Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The 1.9 Å resolution structure of *Mycobacterium tuberculosis* 1-deoxy-D-xylulose 5-phosphate reductoisomerase, a potential drug target

Lena M. Henriksson,[a] Christofer Björkelid,[a] Sherry L. Mowbray[b] and Torsten Unge[a]*

[a]Department of Cell and Molecular Biology, Uppsala University, Biomedical Center, Box 596, SE-751 24 Uppsala, Sweden, and [b]Department of Molecular Biology, Swedish University of Agricultural Sciences, Biomedical Center, Box 590, SE-751 24 Uppsala, Sweden

Correspondence e-mail: torsten@xray.bmc.uu.se

Received 13 March 2006
Accepted 23 May 2006

PDB Reference: 1-deoxy-D-xylulose 5-phosphate reductoisomerase, 2c82, r2c82sf.

1-Deoxy-D-xylulose 5-phosphate reductoisomerase catalyzes the NADPH-dependent rearrangement and reduction of 1-deoxy-D-xylulose 5-phosphate to form 2-*C*-methyl-D-erythritol 4-phosphate, as the second step of the deoxyxylulose 5-phosphate/methylerythritol 4-phosphate pathway found in many bacteria and plants. The end product, isopentenyl diphosphate, is the precursor of various isoprenoids vital to all living organisms. The pathway is not found in humans; the mevalonate pathway is instead used for the formation of isopentenyl diphosphate. This difference, combined with its essentiality, makes the reductoisomerase an excellent drug target in a number of pathogenic organisms. The structure of 1-deoxy-D-xylulose 5-phosphate reductoisomerase from *Mycobacterium tuberculosis* (Rv2870c) was solved by molecular replacement and refined to a resolution of 1.9 Å. The enzyme exhibited an estimated k_{cat} of 5.3 s^{-1} and K_m and k_{cat}/K_m values of 7.2 μ*M* and 7.4 × 10^5 M^{-1} s^{-1} for NADPH and 340 μ*M* and 1.6 × 10^4 M^{-1} s^{-1} for 1-deoxy-D-xylulose 5-phosphate. In the structure, a sulfate is bound at the expected site of the phosphate moiety of the sugar substrate. The *M. tuberculosis* enzyme displays a similar fold to the previously published structures from *Escherichia coli* and *Zymomonas mobilis*. Comparisons offer suggestions for the design of specific drugs. Furthermore, the new structure represents an intermediate conformation between the open apo form and the closed holo form observed previously, giving insights into the conformational changes associated with catalysis.

1. Introduction

The high incidence of tuberculosis makes it one of the most serious diseases of our time. According to the World Health Organization (http://www.who.org), one-third of the world's population is currently infected with the causative bacterium *Mycobacterium tuberculosis*, which causes 2 million deaths each year. The need for new and better drugs is urgent owing both to the increased frequency of drug-resistant and multi-drug-resistant strains and to the deadly combination presented by tuberculosis and HIV.

Isoprenoids, such as steroid hormones, carotenoids, cholesterol and ubiquinone, perform vital functions in all living organisms (Sacchettini & Poulter, 1997). The precursor to isoprenoids, isopentenyl diphosphate (IPP), can be formed through the mevalonate pathway (Spurgeon & Porter, 1981) starting from acetyl-CoA. This pathway is present primarily in eukaryotes and archaea (Boucher & Doolittle, 2000). In recent years, an alternative pathway has been found to be present in plants, green algae and many bacteria (Rohmer, 1999; Rohmer *et al.*, 1993; Lichtenthaler, 1999), the so-called deoxyxylulose 5-phosphate/methylerythritol 4-phosphate (DOXP/

MEP) or non-mevalonate pathway (Rohmer, 1999; Rohmer *et al.*, 1993), which instead uses pyruvate and D-glyceraldehyde 3-phosphate as precursors. The first step in the DOXP/MEP pathway is performed by the enzyme 1-deoxy-D-xylulose 5-phosphate synthase (DXS), in which pyruvate and D-glyceraldehyde 3-phosphate are condensed to form 1-deoxy-D-xylulose 5-phosphate (DXP). In the second step, which is NADPH-dependent and requires the presence of a divalent cation such as Mg^{2+}, Co^{2+} or Mn^{2+} (Takahashi *et al.*, 1998), the rearrangement and reduction of DXP is catalyzed by 1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXR; EC 1.1.1.267), resulting in the product 2-*C*-methyl-D-erythritol 4-phosphate (MEP).

