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

Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of DapC (Rv0858c) from *Mycobacterium tuberculosis*

Simone Weyand, Georgia Kefala‡ and Manfred S. Weiss*

EMBL Hamburg Outstation, c/o DESY, Notkestrasse 85, D-22603 Hamburg, Germany

‡ Present address: Salk Institute, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA.

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

Received 29 June 2006
Accepted 10 July 2006

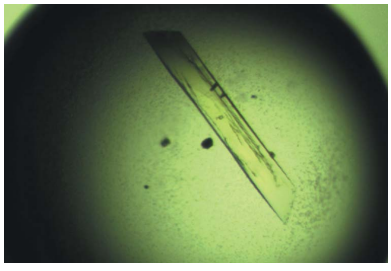
N-Succinyldiaminopimelate aminotransferase from *Mycobacterium tuberculosis* (DAP-AT; DapC; Rv0858c) has been cloned, heterologously expressed in *Escherichia coli*, purified using standard chromatographic techniques and crystallized in two related crystal forms. Preliminary diffraction data analysis suggests the presence of a monomer in the asymmetric unit of the tetragonal crystal form and a dimer in the asymmetric unit of the orthorhombic crystal form.

1. Introduction

In the past few years, the enzymes of the lysine-biosynthetic pathway have drawn considerable attention as potential targets for new antibacterial drugs and herbicides (Hutton *et al.*, 2003). In both plants and bacteria, the lysine-biosynthetic pathway starts from L-aspartate and leads to L-lysine in nine enzymatically catalyzed steps. The sixth step in this pathway is carried out by the enzyme *N*-succinyldiamino pimelate aminotransferase (DAP-AT; DapC; EC 2.6.1.17), which transfers an amino group from L-glutamate to *N*-succinyl-2-L-amino-6-oxoheptanedioate (*N*-succinyl-2-L-amino-6-ketopimelate) in a PLP-dependent reaction, yielding as products *N*-succinyl-L-2,6-diaminoheptanedioate (*N*-succinyl-diaminopimelate) and 2-oxoglutarate (Peterkofsky & Gilvarg, 1961). The reaction proceeds *via* the formation of a Schiff base between the substrate and the cofactor of the enzyme, PLP.

DapC can be efficiently inhibited by hydrazino analogues of the reaction product (Cox *et al.*, 1996) with inhibition constants in the nanomolar range, but the results from antimicrobial activity tests have so far been rather unimpressive (Cox *et al.*, 1998). The reason for this appears to be that the enzyme *N*-acetylornithine aminotransferase (*N*AcO-AT) is promiscuous and can also turn over the substrate of DAP-AT (Cox & Wang, 2001).

DapC has been identified and characterized in some organisms, such as *Escherichia coli* (Peterkofsky & Gilvarg, 1961), *Bordetella pertussis* (Fuchs *et al.*, 2000) and *Corynebacterium glutamicum* (Hartmann *et al.*, 2003), but no three-dimensional structure of a DapC has been reported to date. As part of our efforts to structurally characterize all proteins involved in lysine biosynthesis in *Mycobacterium tuberculosis*, we report here the cloning of DapC (Rv0858c) from genomic DNA, heterologous overexpression of the protein in *E. coli*, purification to homogeneity and successful crystallization in two crystal forms.



2. Experimental methods

2.1. Cloning

The 1200 bp DNA fragment was amplified from genomic DNA of the bacterial strain H37Rv of *M. tuberculosis*. Oligonucleotides were purchased from Invitrogen with the following sequences: 5′-GGGGCATATG**GCT**ACGGTGTCGCGGCTGCGGCCGTATGCG-3′ (forward) and 5′-GGGGCTCGAG**ACCTCGTGGTACACC**GGTAGCGGGTCGCTCGGCGAGCACG-3′ (reverse). The forward primer contains a 5′-*Nde*I restriction site and the reverse primer a 3′-*Xho*I restriction site. In addition, the reverse primer was designed to

Table 1
Data-collection and processing statistics.

