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

Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of the small subunit of isopropylmalate isomerase (Rv2987c) from *Mycobacterium tuberculosis*

Manikandan Karuppasamy,[a] Arie Geerlof,[a] Linda Schuldt,[a] Christoph Mueller-Dieckmann[b] and Manfred S. Weiss[a]*

[a]EMBL Hamburg Outstation, c/o DESY, Notkestrasse 85, D-22603 Hamburg, Germany, and [b]ESRF, 6 Rue Jules Horowitz, F-38043 Grenoble CEDEX 09, France

Correspondence e-mail: msweiss@embl-hamburg.de

Received 6 October 2008
Accepted 13 December 2008

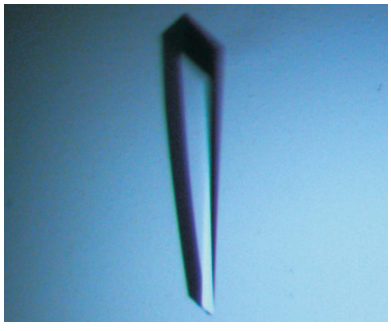
Two C-terminally truncated variants of the small subunit of *Mycobacterium tuberculosis* isopropylmalate isomerase (Rv2987c; LeuD), LeuD_1-156 and LeuD_1-168, have been cloned, heterologously expressed in *Escherichia coli*, purified using standard chromatographic techniques and crystallized. The crystals of LeuD_1-156 belonged to the hexagonal system (space group $P6_122$ or $P6_522$) with up to four subunits in the asymmetric unit, whereas the crystals of LeuD_1-168 belonged to the monoclinic system (space group $P2_1$) with two subunits in the asymmetric unit. Both crystals diffracted X-rays to beyond 2.0 Å resolution and were suitable for further crystallographic analysis.

1. Introduction

Tuberculosis (TB) is an airborne global disease which is mainly caused by the pathogenic bacterium *Mycobacterium tuberculosis* (*Mtb*). Although TB can be effectively cured using antibiotics, every year around 2–3 million people die from TB worldwide. Since some *Mtb* strains have become resistant towards all known antibiotics (Smith *et al.*, 2000; Doherty, 2004), new and faster approaches need to be developed to identify potential new drug targets. In this respect, a promising strategy appears to be to identify biochemical pathways that are crucial for the survival of the pathogen. If such a pathway turns out to be absent in the corresponding host, the proteins involved in the pathway constitute prime potential drug targets. The biosynthetic pathway of the amino acid leucine appears to be a good candidate pathway, since it has been shown by both *in vivo* and *in vitro* studies that leucine-auxotrophic *Mtb* strains are not able to replicate inside their host cells (McAdam *et al.*, 1995; Bange *et al.*, 1996; Guleria *et al.*, 1996; Hondalus *et al.*, 2000). These strains are no longer virulent, but become so when they are complemented with the corresponding gene ectopically. Furthermore, leucine is one of the essential amino acids for humans and other mammals. Consequently, the enzymes of this pathway are to be considered as potential candidates for the design of new drugs against TB.

The complete leucine-biosynthesis pathway of *Mtb* starts with the compound 2-oxoisovalerate and consists of five consecutive reactions catalysed by four different enzymes. The respective enzymes are 2-isopropylmalate synthase (LeuA), the large and small subunit of 3-isopropylmalate isomerase (LeuC and LeuD), 3-isopropylmalate dehydrogenase (LeuB) and branched-chain amino-acid transferase (IlvE). To date, three-dimensional structures have been determined for the leucine-biosynthesis pathway enzymes LeuA (Koon *et al.*, 2004) and LeuB (Singh *et al.*, 2005) from *Mtb*.

The second reaction of the leucine-biosynthesis pathway is the stereospecific conversion of α-isopropylmalate to β-isopropylmalate. This reaction is catalysed by the enzyme isopropylmalate isomerase, which exists in the form of two subunits: a large subunit (LeuC; Rv2988c; 473 amino-acid residues; molecular mass 50.2 kDa) and a small subunit (LeuD; Rv2987c; 198 amino-acid residues; molecular mass 21.8 kDa).