Since humans use such a different way of obtaining the essential isoprenoids, all the enzymes involved in the DOXP/MEP pathway should be considered as potential targets for drugs and herbicides. One known inhibitor of the DOXP/MEP pathway in plants and bacteria is fosmidomycin (Shigi, 1989), which specifically inhibits DXR (Kuzuyama, Shimizu *et al.*, 1998; Zeidler *et al.*, 1998). This compound has been shown to be active against *Plasmodium falciparum* in humans (Missinou *et al.*, 2002; Lell *et al.*, 2003) and *P. vinckei* in mice (Jomaa *et al.*, 1999).

12 DXR structures are presently available in the Protein Data Bank (PDB; Berman *et al.*, 2000), representing enzymes from the bacteria *Escherichia coli* and *Zymomonas mobilis*. Among the *E. coli* entries are two apo structures (PDB codes 1k5h and 1onn; Reuter *et al.*, 2002; Steinbacher *et al.*, 2003), structures containing a bisphosphonate and a sulfate, with and without Mg^{2+} (PDB codes 1t1s and 1t1r; Yajima *et al.*, 2004), and two structures with Mn^{2+} bound, one in complex with the inhibitor fosmidomycin (PDB codes 1ono and 1onp; Steinbacher *et al.*, 2003). Four additional *E. coli* structures have NADPH present: one with bound sulfate, one in complex with the inhibitor fosmidomycin, one with fosmidomycin and citric acid and one in complex with the substrate DXP (PDB codes 1jvs, 1q0l, 1q0h and 1q0q, respectively; Yajima *et al.*, 2002; MacSweeney *et al.*, 2005). Two *Z. mobilis* DXR structures have been determined, one with an acetate ion (PDB code 1r0k) and one in complex with NADPH (PDB code 1r0l; Ricagno *et al.*, 2004).

Here, we report the cloning, expression, purification, kinetic characterization and X-ray structure of *M. tuberculosis* DXR (*Mt*DXR), Rv2870c (http://genolist.pasteur.fr/TubercuList/; Cole *et al.*, 1998), prerequisite steps in future structure-based drug-discovery studies.

2. Materials and methods

2.1. Cloning, protein expression and purification

The sequence corresponding to *Mt*DXR (Rv2870c; amino acids 1–393 of 413, molecular weight 42.0 kDa) was amplified by PCR from a template provided by AstraZeneca India Pvt. Ltd, originating from *M. tuberculosis* strain H37Rv (Cole *et al.*, 1998). The primers 5′-CACCATGGCTCATCATCATCATCATCATGTGACCAACTCGACCGAC-3′ (forward) and 5′-CTACGCCACCGAAGCCATACCAG-3′ (reverse), used together with *Pfu* Ultra DNA polymerase (Stratagene), simultaneously introduced an N-terminal His_6 tag. The Champion pET101 Directional TOPO Expression Kit (Invitrogen) was used for ligation of the PCR product into the pET101D-TOPO vector. Cloning was performed in *E. coli* TOP10F′ cells (Invitrogen). Positive clones were selected by growth on Luria agar plates containing 50 μg ml^{-1} ampicillin and plasmids were isolated following the QIAprep Spin Miniprep kit protocol (Qiagen). An analytical PCR was performed using the T7 forward primer from the Champion pET101 Directional TOPO Expression Kit (Invitrogen) and the reverse primer. The clone was verified by DNA-sequence analysis (Uppsala Genome Center, Rudbeck Laboratory). Expression was performed in *E. coli* BL21-STAR (DE3) cells (Invitrogen) at 310 K. The culture was induced with 100 mg l^{-1} IPTG (Sigma) at OD_{550} = 0.7–1.0 and growth was continued for 3 h. The cells were harvested and washed with 1× SSPE buffer (150 m*M* NaCl, 10 m*M* NaH_2PO_4 pH 7.5, 1 m*M* EDTA) and then stored at 253 K.

