Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structural analysis of mycobacterial branched-chain aminotransferase: implications for inhibitor design

Alina Castell, Christian Mille and Torsten Unge*

Department of Cell and Molecular Biology, Uppsala University, BMC, SE-751 24 Uppsala, Sweden

Correspondence e-mail: torsten.unge@icm.uu.se

Received 30 September 2009
Accepted 8 February 2010

PDB References: *Mycobacterium tuberculosis* branched-chain aminotransferase, apo form, 3dtf; cocrystallized with *O*-benzylhydroxylamine, 3jz6.

The branched-chain aminotransferase (BCAT) of *Mycobacterium tuberculosis* has been characterized as being essential to the survival of the bacterium. The enzyme is pyridoxal 5′-phosphate-dependent and belongs to the aminotransferase IIIa subfamily, to which the human BCATs also belong. The overall sequence similarity is high within the subfamily and the sequence identity among the active-site residues is high. In order to identify structurally unique features of *M. tuberculosis* BCAT, X-ray structural and functional analyses of the closely related BCAT from *M. smegmatis* were carried out. The crystal structures include the apo form at 2.2 Å resolution and a 1.9 Å structure of the holo form cocrystallized with the inhibitor *O*-benzylhydroxylamine (Obe). The analyses highlighted the active-site residues Tyr209 and Gly243 as being structurally unique characteristics of the mycobacterial BCATs relative to the human BCATs. The inhibitory activities of Obe and ammonium sulfate were verified in an inhibition assay. Modelling of the inhibitor Obe in the substrate pocket indicated potential for the design of a mycobacterial-specific inhibitor.

1. Introduction

Tuberculosis causes two million deaths every year and the World Health Organization has estimated that one-third of the world's population is infected (Dye, 2006). Current treatment consists of a combination of several drugs that have to be administered for at least six months. The standard treatment involves the first-line drugs isoniazid, rifampicin, pyrazinamide and ethambutol. Although the effectiveness of these first-line drugs is nearly 95%, patient compliance is poor owing to the lengthy duration of the treatment, which has led to multi-drug-resistant tuberculosis (MDR-TB; Snider & Roper, 1992; Dooley *et al.*, 1992; Kochi *et al.*, 1993; Zignol *et al.*, 2006). The need for new drugs against MDR-TB is therefore compelling.

Essentiality for bacterial survival is an important criterion in selecting an enzyme as a drug target, but additional properties are also required such as selectivity and vulnerability. Furthermore, it should be possible to support the process of optimization of the drug properties of the candidate compounds by structural analysis (Sassetti *et al.*, 2003; Rengarajan *et al.*, 2005; Balganesh & Furr, 2007).

According to analysis using high-density mutagenesis and inhibition studies *in vitro*, the biosynthetic pathways of leucine, isoleucine and valine are required for the growth of *Mycobacterium tuberculosis* (*Mt*; Grandoni *et al.*, 1998; Sassetti *et al.*, 2003; Sassetti & Rubin, 2003). The enzyme

branched-chain aminotransferase (BCAT) transfers an amino group from glutamate to the α-ketoacid of the respective amino acid in the final step in the biosynthesis of branched-chain amino acids. The reaction is reversible and requires that the cofactor pyridoxal 5′-phosphate (PLP) is covalently bound through a Schiff base to a conserved lysine in the active site (Hutson, 1988, 2001). The enzyme coded for by the *rv2210c* gene of *Mt* strain H37Rv is annotated as BCAT IlvE (EC 2.6.1.42; Cole *et al.*, 1998). IlvE belongs to the aminotransferase IIIa subfamily. In addition to its BCAT activity, IlvE catalyzes the formation of methionine from α-ketomethiobutyrate in the last step of the methionine-regeneration pathway (Venos *et al.*, 2004).

The *Homo sapiens* (*Hs*) genome codes for two variants of BCAT. One is localized in the mitochondria (*Hs*-mBCAT) and the other is mainly localized in the cytosol of nervous tissue (*Hs*-cBCAT) (Sweatt *et al.*, 2004). Both variants play important roles in interorgan nitrogen metabolism (Hutson, 2001). Furthermore, *Hs*-mBCAT has been suggested to play a role in the regulation of leucine in the pancreas, whereas the function of *Hs*-cBCAT is less well known, although it might be involved in the regulation of anabolic signals provided by leucine (Sweatt *et al.*, 2004). *Mt*-BCAT and the two *Hs*-BCATs share about 40% sequence identity and the similarity of their active-site residues is high. Thus, in order to design *Mt*-specific inhibitors their structural differences will have to be identified and exploited. The sequence similarity between the mycobacterial BCATs is high and their active sites are conserved. Thus, the *M. smegmatis* BCAT structure may serve as a general model for the design of mycobacterial BCAT inhibitors (Venos *et al.*, 2004).

Aminooxy compounds, including *O*-benzylhydroxylamine (Obe), have been shown to inhibit *Mt* cell growth; the minimum inhibitory concentration (MIC) for Obe is 1.25 m*M* and the competitive K_i is 8.2 m*M* (Braunstein, 1973; Beeler & Churchich, 1976; Venos *et al.*, 2004). This may be a consequence of their inhibitory effect on the enzymatic activity of BCAT, including its methionine-regeneration activity (Berger, 2000; Venos *et al.*, 2004). The inhibitory mechanism has been suggested to take place through reaction with the Schiff base, the formation of a stable oxime product with PLP and the release of the oxime product from the active site of the enzyme (Braunstein, 1973; Beeler & Churchich, 1976).

Here, we report the cloning, expression and activity/inhibition studies of the two homologous BCAT enzymes *Mt*-BCAT (Rv2210c) and *M. smegmatis* BCAT (*Ms*-BCAT; Msmeg4276) and structural studies of *Ms*-BCAT. The activity measurements verify the BCAT activity and the inhibitory properties of Obe and ammonium ions. The structural studies, which are limited to the apo and holo structures of the *Ms* enzyme (owing to a failure to produce crystals of the *Mt* orthologue), provide detailed information about the active site, cofactor binding, substrate pocket and modelling of the binding mode of the inhibitor Obe. Structural comparison with the *Hs* enzyme highlights two significant differences in the active sites that could be exploited in the development of a specific anti-mycobacterial drug. We suggest an alternative inhibition mechanism for Obe and the potential for the design of a specific mycobacterial inhibitor based on the modelling of Obe in the substrate pocket.