	Orthorhombic form	Tetragonal form
No. of crystals	1	1
Wavelength (Å)	0.8048	0.8048
Crystal-to-detector distance (mm)	180	270
Rotation range per image (°)	0.5	0.5
Total rotation range (°)	180	181.5
Resolution range (Å)	99.0–2.00 (2.03–2.00)	99.0–3.20 (3.31–3.20)
Space group	$P222_1$	$P4_122$ or $P4_322$
Unit-cell parameters (Å)	a = 54.34, b = 56.13, c = 247.47	a = 56.00, c = 248.16
Mosaicity (°)	0.54	0.64
Total No. of reflections	384400	76870
Unique reflections	52277	7208
Redundancy	7.4	10.7
$I/\sigma(I)$	18.5 (3.2)	24.9 (2.8)
Completeness (%)	99.8 (100.0)	99.8 (100.0)
R_{merge} (%)	11.3 (64.7)	9.3 (76.5)
$R_{r.i.m.}$ (%)	12.2 (69.5)	9.8 (80.7)
$R_{p.i.m.}$ (%)	4.5 (25.4)	2.9 (25.3)
Overall B factor from Wilson plot (Å^2)	23.7	86.7
Optical resolution (Å)	1.56	2.40

include a thrombin cleavage site. In the forward primer the GCT triplet coding for Ala (bold) was introduced as the second codon to increase the efficiency of expression (Looman *et al.*, 1987). In the reverse primer a thrombin cleavage site (bold) was introduced. This target sequence was amplified by PCR, cloned into the pCR-BluntII-TOPO-vector (Invitrogen) and finally subcloned into the pET22b expression vector (Novagen), which adds a His_6 tag to the C-terminus of the expressed protein. The success of the cloning of the final Rv0858c construct in pET22b was confirmed by sequencing.

2.2. Expression and purification

E. coli BL21 Star (DE3) pRARE cells were used to express the recombinant protein. The cells were cultivated in LB broth medium at 310 K containing ampicillin (50 µg ml^{-1}) and chloramphenicol (30 µg ml^{-1}). Cells were grown to an OD_{600} of approximately 0.6, induced by adding 250 µ*M* isopropyl *β*-D-thiogalactopyranoside (IPTG) and incubated for 16 h at 293 K and 220 rev min^{-1}. Harvested cells were resuspended in lysis buffer [20 m*M* Tris–HCl pH 8.0, 250 m*M* NaCl, 10 m*M* imidazole, 5%(*v*/*v*) glycerol, 3 m*M* *β*-mercaptoethanol and Complete Mini EDTA-free Protease Inhibitor tablets (Roche)] and sonicated three times for 5 min using 0.4 s pulses at 277 K. This cell suspension was spun down and the lysate was filtered through a 0.45 µm membrane and then loaded onto Ni^{2+} beads (Qiagen), which were pre-equilibrated with three column volumes of a buffer containing 20 m*M* Tris–HCl pH 8.0, 250 m*M* NaCl, 10 m*M* imidazole and 3 m*M* *β*-mercaptoethanol. The protein/resin mixture was incubated for 1 h at 277 K by end-over-end rotation, followed by three different washing steps at room temperature. The first washing step used 15 column volumes of washing buffer (20 m*M* Tris–HCl pH 8.0, 250 m*M* NaCl, 10 m*M* imidazole and 3 m*M* *β*-mercaptoethanol), followed by three column volumes of a high-salt wash (20 m*M* Tris–HCl pH 8.0, 800 m*M* NaCl, 10 m*M* imidazole and 3 m*M* *β*-mercaptoethanol) and an imidazole wash of three column volumes of 20 m*M* Tris–HCl pH 8.0, 250 m*M* NaCl, 50 m*M* imidazole and 3 m*M* *β*-mercaptoethanol. Using a stepwise imidazole gradient of 150, 250 and 350 m*M* imidazole in the elution buffer (20 m*M* Tris–HCl pH 8.0, 250 m*M* NaCl and 3 m*M* *β*-mercaptoethanol), the target protein together with its cofactor PLP was eluted within the imidazole wash as well as the 150 m*M* imidazole elution fractions. This was confirmed by SDS–PAGE and mass spectroscopy of the respective bands on the gel. All target protein-containing fractions were pooled and further purified by size-exclusion chromatography using a Superdex 200 16/60 (Amersham Pharmacia Biotech) column and gel-filtration buffer [20 m*M* Tris–HCl pH 8.0, 250 m*M* NaCl, 5%(*v*/*v*) glycerol and 5 m*M* DTT]. Peak fractions of the yellow-coloured protein were analyzed by SDS–PAGE, pooled and concentrated to 15 mg ml^{-1}. The oligomeric state and the monodispersity of the protein were then investigated by dynamic light scattering at a protein concentration of about 2 mg ml^{-1}.