For purification, thawed cells were treated with lysis buffer (50 m*M* NaH_2PO_4 pH 8.0, 300 m*M* NaCl, 10 m*M* imidazole, 10% glycerol) with 0.01 mg ml^{-1} RNase, 0.02 mg ml^{-1} DNase and lysed with a Constant Cell Disruptor (Constant Systems Ltd) operated at 250 MPa. The cell lysate containing soluble *Mt*DXR was incubated for 30 min at 277 K with Ni–NTA agarose slurry (Qiagen). The resin was washed with ten column volumes of lysis buffer containing 20 m*M* imidazole and the protein was then eluted with four column volumes of the same buffer containing 250 m*M* imidazole. The protein was further purified by size-exclusion chromatography on a HiLoad 16/60 Superdex 75 column (Amersham Pharmacia Biotech) developed with a buffer containing 150 m*M* NaCl and 20 m*M* Tris–HCl pH 7.5. The pooled fractions were concentrated to 5 mg ml^{-1} (based on the calculated absorbance of 0.907 for a 1 mg ml^{-1} solution at 280 nm) with a Vivaspin concentrator (Vivascience, molecular-weight cutoff 10 kDa) and stored at 193 K. The material was more than 98% pure as deduced from SDS–PAGE analysis (PhastSystem, Amersham Biosciences). The yield of pure and soluble protein was 1.5 mg per litre of culture.

2.2. Assay

The activity of *Mt*DXR was evaluated in a spectrophotometric assay (Kuzuyama, Takahashi *et al.*, 1998; Takahashi *et al.*, 1998), in which the NADPH-dependent rearrangement and reduction of DXP to form MEP, catalyzed by DXR, is monitored at 340 nm using the absorption of NADPH ($\varepsilon_{340\,nm}$ = 6220 M^{-1} cm^{-1}). Each assay contained 350 μl reaction buffer (50 m*M* HEPES–NaOH, 1.5 m*M* $MnCl_2$, 4% DMSO, 2 m*M* β-mercaptoethanol, 100 m*M* NaCl), 50 μl *Mt*DXR (in 100 m*M* NaCl, 0.01% Brij-35 and 50 m*M* HEPES-NaOH pH 7.5), 50 μl NADPH (in dH_2O) and 50 μl DXP (in dH_2O; DXP was generously provided by AstraZeneca India Pvt. Ltd). Six measurements were made with final concentrations of 357 n*M* and 204 μ*M* for DXR and

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Data-collection statistics	
Unit-cell parameters (Å)	$a = b = 56.9$, $c = 239.0$
Space group	$P4_1$
Resolution range (Å)	47.78–1.90 (2.00–1.90)
No. of reflections measured	200768
No. of unique reflections	59343
Average multiplicity	3.4 (3.3)
Completeness (%)	99.7 (99.8)
$R_{\rm meas}$†	0.075 (0.393)
$\langle I\rangle/\langle\sigma(I)\rangle$	6.2 (2.3)
Refinement statistics	
Resolution range (Å)	40.00–1.90 (1.95–1.90)
No. of reflections used in working set	56321 (4151)
No. of reflections for $R_{\rm free}$ calculation	2962 (198)
R (%)	21.5 (25.2)
$R_{\rm free}$ (%)	24.7 (28.3)
No. of non-H atoms	5906
No. of solvent waters	334
Mean B factor, protein atoms (Å^2)	36.6
Mean B factor, solvent atoms (Å^2)	42.0
Mean B factor, sulfate ligands (Å^2)	50.6
Ramachandran plot outliers‡ (%)	1.5
R.m.s. deviation from ideal bond length§ (Å)	0.009
R.m.s. deviation from ideal bond angle§ (°)	1.15

† Multiplicity-corrected $R_{\rm merge}$ (Diederichs & Karplus, 1997). ‡ Calculated using a strict-boundary Ramachandran plot (Kleywegt & Jones, 1996). § Calculated using small-molecule-based parameters (Engh & Huber, 1991).