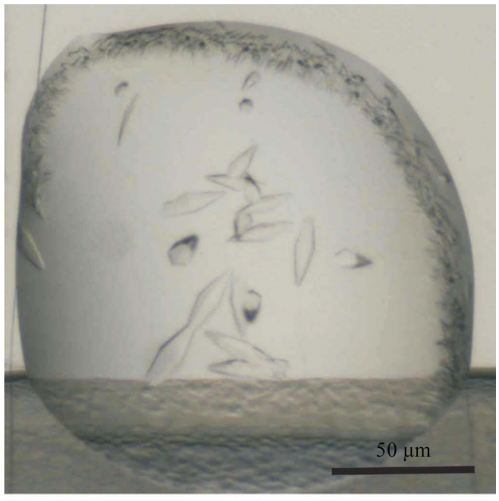


All attempts to crystallize full-length LeuC, full-length LeuD or the entire LeuC/D complex have so far been unsuccessful. Since the small subunit LeuD is predicted to have a flexible C-terminal end which may hinder crystallization, two C-terminal deletion variants were designed. The two variants are termed LeuD_1-156 and LeuD_1-168 as they consist of the N-terminal 156 and 168 amino acids of LeuD, respectively. In this report, we describe the cloning, expression, purification and preliminary crystallographic analysis of these two C-terminally truncated variants of the small subunit of isopropylmalate isomerase.

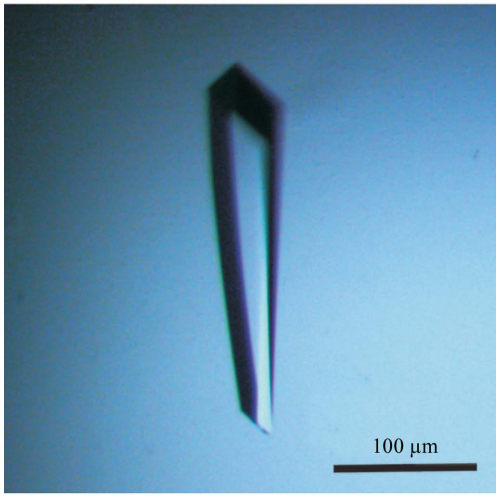
2. Experimental methods

2.1. Cloning

Genomic DNA of the bacterial strain *Mtb* H37Rv was used as a template for the polymerase chain reaction. The following oligonucleotides were used as primers: 5′-AAAA**TCATGA***GC*GAAGCCTTTCACACCCACTCTGGTATTGGCGTGCCG-3′ (forward primer for both variants), 5′-AAAA**CTCGAG**TTAGTCAATCTTTTAGTCAATCTTGAACGGCAGCACCACTGTTGC-3′ (reverse primer for the variant LeuD_1-156) and 5′-AAAA**CTCGAG**TTAGTCGAGACCTTCAAGCAGCCGCCACGC-3′ (reverse primer for the variant LeuD_1-168). In the forward primer, the GC dinucleotide (italicized) was introduced in order to keep the gene in frame. This leads to an additional Ser as the second amino acid of the expressed protein. The restriction sites for the enzymes 5′-*Bsp*HI (forward primer) and 3′-*Xho*I (reverse primers) are marked in bold. The PCR-amplified and doubly digested gene products were ligated into the pETM-11 expression vector, which was linearized by cutting with the restriction enzymes 5′-*Nco*I and 3′-*Xho*I. The pETM-11 vector harbours the sequence for an N-terminal His_6 tag followed by a TEV protease cleavage site. The final purified protein thus contains the dipeptide Gly-Ala at the N-terminus preceding the N-terminal Met. The correctness of the cloned constructs was confirmed by DNA sequencing.



(*a*)



(*b*)

Figure 1
(*a*) Bipyramidal crystals of the variant LeuD_1-156, (*b*) single crystal of the variant LeuD_1-168.