2. Experimental methods

2.1. Cloning, expression and purification of *Ms*-BCAT and *Mt*-BCAT

The *msmeg4276* gene was amplified from genomic *M. smegmatis* DNA by PCR with the oligonucleotides 5′-ATGGCTAATAGCGGTCCGCTCGAG-3′ (forward) and 5′-CTAGTTCAGCCGGGCCATC-3′ (reverse). A sequence coding for an N-terminal six-histidine tag was added in a second PCR reaction with the primer 5′-ATGGCTCATCATCATCATAATAGCGGTCCGCTCGAG-3′ (forward). The *rv2210c* gene was similarly amplified from genomic *M. tuberculosis* H37Rv DNA with the primers 5′-ATGGCTACCAGCGGCTCCCTTCAATTC-3′ (forward), 5′-CAAACCCGCTGCCGCACTAC-3′ (reverse) and 5′-ATGGCTCATCATCATCATCATCATACCAGCGGCTCCCTTCAATTC-3′. The gene was ligated to the pEXP5-CT/TOPO vector (Invitrogen). Cloning and expression were performed in *Escherichia coli* TOP10F cells (Invitrogen) and *E. coli* BL21-AI cells (Invitrogen), respectively. The cell culture was induced with 0.2% arabinose (Fluka) at an OD_{550} of 0.8 and was continued at 289 K for 13 h. 2 g of cells were harvested per litre of cell culture and were suspended in lysis buffer [50 m*M* NaH_2PO_4, 300 m*M* NaCl, 5 m*M* imidazole and 10%(*v*/*v*) glycerol] supplemented with 0.02 mg ml^{-1} DNase I and 0.01 mg ml^{-1} RNase A. Lysis was performed in a Constant Cell Disruptor (Constant Systems Ltd) operated at 250 MPa. The lysate was cleared by centrifugation and the enzyme was purified by Ni–NTA agarose (Qiagen) affinity chromatography. Further purification was performed by size-exclusion chromatography on a HiLoad 16/60 Superdex 75 column (GE Healthcare Biosciences AB), which was equilibrated with 50 m*M* HEPES and 100 m*M* NaCl pH 7.5. Fractions containing the BCAT dimers were collected, saturated with a molar excess of PLP and concentrated to 20 mg ml^{-1} by ultrafiltration in a Vivaspin concentrator (Vivascience). The final material was homogeneous as judged by SDS–PAGE (GE Healthcare Biosciences AB).

2.2. BCAT activity and inhibition measurements for *Ms*-BCAT and *Mt*-BCAT

The BCAT activity was monitored through the formation of L-glutamate by transfer of the amino group from L-leucine, L-valine or L-isoleucine to α-ketoglutarate. L-Glutamic acid was measured using a colorimetric assay (L-glutamic acid colorimetric method; Boehringer Mannheim). The aminotransferase reaction was carried out as described previously (Yvon *et al.*, 1997; Berger *et al.*, 2001). The 40 µl reaction mixture contained 20 m*M* K_2HPO_4 pH 7.5, 3 m*M* α-ketoglutarate, 0.2 m*M* PLP, 6 m*M* substrate (Ile, Leu or Val) and enzyme at a final concentration of 1 µg ml^{-1}. The reaction was carried out at 310 K for 15 min and was then irreversibly

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the outer resolution shell.

	Ms-BCAT-apo (PDB code 3dtf)	*Ms*-BCAT-holo (PDB code 3jz6)
Data collection		
Beamline	ID14 EH2, ESRF	I9-11-3, MAX-lab
Wavelength (Å)	0.933	1.00
Unit-cell parameters (Å, °)	a = 59.9, b = 88.1, c = 73.2, $\alpha = \gamma$ = 90.0, β = 105.6	a = 60.4, b = 88.6, c = 72.9, $\alpha = \gamma$ = 90.0, β = 105.6
Resolution range (Å)	37.32–2.20 (2.32–2.20)	58.22–1.9 (2.0–1.9)
Measured reflections	113560	160773
Unique reflections	36343	50840
$R_{\rm merge}$†	0.072 (0.25)	0.053 (0.14)
$\langle I/\sigma(I)\rangle$	13.0 (4.9)	9.4 (5.3)
Completeness (%)	98.0 (99.3)	87.0 (59.6)
Average redundancy	3.1 (3.1)	3.2 (1.3)
Wilson B value (Å^2)	22.6	22.3
Refinement		
Resolution (Å)	25–2.2	25–1.9
Unique reflections	36301	50724
Reflections in test set	1807	2550
$R_{\rm work}$	0.19	0.20
$R_{\rm free}$	0.24	0.23
No. of atoms		
Protein	5561	5561
Solvent	339	454
PLP	—	15
Sulfate	10	—
Glycerol	12	12
Mean B factors (Å^2)		
Protein	20.8	24.7
Solvent	24.6	32.0
PLP	—	22.0
Sulfate	23.3	—
Glycerol	32.2	40.0
R.m.s. deviations from ideal geometry‡		
Bond distances (Å)	0.006	0.007
Bond angles (°)	1.3	1.1
Ramachandran plot outliers§ (%)	1.5	1.9

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith observation of reflection hkl. ‡ Calculated for the protein using ideal values (Engh & Huber, 1991). § Calculated using a strict-boundary Ramachandran plot (Kleywegt & Jones, 1996).

inactivated by incubation at 368 K for 5 min. The colorimetric reagents were added according to a modified assay protocol. The absorbance was measured at 492 nm at room temperature using a DU 640 spectrophotometer (Beckman). Inhibition of the *Ms*-BCAT and *Mt*-BCAT activity was studied using the *Mt*-BCAT inhibitor Obe, the human BCAT inhibitor gabapentin and ammonium sulfate. Inhibition was measured with Obe at concentrations of 1, 10 and 100 µ*M*, gabapentin at concentrations of 1, 10, 100 and 1000 m*M* and ammonium sulfate at concentrations of 10, 100 and 1000 m*M*. The samples were measured in triplicate. The compounds were added to the reaction mixture 30 min before the reaction was initiated by the addition of L-leucine.

2.3. Crystallization of *Ms*-BCAT

An initial vapour-diffusion crystallization screen was set up with sitting drops at 300 K. The PLP-saturated protein solution (10 mg ml^{-1}) was mixed with precipitant solutions from the JCSG+ suite (Qiagen) in a 1:1 ratio to give a final volume of 2 µl. Crystals grew overnight to dimensions of 0.2 × 0.05 × 0.02 mm in 0.2 *M* ammonium sulfate, 0.1 *M* bis-tris pH 5.5, 25% PEG 3350. The crystallizing agent was optimized to 0.4 *M* ammonium sulfate, 0.2 *M* MES pH 5.5 and 50% PEG 3350 and set up as hanging drops. Cocrystallization of the *Ms*-BCAT holo enzyme with the inhibitor Obe was performed using 100 m*M* MES pH 5.5 and 10% PEG 3350. Obe was incubated with the enzyme solution in a 20:1 molar ratio for 30 min before the crystallization experiment.