NADPH, respectively, and final DXP concentrations of between 25 and 800 µ*M*. Four reactions were also carried out with final concentrations of 18 n*M* and 800 µ*M* for DXR and DXP, respectively, and final NADPH concentrations between 3 and 25 µ*M*. The reactions were followed as the decrease in absorbance at 340 nm at 295 K (DU 640 spectrophotometer, Beckman). Measurements were made every 5 s during a 3 min period. The slope of the linear phase of each reaction was used in calculating the initial velocity. The non-linear data were fitted using *Mathematica* v.3.0 for Macintosh (Wolfram Research Inc.) to the equation

$$v = \frac{k_{\rm cat}[E]_0[N][D]}{K_{iN}K_D + K_D[N] + K_N[D] + [N][D]},$$

where v is the initial velocity, $[N]$ and $[D]$ are the substrate concentrations, $[E]_0$ is the enzyme concentration, K_N and K_D are the Michaelis constants for the two substrates and K_{iN} is the dissociation constant for NADPH binding to the enzyme.

2.3. Crystallization

*Mt*DXR was crystallized in the presence of the known inhibitor fosmidomycin (generously provided by AstraZeneca India Pvt. Ltd) by a microbatch method. The 2 µl sitting drops consisted of 1 µl protein (5 mg ml^{-1} in the buffer used in the size-exclusion chromatography, with the addition of 1.3 m*M* fosmidomycin) mixed with 1 µl crystallization buffer (40% ethylene glycol, 2 m*M* $MgSO_4$, 20 m*M* DTT, 0.2 m*M* EDTA and 100 m*M* acetate pH 5.0) and were streak-seeded immediately with a horse hair. The final conditions in the drop (75 m*M* NaCl, 10 m*M* Tris–HCl pH 7.5, 20% ethylene glycol, 1 m*M* $MgSO_4$, 10 m*M* DTT, 0.1 m*M* EDTA, 50 m*M* acetate pH 5.0, 0.65 m*M* fosmidomycin, 2.5 mg ml^{-1} *Mt*DXR, measured final pH ∼5.1) were stabilized by vapour diffusion against the same solution without fosmidomycin and protein. Pyramid-shaped crystals appeared within a few days and grew to dimensions of 0.1 × 0.1 × 0.3 mm in four weeks. Before the crystals were flash-cooled in liquid nitrogen, they were transferred to a drop of the reservoir solution complemented with an additional 10% ethylene glycol and 1 m*M* fosmidomycin.

2.4. Data collection, structure determination and refinement

X-ray data were collected under cryoconditions at beamline ID23-1 at the European Synchrotron Radiation Facility (ESRF), Grenoble using a MAR Mosaic 225 CCD Detector (MAR USA Inc.) and a wavelength of 1.072 Å. Diffraction data were indexed and integrated with *MOSFLM* (Leslie, 1999) and processed with *SCALA* (Evans, 1993) (as implemented in the *CCP*4 program suite; Collaborative Computational Project, Number 4, 1994). Statistics are shown in Table 1. The crystals possessed the symmetry of space group $P4_1$ or $P4_3$. The Matthews coefficient (Matthews, 1968) suggested that there were two molecules in the asymmetric unit; the value was predicted to be 2.3 Å^3 Da^{-1} with 46% solvent. The program *SOD* (Kleywegt *et al.*, 2001) was used to generate a search model in which the side chains identical to the *E. coli* DXR (PDB code 1q0q; MacSweeney *et al.*, 2005, 40% sequence identity) were preserved and other residues were changed to alanine. Molecular replacement in *Phaser* (McCoy *et al.*, 2005) located two molecules in the asymmetric unit and clearly showed $P4_1$ to be the correct choice of handedness. Refinement was carried out with the program *REFMAC*5 (Murshudov *et al.*, 1997) and rebuilding was performed with the program *O* (Jones *et al.*, 1991), exploiting the maps resulting from the autobuilding routine in *ARP*/*wARP* (Lamzin & Wilson, 1993) in decisions concerning the protein structure. Tight NCS restraints between the *A* and *B* molecules were used in the last rounds of refinement since this resulted in slightly better refinement statistics and also improved the electron density in some places. 334 waters were added after analyzing the results from the *ARP*/*wARP* water-building routine. Different weights were tested in the refinement in order to find the optimal balance between experimental data and stereochemistry. Final refinement statistics (using a matrix weight of 0.1) are shown in Table 1.