2.3. Crystallization

The purified Rv0858c protein with the C-terminal His_6 tag still present was used for crystallization using the hanging-drop vapour-diffusion method. 2 µl protein solution (15 mg ml^{-1}) and 2 µl reservoir solution were equilibrated against 500 µl reservoir solution in 24-well Linbro plates. Crystals were obtained in the presence of 20%(*w*/*v*) PEG 3350, Tris–HCl pH 8.5 and 200 m*M* $MgCl_2$ at room temperature within 2–3 d. They grew to maximum dimensions of about 600 × 120 × 100 µm and diffracted X-rays to around 1.9 Å resolution. A second crystal form was grown from very similar conditions: 24%(*w*/*v*) PEG 4000, Tris–HCl pH 8.0 and 200 m*M* $MgCl_2$. These crystals grew to maximum dimensions of up to 800 × 150 × 150 µm at room temperature within 2–3 d and diffracted X-rays very anisotropically to about 3.0 Å resolution

2.4. Diffraction data collection and processing

Two crystals were used for data collection. They were mounted in a cryoloop, soaked quickly in 15%(*v*/*v*) MPD in reservoir solution and flash-cooled to 100 K in a nitrogen stream. Diffraction data sets were collected at beamline X13 at EMBL Hamburg, Germany using a MAR CCD detector. The data were indexed and integrated using *DENZO* (Otwinowski & Minor, 1997) and scaled using *SCALEPACK* (Otwinowski & Minor, 1997). The redundancy-independent merging *R* factor $R_{r.i.m.}$ as well as the precision-indicating merging *R* factor $R_{p.i.m.}$ (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 relevant data-collection and processing parameters are given 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 set was calculated using the program *SFCHECK* (Vaguine *et al.*, 1999). Calculation of the self-rotation function was carried out using the program *MOLREP* (Collaborative Computational Project, Number 4, 1994) and structure-factor amplitudes to a maximum resolution of 4.0 Å.

3. Results and discussion

From the elution volume of the size-exclusion chromatography experiment as well as from the dynamic light-scattering experiment, the protein appeared to be monomeric in solution. This is to some extent in contrast to what might have been expected for a member of the class I PLP-dependent enzymes (Schneider *et al.*, 2000), which typically occur as homodimers. However, it is possible that at the low protein concentrations during the chromatography and the DLS experiment monomers are the preferred species.

Two different crystal forms were obtained under two very similar crystallization conditions. The most significant difference between the two conditions is the difference in pH of half a pH unit. The first modification grew in clusters (Fig. 1*a*), which had to be broken apart for data collection. The second modification (Fig. 1*b*) appeared to

grow as a single crystal, but a diffraction experiment confirmed the presence of multiple lattices. Both crystal forms exhibited a distinct yellow colour, which is probably the result of PLP being present in the crystallized material. PLP was not added to the crystallization experiments and must have copurified with the protein.

For data collection on a crystal of form 1, a needle was broken off from a cluster. Although the needle looked like a single crystal, the diffraction pattern exhibited the presence of many lattices in similar orientations. Nevertheless, since the diffraction limit extended to beyond 2.0 Å, a complete data set was then collected. Despite the presence of many lattices, it was possible to process the data by selecting a single lattice. The resulting statistics are shown in Table 1. The quality of the data set is not outstanding, as indicated by the relatively high merging R factor, which is certainly a consequence of the multiple lattices present. The data could be reduced in an orthorhombic space group and the systematic extinctions along the longest axis identified the space group as $P222_1$. Taken together, the unit-cell parameters and the molecular weight of a monomer of Rv0858c of 42 210 Da suggest the presence of two molecules per asymmetric unit. The corresponding Matthews parameter V_M (Matthews, 1968) is 2.24 Å^3 Da^{-1} and the solvent content is 45%. The presence of two molecules per asymmetric unit could be confirmed by examining the peaks present in the self-rotation function (Fig. 2). From this, it appears that one homodimer of the protein is present in the asymmetric unit of the orthorhombic crystal arranged in a pseudo-tetragonal packing. This would be consistent with the notion that the biologically active form of most PLP-dependent enzymes is a homodimer, but it is in contrast to the findings that in solution Rv0858c appears to be a monomer.