2.4. Structure determination and refinement of *Ms*-BCAT

X-ray diffraction data were collected for the *Ms*-BCAT apo enzyme (*Ms*-BCAT-apo) and for the holo enzyme cocrystallized with the inhibitor Obe (*Ms*-BCAT-holo). The data were indexed and integrated with *MOSFLM* (Leslie, 2006) and scaled with *SCALA* (Evans, 1993) in point group *P*2. An *Ms*-BCAT monomer model was derived from the *Hs*-mBCAT structure (PDB code 1ekf; Yennawar *et al.*, 2001) using the programs *SOD* (Kleywegt *et al.*, 2001) and *O* (Jones *et al.*, 1991). Initial phases for the apo structure were calculated by molecular replacement using *MOLREP* (Vagin & Teplyakov, 1997). Initial rigid-body refinement with *REFMAC*5 (Murshudov *et al.*, 1997) resulted in $R_{\rm work}$ and $R_{\rm free}$ values of 36% and 50%, respectively. Repeated cycles of restrained refinement and rebuilding in *O* (Jones *et al.*, 1991) lowered the $R_{\rm free}$ value to 24%. *ARP*/*wARP* (Lamzin & Wilson, 1993) was used to add water molecules. The space group was determined to be $P2_1$.

The apo structure was used as a search model in the molecular-replacement calculations for the holo structure. Initial refinement of the *Ms*-BCAT-holo structure gave an $R_{\rm free}$ of 27%. Further refinements were performed in the *CNS* program suite using a simulated-annealing protocol followed by iterative cycles of energy minimization (Brünger *et al.*, 1998). Data-collection and refinement statistics are shown in Table 1. The PLP-modified Lys204 was treated as a new amino-acid residue in the refinement with *CNS*. Refinement parameters for PLP and the Schiff base were generated using the *XPLO2D* program (Kleywegt & Jones, 1998). *BLAST* (Altschul *et al.*, 1990) was used for sequence comparison and *Indonesia* (http://xray.bmc.uu.se/~dennis/manual) was used for structure-based sequence alignment. *DALI* (Holm & Park, 2000) and the Lsq_explicit option in *O* (Jones *et al.*, 1991) were used for structure comparisons and figures were produced with *O* and rendered with *Molray* (Harris & Jones, 2001). The

total contact area between subunits *A* and *B* was calculated using *AREAIMOL* (Lee & Richards, 1971). Coordinates and structure-factor data for *Ms*-BCAT-apo and *Ms*-BCAT-holo have been deposited in the PDB with entry codes 3dtf and 3jz6, respectively.

3. Results

3.1. BCAT activity and inhibitor studies

The BCAT activity of the *Ms*-BCAT and *Mt*-BCAT enzymes was evaluated on a relative basis by normalizing the activity of the most effective amino donor L-leucine to 100%. Using this procedure, we found that the amino acids could be ranked in the following order with respect to their efficiency as amino donor: Leu > Ile > Val (Table 2). α-Ketoglutarate was used as an acceptor in all experiments.

The inhibition studies were performed with leucine as the substrate. Obe exhibited inhibitory activity against both the *Mt* and *Ms* enzymes (Table 2). The human anticonvulsant drug gabapentin, which is an inhibitor of the human cytosolic BCAT (Goto *et al.*, 2005), had no inhibitory effect on the activity of *Ms*-BCAT or *Mt*-BCAT. Ammonium sulfate, which was used in the crystallization experiments of the apo structure, had inhibitory activity at millimolar concentrations.

3.2. Similarity to other structures

A structure-based amino-acid sequence alignment was made of the BCATs from *M. smegmatis*, *H. sapiens* and *E. coli* (Fig. 1). In addition, the sequence of *M. tuberculosis* BCAT was aligned with that of *M. smegmatis* BCAT and included in the figure. The PLP-binding amino acids Arg101, Lys204, Tyr209, Glu240 and Thr314 (*Ms* numbering) are conserved. The *Ms*-BCAT, *Mt*-BCAT and *Hs*-mBCAT sequences share 39% sequence identity. Interestingly, not all of the active-site residues are conserved. The human enzyme contains the substitutions Gly243→Thr, Ile271→Val and Thr272→Val (*Ms* numbering), whereas the cytosolic *Hs*-BCAT contains a threonine at position 272. The *Ms*-BCAT, *Mt*-BCAT and *Ec*-BCAT sequences share 31% sequence identity. The active-site residues are highly conserved.

Table 2
Substrate specificity and inhibition of *M. smegmatis* and *M. tuberculosis* BCAT.

The amino-group acceptor was α-ketoglutarate. ND, not detected. NA, not applicable.

	Mt-BCAT	*Ms*-BCAT
Substrate†, relative activity (%) ± s.d.		
L-Leucine	100 ± 10	100 ± 7
L-Isoleucine	96 ± 7	90 ± 6
L-Valine	50 ± 5	38 ± 9
L-Phenylalanine	10 ± 6	ND
L-Methionine	N.D	N.D
Inhibitor‡, IC_{50} (μM)		
Obe	65 ± 5	42 ± 3
Gabapentin	>10000	>10000
Other ligands§, IC_{50} (mM)		
Ammonium sulfate	10 ± 2	NA

† Amino-group donor. ‡ L-Leucine was used as the amino donor. § Other molecules found in or near the active site.

3.3. Sequence similarity between the mycobacterial BCATs

The amino-acid sequences of *Ms*-BCAT and *Mt*-BCAT have 84% identity. The difference in sequence identity corresponds to 56 residues that are evenly distributed over the sequence.