The close similarity between the two subunits and the nearly perfect twofold non-crystallographic axis relating them suggested early on that the higher symmetry space group $P4_122$ should also be considered. However, the statistics during both molecular replacement and refinement were significantly better using the lower symmetry choice; differences seem to lie in the conformations of some side chains.

The accessible surface areas as well as the subunit-contact surfaces were calculated with the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994). *BLAST* (Altschul *et al.*, 1997) was used for identifying similar

sequences and structures. *ClustalW* (Thompson *et al.*, 1994) was used for sequence comparisons and figures were prepared with *O* and *MOLRAY* (Harris & Jones, 2001).

3. Results

3.1. Enzyme properties

The protein corresponding to *Mt*DXR with an N-terminal His_6 tag was overexpressed in *E. coli*, purified and crystallized with the N-terminal His_6 tag intact. The slightly shortened construct used was designed based on observations of proteolysis at the C-terminus of the full-length protein. In order to test the enzymatic activity, the NADPH-dependent rearrangement and reduction of DXP to form MEP was evaluated using a spectrophotometric assay. In the presence of saturating concentrations of Mn^{2+} and monitoring the reaction at room temperature and pH 7.5, *Mt*DXR was estimated to have a k_{cat} value of 5.3 s^{-1}, with K_m values of 7.2 μM for NADPH and 340 μM for DXP.