A single crystal of form 2 was also used for data collection. As in the case of form 1, the diffraction pattern exhibited the presence of many lattices, despite the fact that the crystal appeared to be a single crystal. Although the diffraction did not extend as far as for form 1, a complete data set was collected to 3.2 Å resolution. The data could be reduced in the tetragonal space groups $P4_122$ or $P4_322$, with unit-cell parameters very similar to those found for the orthorhombic crystal form 1 (Table 1). Consequently, in this crystal form there is room for only one monomer per asymmetric unit (V_M = 2.30 Å^3 Da^{-1}; solvent content 47%).

The similar unit-cell parameters of the two crystal forms and the self-rotation function of the orthorhombic form suggest that the packing of the molecules in the two space groups is related. This is consistent with the fact that the space groups $P4_122$ and $P4_322$ constitute the minimal non-isomorphic supergroups of space group $P222_1$ (Hahn, 2002).

The detailed appearance of Rv0858c in both crystal forms will be revealed by the determination of its three-dimensional structure, which is currently in progress.

4. Conclusion

The protein Rv0858c from *M. tuberculosis* has been cloned, heterologously expressed in *E. coli* and purified to homogeneity. Crystals were grown in two related orthorhombic and tetragonal crystal forms, both of which exhibit a pale yellow colour, indicating the presence of PLP in the crystal.

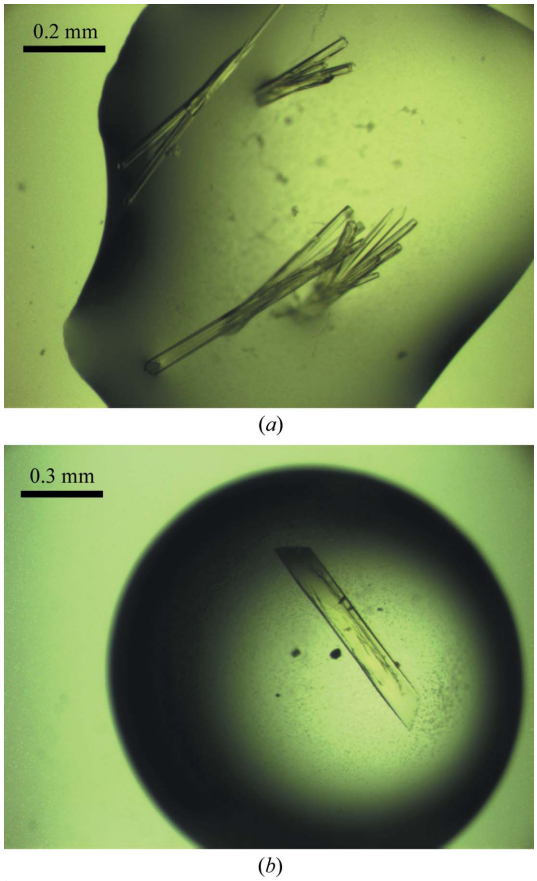


Figure 1
Crystals of Rv0858c from *M. tuberculosis*. (*a*) Clusters of crystals of Rv0858c in the orthorhombic crystal form and (*b*) a crystal of Rv0858c in the tetragonal crystal form.