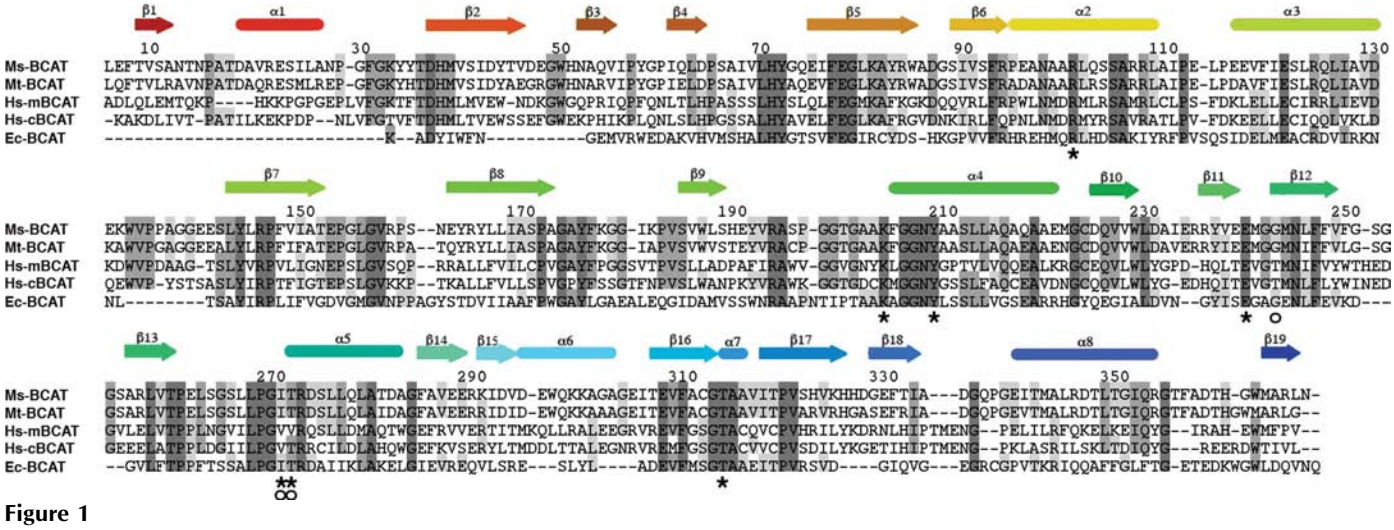

Figure 1
Structure-based sequence alignment of *M. smegmatis* BCAT (*Ms*-BCAT; PDB code 3dtf), *H. sapiens* mitochondrial BCAT (*Hs*-mBCAT; PDB code 1ekf), *H. sapiens* cytosolic BCAT (*Hs*-cBCAT; PDB code 2coj) and *E. coli* BCAT (*Ec*-BCAT; PDB code 1a3g). The sequence of *M. tuberculosis* BCAT (*Mt*-BCAT) was aligned with the sequence of *Ms*-BCAT and included in the figure. Residues that bind directly to the PLP cofactor in the active site are denoted by an asterisk (*). Active-site residues that differ between the mycobacterial and *Hs*-BCATs are denoted by a circle. Lys204 is the active-site residue that forms a Schiff base with PLP. A few active-site residues differ between the sequences, the most important of which is Gly243 (*Ms*-BCAT numbering), which is a threonine in *Hs*-BCAT. The secondary-structure elements of *Ms*-BCAT are coloured as in Fig. 2 and presented as arrows (β-structure) and cylinders (α-helices). The alignment was made with *Indonesia* (http://xray.bmc.uu.se/~dennis/manual). Residues of the linker 1 region that were not present in the structure files (*Hs*-cBCAT and *Ec*-BCAT) were added manually.

However, 40 of these have conserved chemical properties. The residue Ser196 is positioned in the dimerization site of *Ms*-BCAT. This residue is a cysteine in *Mt*-BCAT and thereby enables the formation of a disulfide bridge in *Mt*-BCAT. However, the presence of a disulfide bond could not be experimentally verified by SDS–PAGE under reducing and nonreducing conditions.

3.4. Overall structure

The BCAT structural results are based on two *Ms*-BCAT structures: an apo structure refined to 2.2 Å resolution and a holo structure refined to 1.9 Å resolution. The asymmetric unit contains one *Ms*-BCAT homodimer (Fig. 2). Each subunit has a molecular mass of approximately 40 kDa and is built up of 368 amino acids. Superposition of 363 C^{α} atoms from the two subunits resulted in an r.m.s. deviation of 0.9 Å (Lsq_explicit option in *O*; Jones *et al.*, 1991). The position of α-helix α1 in subunit *B* is shifted by ∼1 Å, which in turn induces a small positional shift of linker 1 in the substrate pocket. This structural difference between the subunits appears to be a consequence of crystal-packing contacts. The N-terminal six-histidine affinity tag and the first five amino-acid residues were not visible in the electron-density map.

On the basis of the structural similarities to *Hs*-BCAT (PDB code 1ekf), the fold of the *Ms*-BCAT structure should be classified as a class IV PLP-dependent aminotransferase (analysis with *DALI*; Holm & Park, 2000) according to the fold-type classification of Grishin *et al.* (1995). However, based on sequence comparison of aminotransferases (Mehta *et al.*, 1993) and the methionine-regenerating activity of mycobacterial BCAT (Berger, 2000) it should be classified as a subfamily IIIa aminotransferase.

Each subunit in the *Ms*-BCAT structure is built up by two α/β domains. The N-terminal domain (residues 6–173 and 364–368) consists of nine β-strands and three α-helices. The C-terminal domain (residues 184–333) consists of a ten-stranded β-sheet and four α-helices. The two domains are connected by two linker sequences (residues 174–184 and 334–363). The *B* factors for the linker 1 residues are high and the quality of the structure differs between the two subunits in this region. Tyr176 is part of the substrate pocket. The contact area between the two subunits is approximately 4000 $Å^2$ as calculated with *AREAIMOL* (Lee & Richards, 1971) and involves polar and hydrophobic residues from both domains of the subunits. The most prominent interactions are the loops 64–74 and 152–163 that reach from one subunit into the substrate-binding area of the other subunit and the formation of a small β-sheet by the β4 strand from each subunit.