2.2. Expression and purification

The recombinant plasmid was transformed into the expression strain *Escherichia coli* BL21(DE3) RP. Cells from a 10 ml overnight culture were used to inoculate 1 l ZYP-5052 auto-induction medium (Studier, 2005) containing kanamycin (50 µg ml^{-1}) and chloramphenicol (34 µg ml^{-1}). Cultivation was carried out at 310 K and 220 rev min^{-1} until an $OD_{600\,nm}$ of 0.6–0.7 was reached. Subsequently, the temperature was lowered to 293 K for protein expression. Cells were grown for a further 20 h at 293 K and then harvested by centrifugation. The cell pellet was stored at 253 K until further use. 1 g of cells was resuspended in 10 ml lysis buffer [20 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 20 m*M* imidazole, 5%(*v*/*v*) glycerol and 2 m*M* *β*-mercaptoethanol (*β*-ME)] containing one Complete EDTA-free protease-inhibitor cocktail tablet (Roche) per 25 ml and lysed by sonication for 15 min using 0.3 s pulses at 277 K with a 2 min interval after every 5 min of sonication. The lysates were clarified by centrifugation (39 000*g* for 60 min at 277 K) and after discarding the pellet the supernatant was stored on ice. The supernatant was sterile-filtered through a 0.45 µm membrane and then loaded onto a 3 ml Ni–NTA column (Qiagen) pre-equilibrated with lysis buffer. The column with the protein bound was washed in two steps. Initially, three column volumes of high-salt wash buffer [20 m*M* Tris–HCl pH 8.0, 300 m*M* NaCl, 20 m*M* imidazole, 5%(*v*/*v*) glycerol and 2 m*M* *β*-ME] were used for washing. One column volume of imidazole wash buffer [20 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 30 m*M* imidazole, 5%(*v*/*v*) glycerol and 2 m*M* *β*-ME] was then used to remove nonspecifically bound proteins. Finally, the protein was eluted with one column volume of elution buffer [20 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 300 m*M* imidazole, 5%(*v*/*v*) glycerol and 2 m*M* *β*-ME]. All these purification steps were carried out at 277 K. The N-terminal His_6 tag was removed from the protein by incubating the eluted protein with 1 ml TEV protease solution [1 mg ml^{-1} TEV protease in buffer containing 50 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 0.01%(*v*/*v*) 1-thioglycerol and 50%(*v*/*v*) glycerol] overnight (at 277 K) and simultaneously dialysing against dialysis buffer [20 m*M* Tris–HCl pH 8.0, 50 m*M* NaCl, 5%(*v*/*v*) glycerol and 2 m*M* *β*-ME]. Undigested protein and TEV protease were separated from the digested protein by loading the overnight dialysed sample onto a 0.8 ml Ni–NTA column pre-equilibrated with dialysis buffer. The flow-through, which only contains the digested protein, was collected. The concentrated protein was further subjected to size-exclusion chromatography using a HiLoad 16/60 Superdex 75 (GE Healthcare) column. Prior to loading the protein sample, the gel-filtration column was equilibrated with dialysis buffer. The protein was eluted from the column and collected in 2 ml fractions. Protein-containing fractions were analysed by SDS–PAGE and the fractions containing LeuD

were pooled, concentrated, flash-frozen in liquid N_2 and stored at 193 K.