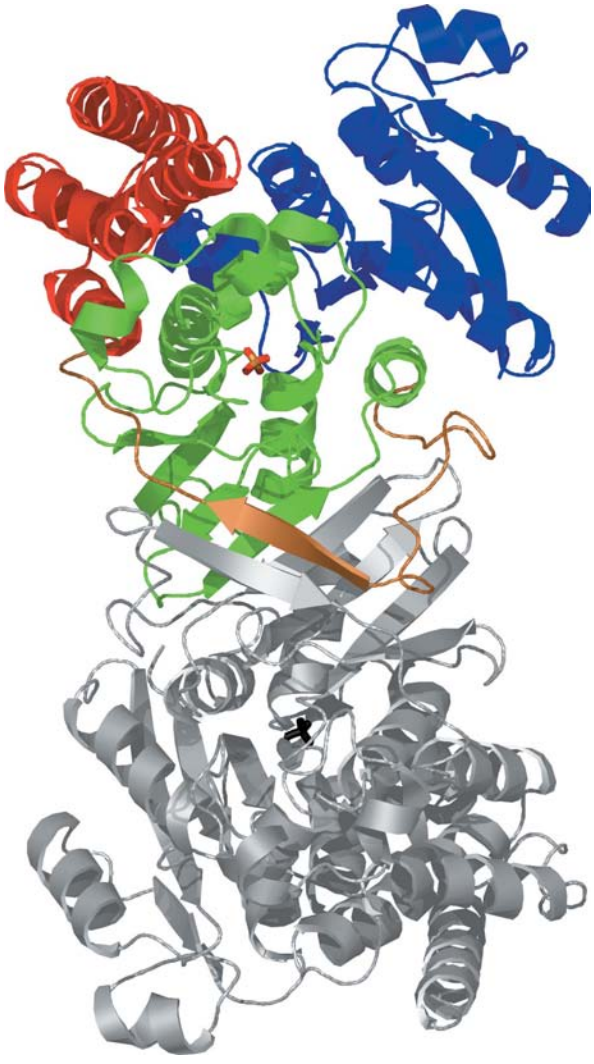

Figure 1
Structure of *Mt*DXR. The overall structure of the *Mt*DXR homodimer is illustrated as a ribbon drawing. The subunit is comprised of an N-terminal NADPH-binding domain (blue), a central catalytic domain (green), a connecting region (gold) and a C-terminal all-α-helical domain (red). A sulfate ion positioned in the active site is shown in gold and red. The dimer interaction with the related subunit (grey, with the sulfate ion in black) is formed by an eight-stranded β-sheet, which is joined to an additional strand from each subunit to form a distorted ten-stranded β-barrel.

3.2. Overall structure

The structure of *Mt*DXR was solved in space group $P4_1$ by molecular replacement using a modified model of the *E. coli* DXR as a search model (MacSweeney *et al.*, 2005) and was refined to 1.9 Å resolution. Data-collection and refinement statistics are summarized in Table 1. The asymmetric unit contains one homodimer. The r.m.s. difference between C^α atomic positions in the two subunits when they are superimposed is 0.015 Å when comparing all C^α atoms with a cutoff of 3.5 Å. Residues 11–389 could be modelled for both molecules from the electron-density map; however, the electron density for amino acids 200–202 is poor. The His_6 tag, the N-terminal residues (1–10) and the last four residues in the construct (390–393) could not be modelled in either subunit.

The subunit is composed of three domains: an N-terminal NADPH-binding domain, a central catalytic domain and a C-terminal α-helical domain (Fig. 1). These are arranged in a V shape, where the N-terminal and C-terminal domains form the two arms and the central domain lies at the vertex. The N-terminal domain (residues 11–150) consists of a Rossmann fold, with a β-sheet containing seven parallel β-strands (β1–β7, including a kink at residue 42 in β2), which is flanked by a total of seven α-helices (α1–α7). The catalytic domain (residues 151–275) includes a four-stranded β-sheet ordered β9-β8-β10-β11, where β10 is positioned antiparallel to the other strands. This sheet adheres to the other two domains by virtue of a layer of helices that line its convex surface. A connecting region (residues 276–306) stretches along the catalytic domain, containing β12 (residues 293–298). The C-terminal domain (residues 307–389) features a four-helix bundle.

The dimer interface is created by interactions between the catalytic and connecting regions of each subunit. A twisted eight-stranded β-sheet is formed using the four β-strands of each catalytic domain, with the respective β11 strands positioned antiparallel at its central point (Fig. 1). Further antiparallel interactions between the β12 strands of each subunit are found on the concave surface of this larger sheet; interactions at the C-terminal ends of the β12 strands link them to the sheet, thus forming an imperfect ten-stranded β-barrel as the core of the dimer interface. A surface area of 4300 $Å^2$ from each subunit is buried by the dimer formation. This corresponds to 30% of the total surface area.

3.3. Active site

A positively charged pocket has been shown previously (MacSweeney *et al.*, 2005) to interact with the phosphate/phosphonate moieties of DXP and fosmidomycin. In our structure, a sulfate ion is bound in this pocket at the expected