3.5. Active site

Each subunit has an active site located at the interface between its domains. PLP is bound to the enzyme *via* four attachment points: the aldehyde C atom, the pyridine hydroxyl group, the pyridine N atom and the phosphate group (Fig. 3). The aldehyde C atom is bound to the ε-amino group of Lys204 through a Schiff base with a distance of 1.4 Å. The lysine approaches the PLP ring from the *Re* face, similar to the other enzymes belonging to aminotransferase subgroup III (Mehta *et al.*, 1993). The Schiff-base plane (C—N—C) deviates from coplanarity with the pyridine ring by an angle of 40°. Although not with optimal geometry, a hydrogen bond with a distance of 2.6 Å between the donor and acceptor atoms is formed between the Lys N^{ε} atom of the Schiff base and the pyridine hydroxyl group (O3). The orientation of the pyridine plane of the PLP cofactor is further determined by a strong hydrogen

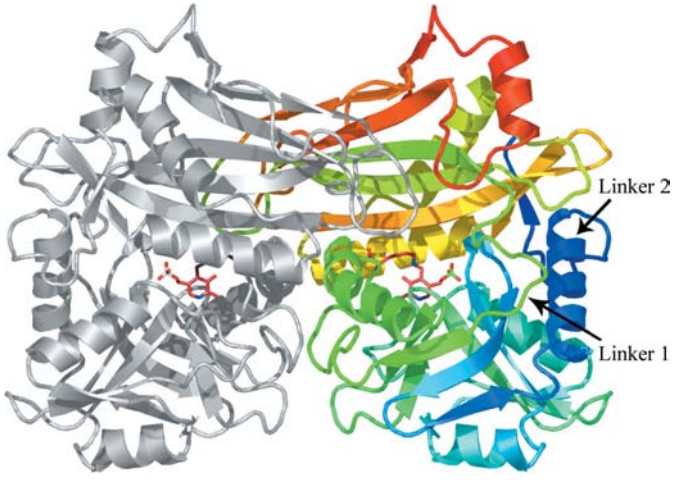

Figure 2
Ribbon diagram of the *M. smegmatis* BCAT homodimer. One subunit is coloured grey and the other is coloured red to blue from the N-terminus to the C-terminus. Each subunit is built up by two domains, with the active site at the interface, in which the cofactor PLP is covalently linked to Lys204 through a Schiff base. Two linkers associate the domains. Linker 1 is inserted in the substrate pocket, whereas the C-terminal linker (linker 2) folds back onto the N-terminal domain.

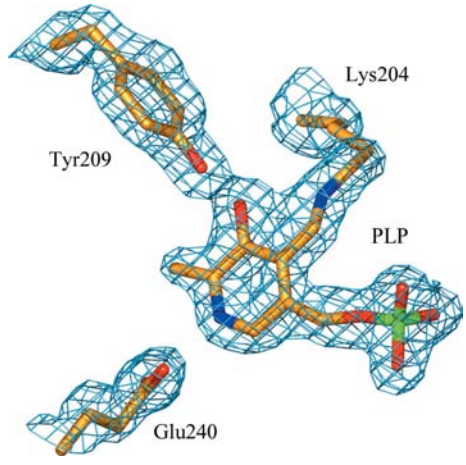

Figure 3
Arrangement of the cofactor PLP in the catalytic reaction centre of *Ms*-BCAT. PLP is bound to Lys204 through a Schiff base. The lysine approaches the PLP ring from the *Re* face. Tyr209 and Glu240 assist in the orientation of the PLP pyridine ring. Tyr209 binds to PLP on the *Si* face of the pyridine ring and with a distance of 2.4 Å between the hydrogen donor and acceptor atoms Tyr209 O and PLP O3. Residues binding to the phosphate group are not shown in the figure. The $2mF_o - DF_c$ density is contoured at a level of 0.30 e $Å^{-3}$.

bond between the hydroxyl group of Tyr209 O and PLP O3 (2.4 Å). Furthermore, a hydrogen bond between Glu240 OE1 and the pyridine N atom of the PLP (2.6 Å) assists in the orientation of the PLP ring. The phosphate group of PLP is bound to Arg101 NH1 and NH2, Ile271 N, Thr272 N and OG1 and Thr314 OG1 and N. In addition, three water molecules are involved in the hydrogen-bond network coordinating the phosphate group.

Comparison of the apo structure and the holo structure indicates that the active-site residues are maintained in their positions through polar and hydrophobic interactions. Despite the absence of PLP, there are only minor structural differences between the apo and the holo structures. Lys204 N^{ε} has only moved by 0.4 Å, the Tyr209 phenolic O atom by 0.5 Å and C^{α} of Gly243 by 0.6 Å. The position of the Lys204 aliphatic side chain is kept in place by a hydrophobic interaction with Phe77. A sulfate ion occupies the position of the phosphate and the position of the Arg101 side chain is maintained through hydrogen bonds to the two carboxylate O atoms of Glu78, Pro269 O and the sulfate ion.

3.6. Structure similarity between *Ms*-BCAT and *Hs*-BCAT

The *Ms*-BCAT structure is similar to previously published structures of BCAT from *H. sapiens* (PDB codes 1ekf and 2coj; Yennawar *et al.*, 2001; Goto *et al.*, 2005) and *E. coli* (PDB code 1a3g; Okada *et al.*, 1997). The two unpublished structures of BCAT from the *Thermus thermophilus* (PDB code 1wrv; M. Goto, I. Miyahara & K. Hirotsu, unpublished work) and *T. maritima* (PDB code 3csw; Joint Center for Structural Genomics, unpublished work) are similar to the *E. coli* BCAT structure. We have compared these known structures with the holo structure of *Ms*-BCAT. A sequence comparison between *Ms*-BCAT and *Hs*-BCAT indicated high similarity between the enzymes. This similarity was reflected to an even greater degree in a comparison of their three-dimensional structures (Fig. 4). Including all C^{α} atoms, comparison of *Ms*-BCAT with cytosolic and mitochondrial *Hs*-BCAT resulted in r.m.s. deviations of 1.6 and 1.7 Å, respectively, as calculated with the Lsq_explicit option in *O* (Jones *et al.*, 1991). When 107 amino acids from the active-site area were used in the comparison, the r.m.s. deviations were 0.7 and 0.8 Å, respectively. The secondary-structure elements are generally well conserved, whereas the loop regions showed greater structural variation.

The amino-acid residues surrounding the active site are conserved, with three exceptions: positions 243, 271 and 272 (*Ms*-BCAT numbering). Amino acid 243 is a glycine in *Ms*-BCAT, whereas it is a threonine in both the cytoplasmic and mitochondrial *Hs*-BCAT enzymes. Amino acid 271 is an isoleucine in *Ms*-BCAT and a valine in *Hs*-BCAT, but the side chain is turned away from the active site. In *Ms*-BCAT amino acid 272 is a threonine, whereas it is a valine in mitochondrial *Hs*-BCAT. The presence of the threonine leads to differences in the coordination of the PLP phosphate group to which it forms a hydrogen bond. In the cytoplasmic enzyme the threonine is conserved.