2.3. Crystallization

The pure protein (8 mg ml^{-1}) in a buffer solution consisting of 20 m*M* Tris–HCl pH 8.0, 50 m*M* NaCl, 5%(*v*/*v*) glycerol and 2 m*M* β-ME was used to set up initial crystallization screening experiments. The sitting-drop vapour-diffusion method was used with a drop size of 600 nl (300 nl protein solution and 300 nl reservoir solution) and 50 µl reservoir solution. Screening with various commercially available kits was carried out at the High Throughput Crystallization Facility at EMBL Hamburg (Mueller-Dieckmann, 2006). Bipyramidal crystals of LeuD_1-156 (Fig. 1*a*) appeared after 3 d in two conditions, both of which contained sodium citrate as the precipitant. The refined crystallization conditions were 1 *M* sodium citrate tribasic, 0.1 *M* sodium cacodylate pH 6.5 (condition 1) and 1 *M* sodium citrate tribasic, 0.1 *M* bis-tris propane pH 7.0 (condition 2). Rectangular-shaped crystals of LeuD_1-168 (Fig. 1*b*) also appeared after 3 d in 0.2 *M* lithium sulfate, 0.1 *M* Tris–HCl pH 8.5 and 20%(*w*/*v*) PEG 3350.

2.4. Diffraction data collection and processing

A few single crystals of LeuD_1-156 were harvested using a suitably sized nylon loop and soaked in a cryoprotectant solution containing the mother liquor plus 20%(*v*/*v*) ethylene glycol for 10–15 s. The crystals were then flash-cooled and stored at 100 K. Similarly, single crystals of LeuD_1-168 were harvested and flash-cooled in mother-liquor solution containing 15%(*v*/*v*) glycerol as a cryoprotectant and stored at 100 K. Diffraction data collection from a crystal of LeuD_1-156 grown from condition 2 was carried out on beamline ID29 at the ESRF synchrotron (Grenoble, France). Beamline ID29 is equipped with an ADSC Q315r CCD detector. Diffraction data collection from a crystal of LeuD_1-168 was carried out on beamline PXIII at the SLS synchrotron (Villigen, Switzerland). This beamline is equipped with a MAR Mosaic CCD 225 detector. Indexing and integration of the data was performed using the program *DENZO* (Otwinowski & Minor, 1997) followed by scaling and merging using *SCALEPACK* (Otwinowski & Minor, 1997). The *R* factors $R_{\rm r.i.m.}$ (redundancy-independent merging *R* factor) as well as $R_{\rm p.i.m.}$ (precision-indicating merging *R* factor; Weiss, 2001) were calculated using the program *RMERGE* (available from http://www.embl-hamburg.de/~msweiss/projects/msw_qual.html or from MSW upon request). The data-collection parameters and processing statistics are summarized in Table 1. Intensities were converted to structure-factor amplitudes using the program *TRUNCATE* (French & Wilson, 1978; Collaborative Computational Project, Number 4, 1994). The optical resolution of the data sets was calculated with *SFCHECK* (Vaguine *et al.*, 1999) and the self-rotation functions were computed using the program *MOLREP* (Collaborative Computational Project, Number 4, 1994) based on structure-factor amplitudes to a maximum resolution of 4.0 Å.

3. Results and discussion

3.1. Yield and purity of the sample

From 8 g of wet cell pellet, nearly 10 mg of pure protein could be purified for both variants. The purity of the sample was higher than 95% as judged by SDS–PAGE. Both protein variants LeuD_1-156 and LeuD_1-168 eluted as a monomer from the gel-filtration chromatography column and were soluble up to a concentration of 20 mg ml^{-1}. Nevertheless, a native PAGE shows two clear bands for both LeuD variants (data not shown), which may indicate the presence of higher order oligomers.

Table 1
Data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

	LeuD_1-156	LeuD_1-168
Beamline	ID29, ESRF	PXIII, SLS
No. of crystals	1	1
Wavelength (Å)	1.0724	1.0000
Crystal-to-detector distance (mm)	274.29	90.0
Rotation range per image (°)	0.25	0.40
Total rotation range (°)	50	180
Resolution range (Å)	99.0–2.00 (2.05–2.00)	99.0–1.36 (1.38–1.36)
Unit-cell parameters (Å, °)	$a = b = 64.62$, $c = 419.82$	$a = 41.59$, $b = 75.89$, $c = 49.04$, $\beta = 103.0$
Space group	$P6_122$ or $P6_522$	$P2_1$
Mosaicity (°)	0.35	0.62
Total No. of reflections	190742	219756
Unique reflections	36296	63604
Redundancy	5.3 (3.5)	3.5 (3.3)
Completeness (%)	98.1 (92.3)	99.9 (100)
$I/\sigma(I)$	21.8 (2.2)	16.5 (2.1)
$R_{\rm merge}$ (%)	6.7 (43.6)	7.0 (64.8)
$R_{\rm r.i.m.}$† (%)	7.5 (50.3)	8.3 (77.5)
$R_{\rm p.i.m.}$† (%)	3.3 (24.1)	4.4 (41.8)
Overall *B* factor from Wilson plot (Å^2)	52.7	21.4
Optical resolution‡ (Å)	1.69	1.25