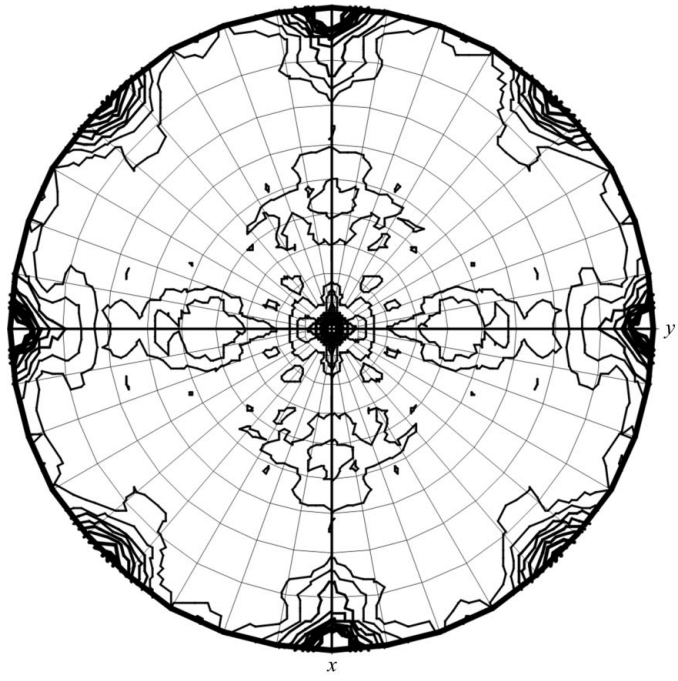


Figure 2
Self-rotation function, κ = 180° section, based on data collected from an orthorhombic crystal of Rv0858c. The pseudo-tetragonal packing of the molecules in this crystal form is apparent from the location of the additional twofold axes in between the crystallographic axes in this section. In addition, the self-rotation function exhibits a large peak along the c direction on the κ = 90° section (not shown). The figure was produced using the program *MOLREP* (Collaborative Computational Project, Number 4, 1994).

We would like to thank Dr L. Jeanne Perry (University of California at Los Angeles, USA) for providing genomic Mtb H37Rv DNA, Dr Arie Geerlof (EMBL) for providing the *E. coli* BL21 Star (DE3) pRARE cells, the TB consortium (http://www.doe-mbi.ucla.edu/TB) for postdoctoral exchange grants to GK and the X-Mtb consortium (http://www.xmtb.org) for funding through BMBF/PTJ grant No. BIO/0312992A. SW is supported by an 'E-STAR' fellowship funded by the EC's FP6 Marie Curie Host fellowship for Early Stage Research Training under contract No. MEST-CT-2004-504640.

References

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cox, R. J., Schouten, J. A., Stentiford, R. A. & Wareing, K. J. (1998). *Bioorg. Med. Chem. Lett.* **8**, 945–950.
Cox, R. J., Sherwin, W. A., Lam, L. K. P. & Vederas, J. C. (1996). *J. Am Chem. Soc.* **118**, 7449–7460.
Cox, R. J. & Wang, P. S. H. (2001). *J. Chem. Soc. Perkin Trans.*, pp. 2006–2008.
French, G. S. & Wilson, K. S. (1978). *Acta Cryst.* A**34**, 517–525.
Fuchs, T. M., Schneider, B., Krumbach, K., Eggeling, L. & Gross, R. (2000). *J. Bacteriol.* **182**, 3626–3631.
Hartmann, M., Tauch, A., Eggeling, L., Bathe, B., Möckel, B., Pühler, A. & Kalinowski, J. (2003). *J. Biotechnol.* **104**, 199–211.
Hahn, T. (2002). Editor. *International Tables for Crystallography*, Vol. *A*, 5th ed. Dordrecht: Kluwer Academic Publishers.
Hutton, C. A., Southwood, T. J. & Turner, J. J. (2003). *Mini Rev. Med. Chem.* **3**, 115–127.
Looman, A. C., Bodlaender, J., Comstock, L. J., Eaton, D., Jhurani, P., de Boer, H. A. & van Knippenberg, P. H. (1987). *EMBO J.* **6**, 2489–2492.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Peterkofsky, B. & Gilvarg, C. (1961). *J. Biol. Chem.* **236**, 1432–1438.
Schneider, G., Käck, H. & Lindqvist, Y. (2000). *Structure*, **8**, R1–R6.
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.