecules of ACCD6 in the asymmetric unit. The calculated Matthews coefficient and solvent content for crystal form II are 2.57 $Å^3\,Da^{-1}$ and 52.23%, respectively. Structural solution of this essential *Mtb* enzyme is currently under way.

X-ray diffraction data were collected on beamline CMCF1 at the Canadian Light Source and on beamline 9 at Stanford Synchrotron Radiation Lightsource (SSRL). Portions of this research were carried out at SSRL, which is a national user facility operated by Stanford University on behalf of the US Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program and the National Institute of General Medical Sciences. Part of the research described in this paper was performed at the Canadian Light Source, which is supported by NSERC, NRC, CIHR and the University of Saskatchewan. Research in the laboratory of MNGJ is funded by the Alberta Heritage Foundation for Medical Research (AHFMR) and the Canadian Institute of Health Research (CIHR). We thank Sheraz Khan for his excellent technical assistance.

References

Bradford, W. Z. & Daley, C. L. (1998). *Infect. Dis. Clin. North Am.* **12**, 157–172.
Brosch, R., Gordon, S. V., Billault, A., Garnier, T., Eiglmeier, K., Soravito, C., Barrell, B. G. & Cole, S. T. (1998). *Infect. Immun.* **66**, 2221–2229.
Cole, S. T. *et al.* (1998). *Nature (London)*, **393**, 537–544.
Cronan, J. E. & Waldrop, G. L. (2002). *Prog. Lipid Res.* **41**, 407–435.
Daniel, J., Oh, T.-J., Lee, C.-M. & Kolattukudy, P. E. (2007). *J. Bacteriol.* **189**, 911–917.
Gordon, S. V., Brosch, R., Billault, A., Garnier, T., Eiglmeier, K. & Cole, S. T. (1999). *Mol. Microbiol.* **32**, 643–655.
Gouet, P., Courcelle, E., Stuart, D. I. & Métoz, F. (1999). *Bioinformatics*, **15**, 305–308.
Holton, S. J., King-Scott, S., Nasser Eddine, A., Kaufmann, S. H. & Wilmanns, M. (2006). *FEBS Lett.* **580**, 6898–6902.
Kurth, D. G., Gago, G. M., de la Iglesia, A., Bazet Lyonnet, B., Lin, T.-W., Morbidoni, H. R., Tsai, S.-C. & Gramajo, H. (2009). *Microbiology*, **155**, 2664–2675.
Lin, T.-W., Melgar, M. M., Kurth, D., Swamidass, S. J., Purdon, J., Tseng, T., Gago, G., Baldi, P., Gramajo, H. & Tsai, S.-C. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 3072–3077.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Sassetti, C. M., Boyd, D. H. & Rubin, E. J. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 12712–12717.
Takayama, K., Wang, C. & Besra, G. S. (2005). *Clin. Microbiol. Rev.* **18**, 81–101.
Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
Tong, L. (2005). *Cell. Mol. Life Sci.* **62**, 1784–1803.
World Health Organization (2006). *Tuberculosis*. http://www.who.int/tb/en.