† The redundancy-independent merging *R* factor $R_{\rm r.i.m} = 100\sum_{hkl}[N/(N-1)]\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, with *N* being the number of times a given reflection *hkl* was observed; the precision-indicating merging *R* factor $R_{\rm p.i.m} = 100\sum_{hkl}[1/(N-1)]\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$ (Weiss, 2001). ‡ Defined as in Vaguine *et al.* (1999).

3.2. Data collection and analysis of the diffraction data

A crystal of LeuD_1-156 which had been flash-cooled and stored for transport to the synchrotron was mounted onto the goniostat in a cold nitrogen stream at 100 K for data collection. The crystal diffracted X-rays to about 2.0 Å resolution. A total of 200 images spanning a 50° rotation of the crystal were collected. An oscillation width of 0.25° per image had to be used in order to avoid overlapping reflections owing to the presence of the rather long unit-cell *c* axis of about 420 Å. The data-collection parameters and the final statistics of the processed data are summarized in Table 1. The crystal belonged to the hexagonal space group $P6_122$ or $P6_522$, with up to four molecules in the asymmetric unit. Assuming the presence of three molecules in the asymmetric unit and a molecular mass of 17 102 Da, the Matthews coefficient $V_{\rm M}$ (Matthews, 1968) was 2.44 Å^3 Da^{-1}. This would correspond to a solvent content of approximately 50%. However, neither the self-rotation function (data not shown) nor the native Patterson synthesis (data not shown) shows any features apart from crystallographic symmetry and the origin peak, respectively. Therefore, no further conclusion about the number of molecules in the asymmetric unit can be drawn. A high-resolution diffraction data set was also collected from a flash-cooled crystal of LeuD_1-168. The crystal diffracted X-rays to about 1.4 Å resolution. It belonged to the monoclinic space group $P2_1$ and presumably contains two monomers in the asymmetric unit. With a molecular mass of 18 496 Da per subunit, this would result in a solvent content of 40% and a $V_{\rm M}$ of 2.04 Å^3 Da^{-1}. The assumption of two molecules per asymmetric unit is supported by a clear peak in the $\kappa = 180°$ section of the self-rotation function (Fig. 2).