There is a significant difference in the orientation of Tyr209 in the *Ms*-BCAT structure compared with *Hs*-BCAT. The tyrosine interacts from the *Si* face of the PLP ring in *Ms*-BCAT, whereas in *Hs*-BCAT it interacts from the *Re* face. In this respect, *Ms*-BCAT is identical to *E. coli* BCAT (PDB code 1a3g).

3.7. Model of the inhibitor Obe

The inhibitor Obe was cocrystallized with holo *Ms*-BCAT at pH 5.5 and the collected data set was called *Ms*-BCAT-holo (Table 1). The calculated electron-density map contained weak electron density in the substrate pocket close to the Schiff base and with the shape of the Obe molecule, indicating that Obe was bound to the protein but at low occupancy. Owing to its low occupancy Obe was not included in the refinement, but it could be modelled in the $mF_o - DF_c$ OMIT map if contoured at a level of 0.20 e Å^{-3}. The shape of the electron density fitted the Obe molecule well but does not fit MES or any of the other buffer components. In a further attempt to verify that this electron density originated from Obe and not from MES new crystals were produced without Obe in the presence of 200 m*M* MES. The unit cell of the new crystals was larger and the asymmetric unit contained three dimers. The 2.4 Å electron-density map from these crystals did not contain the Obe density or any density that could be assigned to MES (data not shown).

According to our modelling, the Obe molecule is buried in the substrate pocket and is in contact with residues Phe31, Tyr72*B*, Tyr144, Arg146, Leu156*B*, Val158*B*, Tyr176, Lys204, Tyr209, Gly243 and Ala315, all of which

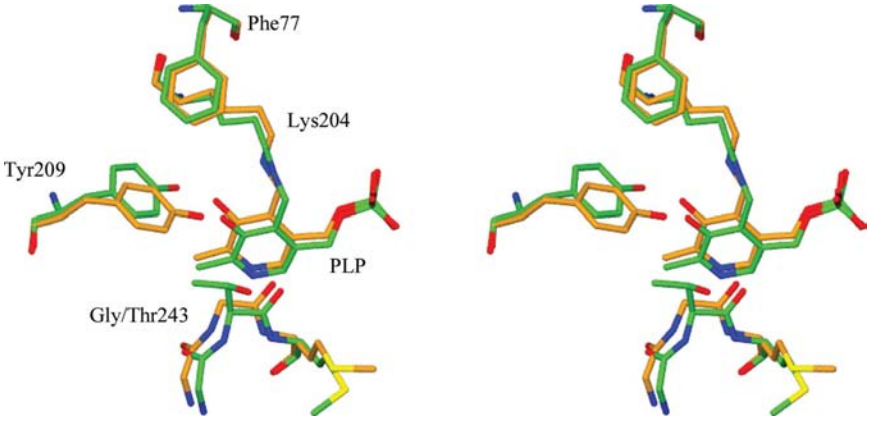

Figure 4
Stereoview of the structural differences between the active sites of *Ms*-BCAT (gold) and the mitochondrial *Hs*-BCAT (green; PDB code 1ekf). At the position of Gly243 in the *M. smegmatis* enzyme the human BCAT enzymes contain a threonine residue. This difference in turn has an effect on the orientation of the side chain of Tyr209, which binds to the PLP pyridine plane on the *Si* face in *Ms*-BCAT and on the *Re* face in *Hs*-BCATs. Gly243 is one of the substrate pocket-lining residues.

are within 4 Å of the Obe molecule (Fig. 5; *B* designates a residue from the other subunit of the dimer). The O—NH_2 moiety of Obe is rotated away from the Schiff base and is within hydrogen-bonding distance of Tyr209 O and Gly243 O. Furthermore, there are three water molecules in contact with the modelled inhibitor; one of these is positioned above Gly243 and forms a hydrogen bond to its main-chain N atom. A threonine residue at the position of Gly243, as in the case of *Hs*-BCAT, would clash with the Obe molecule. In the apo structure the substrate pocket and the entrance of the pocket are filled with water molecules. The electron-density map of the apo structure does not contain any density at the position of the Obe inhibitor, but contains unidentified electron density close to Obe and Gly243 (Fig. 6).

4. Discussion

Branched-chain aminotransferase is involved in the biosynthesis of leucine, isoleucine and valine. The amino-acid sequence similarity between *Ms*-BCAT, *Mt*-BCAT, *Hs*-BCAT and *Ec*-BCAT is high. As expected, the residues involved in PLP binding and substrate-pocket formation are highly conserved. However, the enzymes exhibit some variations in substrate specificity. *Hs*-BCAT favours leucine, isoleucine and valine as substrates; it has low activity towards methionine as a substrate and no significant activity towards aromatic amino acids (Hutson, 2001). However, the prokaryotic BCATs can utilize methionine and also the aromatic amino acids phenylalanine and tryptophan as substrates (Hutson, 2001).

We have recombinantly produced BCAT from *M. tuberculosis* and *M. smegmatis* in *E. coli*. The enzyme assay verified BCAT activity with leucine, isoleucine or valine as amino-group donors. Low activity was detected with phenylalanine and no activity was detected for *Mt*-BCAT with methionine. However, Venos and coworkers used a more sensitive HPLC-based assay and found that *Mt*-BCAT uses phenylalanine as an amino donor in the formation of methionine (Venos *et al.*, 2004). They also showed that *Mt*-BCAT may be involved in the regeneration of methionine in the bacterium and that Obe and other aminooxy compounds inhibit this reaction.

Our experiments showed that Obe inhibited both *Ms*-BCAT and *Mt*-BCAT, while gabapentin, which is an inhibitor of *Hs*-cBCAT, had no inhibitory effect against the two mycobacterial enzymes. On the basis of this observation, we selected Obe as a model inhibitor compound for structural studies. As expected from the chemical properties of the Schiff base, the BCAT reaction was also inhibited by ammonium ions.

The essentiality of an enzyme and the uniqueness of its inhibition pattern are two critical criteria for an enzyme to be classified as a potential drug target. *Mt*-BCAT has been predicted to be essential by the mutation studies of Rubin and coworkers (Sassetti *et al.*, 2003). Although the availability of valine, leucine and isoleucine in the human plasma might reduce the efficacy of an anti-BCAT inhibitor, intracellular bacteria might have limited access to the necessary supplements (Suresh Babu *et al.*, 2002). Further studies to confirm the essentiality, including knockout mutations, remain to be performed.