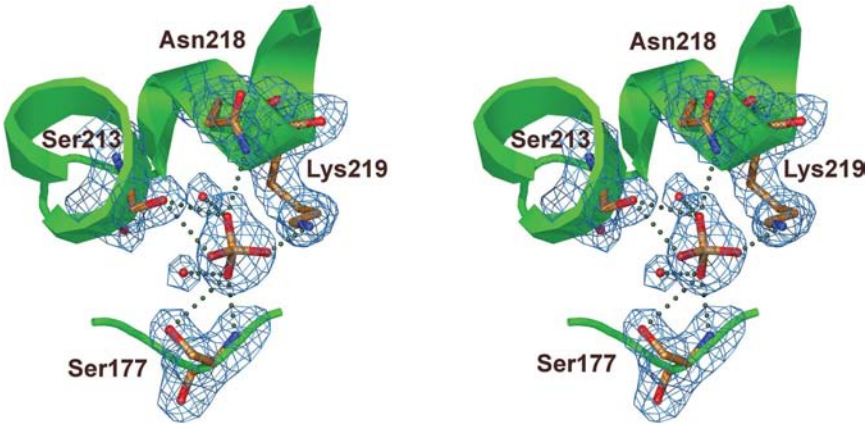

Figure 2
Stereoview of the sulfate ion in the active site of *Mt*DXR. The interactions with residues Ser177 (N, O^{γ}), Ser213 (O^{γ}), Asn218 ($N^{\delta 2}$) and Lys219 (N^{ζ}) are similar to those of the phosphate of the natural substrate, DXP. The *SIGMAA*-weighted (Read, 1986) $|2F_o - F_c|$ map was contoured at 1σ.

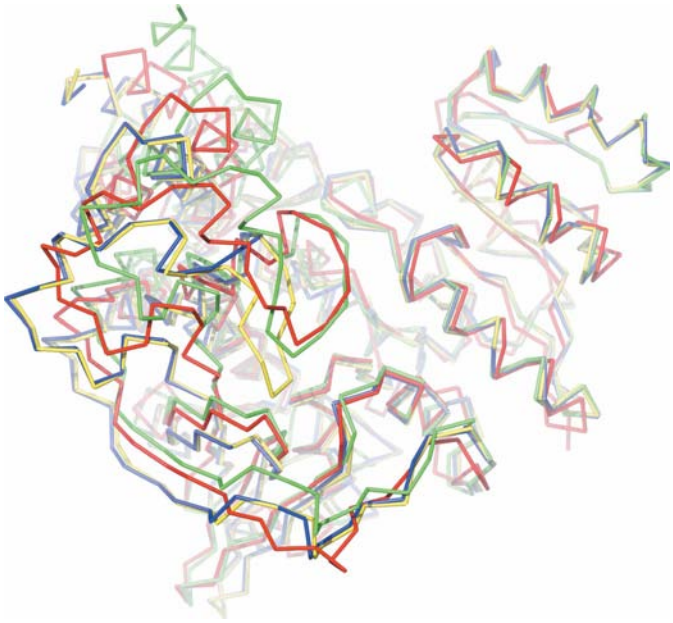

Figure 3
Structural comparisons. The structure of *Mt*DXR (red, complex with SO_4^{2-}) and the *E. coli* DXR structures 1q0q (green, complex with NADPH and DXP; MacSweeney *et al.*, 2005), 1onn (blue, apo structure; Steinbacher *et al.*, 2003;) and 1onp (yellow, complex with fosmidomycin and Mn^{2+}; Steinbacher *et al.*, 2003) have been superimposed based on their N-terminal domains (C^{α} atoms) with the lsq-explicit command in *O* (Jones *et al.*, 1991). The most compact structure is the NADPH–DXP complex 1q0q, as revealed by the position of the C-terminal domain relative to the N-terminal domain. It is notable that the flap over the active site of the *Mt*DXR is positioned similarly to the flap in the NADPH–DXP complex, even though no substrate or cofactor is bound.

position of the phosphate (Fig. 2). The sulfate interacts with two water molecules and with residues Ser177, Ser213, Asn218 and Lys219, which are conserved in the phosphate/phosphonate-binding pocket in previous structures (PDB codes 1q0q and 1q0l; MacSweeney *et al.*, 2005). A distance of 34 Å separates the sulfates in the active sites of the two subunits in the DXR dimer. This, together with the fact that the distorted β-barrel of the dimer interface is expected to be a comparatively rigid structure, suggests that the two active sites function independently.