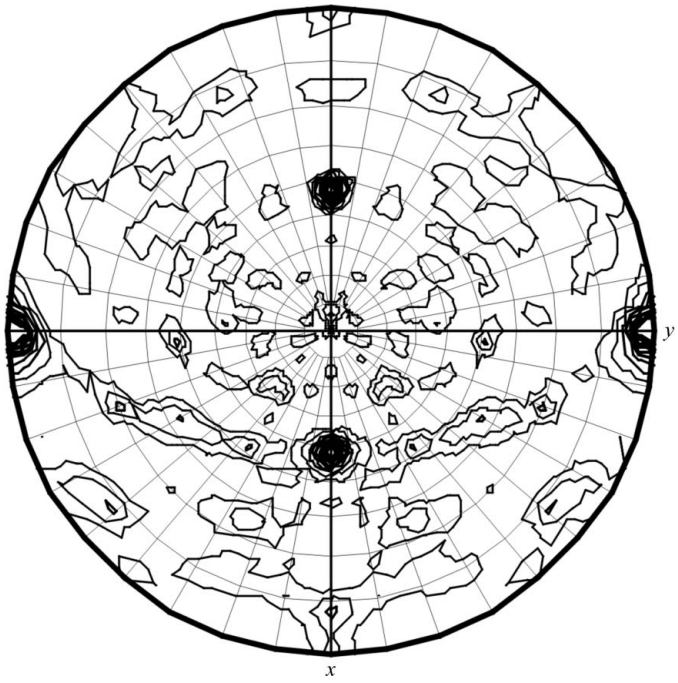


Figure 2
The $\kappa = 180^{\circ}$ section of a self-rotation function map calculated based on structure-factor amplitudes collected from a crystal of LeuD_1-168. The two peaks in the *xz* plane indicate the presence of a noncrystallographic twofold symmetry axis perpendicular to the *y* axis of the crystal.

3.3. Attempts at structure determination

The current version of the Protein Data Bank (Berman *et al.*, 2000) contains three structures related to *Mtb* LeuD. These are the LeuDs from *Streptococcus mutans* (PDB code 2hcu; Z.-Q. Gao, H.-F. Hou, L.-F. Li, Y.-H. Liang, X.-D. Su & Y.-H. Dong, unpublished work), *Methanocaldococcus jannaschii* (PDB code 2pkp; J. Jeyakanthan, D. Gayathri, D. Velmurugan, Y. Agari, A. Ebihara, S. Kuramitsu, A. Shinkai, Y. Shiro & S. Yokoyama, unpublished work) and *Pyrococcus horikoshii* (PDB code 1v7l; Yasutake *et al.*, 2004). With 41% sequence identity, LeuD from *S. mutans* is the most similar to *Mtb* LeuD. Consequently, structure solution was initially attempted by molecular replacement. However, none of the search models tried yielded a molecular-replacement solution. Therefore, structure solution by SeMet MAD phasing has been initiated and is currently in progress.

KM thanks the Alexander-von-Humboldt Foundation for a post-doctoral research fellowship. We would also like to thank the ESRF (Grenoble, France) and the SLS (Villigen, Switzerland) for the allocation and provision of synchrotron beam time.

References

Bange, F.-C., Brown, A. M. & Jacobs, W. R. Jr (1996). *Infect. Immun.* **64**, 1794–1799.
Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Doherty, T. M. (2004). *Trop. Med. Int. Health*, **9**, 818–826.
French, S. & Wilson, K. (1978). *Acta Cryst.* A**34**, 517–525.
Guleria, I., Teitelbaum, R., McAdam, R. A., Kalpana, G., Jacobs, W. R. Jr & Bloom, B. R. (1996). *Nature Med.* **2**, 334–337.
Hondalus, M. K., Bardarov, S., Russell, R., Chan, J., Jacobs, W. R. Jr & Bloom, B. R. (2000). *Infect. Immun.* **68**, 2888–2898.
Koon, N., Squire, C. J. & Baker, E. N. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 8295–8300.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McAdam, R. A., Weisbrod, T. R., Martin, J., Scuderi, J. D., Brown, A. M., Cirillo, J. D., Bloom, B. R. & Jacobs, W. R. Jr (1995). *Infect. Immun.* **63**, 1004–1012.
Mueller-Dieckmann, J. (2006). *Acta Cryst.* D**62**, 1446–1452.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Singh, R. K., Kefala, G., Janowski, R., Mueller-Dieckmann, C., von Kries, J. P. & Weiss, M. S. (2005). *J. Mol. Biol.* **346**, 1–11.
Smith, D., Wiegeshaus, E. & Balasubramanian, V. (2000). *Clin. Infect. Dis.* **31**, S77–S80.
Studier, F. W. (2005). *Protein Expr. Purif.* **41**, 207–234.
Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). *Acta Cryst.* D**55**, 191–205.
Weiss, M. S. (2001). *J. Appl. Cryst.* **34**, 130–135.
Yasutake, Y., Yao, M., Sakai, N., Kirita, T. & Tanaka, I. (2004). *J. Mol. Biol.* **344**, 325–333.