To investigate the uniqueness of the active site of *Mt*-BCAT primarily relative to the human enzyme but also in comparison with that of the *E. coli* enzyme, we subjected the *Ms*-BCAT and *Mt*-BCAT proteins to crystallization experiments; only the *Ms* enzyme crystallized. Our structural studies of *Ms*-BCAT include the apo and the holo form. A comparison of the apo structure with the holo structure revealed rigidity of the active-site residues. Even in the absence

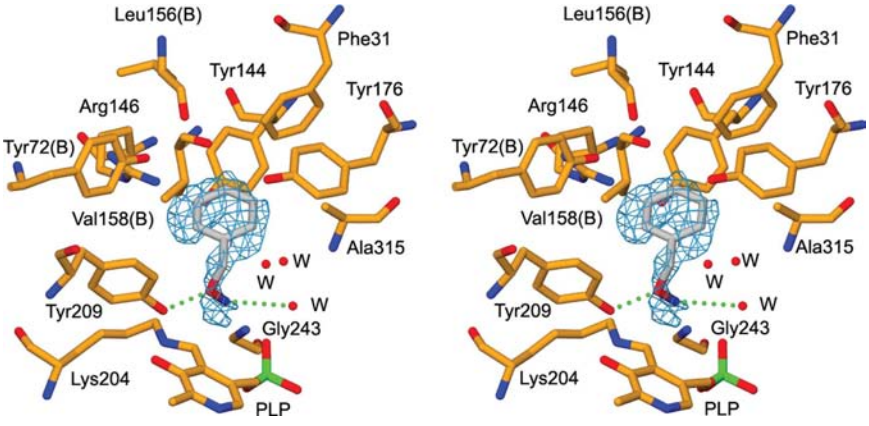

Figure 5
Model of the inhibitor Obe in the substrate pocket of the *Ms*-BCAT-holo structure (stereoview). Obe (grey) was modelled with its O atom above the Schiff base, within hydrogen-bonding distance of a water molecule (W). The residues in the figure are within 4.0 Å of the Obe molecule (*B* designates residues from the other subunit of the dimer). The $mF_o - DF_c$ OMIT map was calculated with simulated annealing in *CNS* and is contoured at a level of 0.20 e $Å^{-3}$.

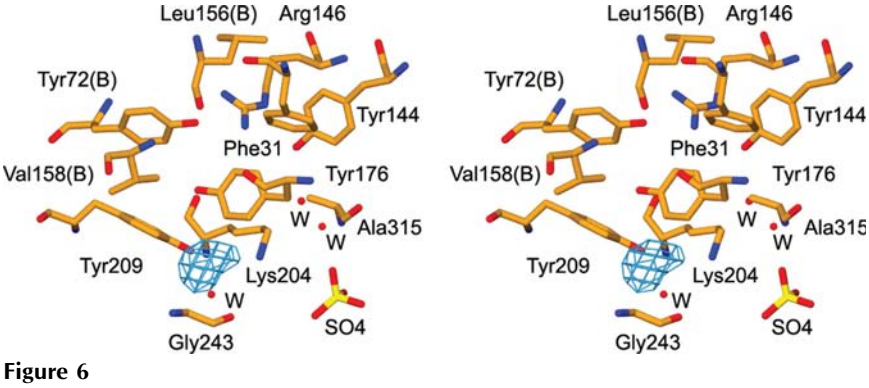

Figure 6
The active site of the apo structure (stereoview). The same residues are shown as in Fig. 5. The figure shows unidentified electron density positioned over Gly243. This density and the water molecules shown in this figure and in Fig. 5 indicate volumes that are accessible to a larger inhibitor compound. The $2mF_o - DF_c$ density is contoured at a level of 0.28 e $Å^{-3}$.

of the PLP molecule the surrounding side chains are held in place by hydrophobic and polar interactions, ready to receive the cofactor. This is in agreement with our observation that PLP binding is reversible and fully substituted molecules could be generated in the presence of excess PLP.

Hs-BCAT has a unique motif of two cysteines at positions 316 and 319 which are essential for enzyme activity (Yennawar *et al.*, 2006). *Ms*-BCAT and *Ec*-BCAT instead have an alanine and a threonine at these positions. Whether these residues play the same role in the bacterial enzymes is not clear.

A structure-based amino-acid sequence comparison of *Ms*-BCAT, *Hs*-BCAT and *Ec*-BCAT and a model of *Mt*-BCAT highlighted position 243 (*Ms*-BCAT numbering) as being an important site of sequence variability. *Ms*-BCAT, *Mt*-BCAT and *Ec*-BCAT contain a glycine at this position, but a threonine is present in *Hs*-BCAT. The *Ms*-BCAT structure shows that Gly243 is one of the residues that line the substrate pocket. The effect of glycine at position 243 is twofold. Firstly, the absence of a side chain allows Tyr209 to approach the PLP pyridine plane from the *Si* face. This arrangement is distinct from that in the human BCAT enzymes, in which the corresponding tyrosine interacts from the *Re* face owing to spatial limitations imposed by the threonine residue. Thus, when the tyrosine is located on the *Si* face of PLP the side chain in general and the phenolic oxygen in particular can interact with the substrate or potential inhibitor molecule in a manner which is completely different from the situation when the tyrosine approaches from the *Re* face. Apart from these differences, the arrangement of the active-site main-chain atoms as well as side-chain atoms is highly conserved.

Many crystallization trials were performed in an effort to trap *Ms*-BCAT in complex with the inhibitor Obe. The difficulty in achieving this is probably a consequence of the reactivity of the compound and the fact that once the oxime product has been formed it diffuses out from the enzyme active site (Beeler & Churchich, 1976; Braunstein, 1973). Although we were unable to trap the oxime product at pH 5.5, we found electron density with the shape of the Obe molecule at a position and orientation that indicated an intermediary complex. Furthermore, the electron-density map indicated that Obe was present at such a low degree of occupancy that full crystallographic refinement of the compound could not be performed. Modelling of Obe in the $mF_o - DF_c$ OMIT map of the holo structure indicated a binding site for Obe in the same position in the substrate pocket as was obtained for gabapentin in *Hs*-cBCAT (Goto *et al.*, 2005). Comparison with the binding mode of gabapentin indicates that the binding sites for the aromatic ring of Obe and the aliphatic ring of gabapentin partly overlap, but the details of their coordination are different, as expected owing to their differences in chemistry. By binding in this position, the Obe molecule may act as an inhibitor by impeding substrate binding. Our model suggests that a potential hydrogen bond can be formed between the Obe O atom and the phenolic O atom of the uniquely oriented Tyr209 and that a hydrogen bond can also be formed between the Obe amino group and a water molecule. The model also indicates that the side chain of a threonine at position 243, as in *Hs*-cBCAT, would clash with the inhibitor. This single amino-acid difference may explain why Obe is inactive towards the human enzyme and indicates the possibility of the design of inhibitors with activity towards *Mt*-BCAT but not *Hs*-cBCAT. The specificity of gabapentin towards *Hs*-cBCAT may be explained by the presence of a hydrogen bond to the threonine.

The presence of unidentified electron density at the 'β-carbon side' of residue Gly243 in the apo structure and a water molecule close to this position in the holo structure indicates that it might be possible to exploit this void for extension of an *Mt*-BCAT inhibitor. Furthermore, the water molecules within 4 Å of the modelled Obe molecule indicate accessible volumes with polar properties. One of these water molecules occupies the same position as one of the gabapentin carboxylate O atoms.

In conclusion, our structural studies indicate that structural differences at positions 209 and 243 offer unique differences between the human and mycobacterial enzymes which might be exploited for the generation of specific inhibitors.

Total DNA of *M. smegmatis* was a gift from Dr Mary Jackson, Institut Pasteur, France. X-ray data were collected at ID14 EH2, ESRF, Grenoble and at I9-11-3, MAX-lab, Lund. We thank our group colleagues Professor T. Alwyn Jones for fruitful discussions in the model-building work and Ms Terese Bergfors for linguistic corrections. Valuable comments on the structure determination were received from Dr Patrik Johansson, Max Planck Institute of Biochemistry, Germany. This tuberculosis-related research was supported by grants from the Swedish Foundation for Strategic Research (SSF), the Swedish Natural Science Research Council and European Commission programs X-TB and NM4TB.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). *J. Mol. Biol.* **215**, 403–410.
Balganesh, T. S. & Furr, B. J. (2007). *Infect. Disord. Drug Targets*, **7**, 120–126.
Beeler, T. & Churchich, J. E. (1976). *J. Biol. Chem.* **251**, 5267–5271.
Berger, B. J. (2000). *Antimicrob. Agents Chemother.* **44**, 2540–2542.
Berger, L. C., Wilson, J., Wood, P. & Berger, B. J. (2001). *J. Bacteriol.* **183**, 4421–4434.
Braunstein, A. E. (1973). *The Enzymes*, 3rd ed., edited by P. D. Boyer, Vol. IX, pp. 379–471. New York: Academic Press.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Cole, S. T. *et al.* (1998). *Nature (London)*, **393**, 537–544.
Dooley, S. W., Jarvis, W. R., Martone, W. J. & Snider, D. E. Jr (1992). *Ann. Intern. Med.* **117**, 257–259.
Dye, C. (2006). *Lancet*, **367**, 938–940.
Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.
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.
Goto, M., Miyahara, I., Hirotsu, K., Conway, M., Yennawar, N., Islam, M. M. & Hutson, S. M. (2005). *J. Biol. Chem.* **280**, 37246–37256.

Grandoni, J. A., Marta, P. T. & Schloss, J. V. (1998). *J. Antimicrob. Chemother.* **42**, 475–482.
Grishin, N. V., Phillips, M. A. & Goldsmith, E. J. (1995). *Protein Sci.* **4**, 1291–1304.
Harris, M. & Jones, T. A. (2001). *Acta Cryst.* D**57**, 1201–1203.
Holm, L. & Park, J. (2000). *Bioinformatics*, **16**, 566–567.
Hutson, S. (2001). *Prog. Nucleic Acid Res. Mol. Biol.* **70**, 175–206.
Hutson, S. M. (1988). *J. Nutr.* **118**, 1475–1481.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kleywegt, G. J. & Jones, T. A. (1996). *Structure*, **15**, 1395–1400.
Kleywegt, G. J. & Jones, T. A. (1998). *Acta Cryst.* D**54**, 1119–1131.
Kleywegt, G. J., Zou, J.-Y., Kjeldgaard, M. & Jones, T. A. (2001). *International Tables for Crystallography*, Vol. *F*, edited by M. G. Rossmann & E. Arnold, pp. 353–356. Dordrecht: Kluwer Academic Publishers.
Kochi, A., Vareldzis, B. & Styblo, K. (1993). *Res. Microbiol.* **144**, 104–110.
Lamzin, V. S. & Wilson, K. S. (1993). *Acta Cryst.* D**49**, 129–147.
Lee, B. & Richards, F. M. (1971). *J. Mol. Biol.* **55**, 379–400.
Leslie, A. G. W. (2006). *Acta Cryst.* D**62**, 48–57.
Mehta, P. K., Hale, T. I. & Christen, P. (1993). *Eur. J. Biochem.* **214**, 549–561.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Okada, K., Hirotsu, K., Sato, M., Hayashi, H. & Kagamiyama, H. (1997). *J. Biochem.* **121**, 637–641.
Rengarajan, J., Bloom, B. R. & Rubin, E. J. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 8327–8332.
Sassetti, C. M., Boyd, D. H. & Rubin, E. J. (2003). *Mol. Microbiol.* **48**, 77–84.
Sassetti, C. M. & Rubin, E. J. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 12989–12994.
Snider, D. E. Jr & Roper, W. L. (1992). *N. Engl. J. Med.* **326**, 703–705.
Suresh Babu, S. V., Shareef, M. M., Pavan Kumar Shetty, A., Taranath Shetty, K. (2002). *Indian J. Clin. Biochem.* **17**, 7–26.
Sweatt, A. J., Wood, M., Suryawan, A., Wallin, R., Willingham, M. C. & Hutson, S. M. (2004). *Am. J. Physiol. Endocrinol. Metab.* **286**, E64–E76.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Venos, E. S., Knodel, M. H., Radford, C. L. & Berger, B. J. (2004). *BMC Microbiol.* **4**, 39.
Yennawar, N., Dunbar, J., Conway, M., Hutson, S. & Farber, G. (2001). *Acta Cryst.* D**57**, 506–515.
Yennawar, N. H., Islam, M. M., Conway, M., Wallin, R. & Hutson, S. M. (2006). *J. Biol. Chem.* **281**, 39660–39671.
Yvon, M., Thirouin, S., Rijnen, L., Fromentier, D. & Gripon, J. C. (1997). *Appl. Environ. Microbiol.* **63**, 414–419.
Zignol, M., Hosseini, M. S., Wright, A., Weezenbeek, C. L., Nunn, P., Watt, C. J., Williams, B. G. & Dye, C. (2006). *J. Infect. Dis.* **194**, 479–485.