3.4. Comparison to other DXR structures

Superposition of the *A* subunits of the 13 currently available DXR structures using all C^{α} atoms with the default cutoff of 3.5 Å in the program *LSQMAN* (Kleywegt & Jones, 1997) shows that the r.m.s. differences between C^{α} atomic positions vary between 1.3 and 1.7 Å (with 80–90% of the C^{α} atoms matching and ∼40% amino-acid sequence identity). Although the three-dimensional structural arrangement is clearly conserved, *Mt*DXR exhibits some differences compared with the other structures, including having a shorter $\alpha 3/\beta 4$ unit (residues 70–84) and shorter $\alpha 7/\beta 7$ and $\alpha 8/\beta 8$ loops (residues 142–145 and 164–169, respectively).

Comparison of the various structures shows that DXR is a flexible enzyme in which the C-terminal domain moves relative to the other two, triggered by the binding of cofactors, substrates and inhibitors (MacSweeney *et al.*, 2005). Corresponding points in the structures are shifted by approximately 7 Å as a result of the domain movements. The previously published structures may be grouped into closed, open and super-open forms. One apo structure (PDB code 1k5h; Reuter *et al.*, 2002) at present stands alone as the only representative of the super-open conformation. Structures with a single substrate or inhibitor bound (in either NADPH- or DXP-binding sites) are open, whether or not metal ions are present. Only after both binding sites are occupied does the enzyme assume the closed conformation, as observed in the 1q0h, 1q0l and 1q0q structures (MacSweeney *et al.*, 2005). A superposition of the *Mt*DXR structure (complex with SO_4^{2-}) with the *E. coli* DXR structures 1q0q (closed form in complex with NADPH and DXP; MacSweeney *et al.*, 2005), 1onn (open apo structure; Steinbacher *et al.*, 2003) and 1onp (open form in complex with fosmidomycin and Mn^{2+}; Steinbacher *et al.*, 2003), based on their N-terminal domains, is shown in Fig. 3. Notably, the *Mt*DXR structure, here in complex with only a sulfate ion, represents an intermediate conformation. Although our structure did not contain NADPH, DXP or inhibitor, the flap that covers the active site (residues 198–205) is located in approximately the same conformation as is observed in the ternary complex.

3.5. Comparison of DXR sequences

A sequence alignment of the DXRs from *M. tuberculosis*, *E. coli* and *Z. mobilis* is shown in Fig. 4. The sequence identity between these is 40–41%, although there are some differences within the individual domains. The central catalytic domain is most tightly conserved, with 45–50% amino-acid sequence identity. The N-terminal and C-terminal domains are slightly

less conserved, at ∼40 and 35%, respectively. Thus, the residues in the *E. coli* structure (PDB code 1q0q; MacSweeney *et al.*, 2005) interacting with the substrate DXP are strictly conserved in the three sequences, whereas only 80% of the residues interacting with NADPH are identical.

A 22-residue segment at the C-terminal end of *Mt*DXR is not covered by the structural data. *BLAST* searches with this segment (using the algorithm for short nearly exact matches) uncovered significant similarities to segments of the mycobacterial isopropylmalate synthases (LeuA), as well as to several hypothetical proteins.

4. Discussion

We have shown that the sequence corresponding to *M. tuberculosis* Rv2870c codes for a fully functional DXR. The kinetic parameters measured for the shortened construct of *Mt*DXR used here (k_{cat} = 5.3 s^{-1}, K_m = 7.2 μ*M* for NADPH, K_m = 340 μ*M* for DXP) give rise to calculated k_{cat}/K_m values of 7.4×10^5 M^{-1} s^{-1} for NADPH and 1.6×10^4 M^{-1} s^{-1} for DXP. These values are similar to those previously published for a different construct of the same enzyme, *i.e.* a k_{cat} of 5.0 s^{-1} and K_m values of 3.3 μ*M* for NADPH and 100 μ*M* for DXP, under similar conditions (Argyrou & Blanchard, 2004). Our calculations were performed assuming that NADPH binds first; however, the values obtained do not depend greatly on the kinetic model used, also in agreement with earlier studies (Argyrou & Blanchard, 2004).

The three major domains of DXR have well known folds with defined functions: the N-terminal domain consists of a Rossmann fold, a structure frequently associated with nucleotide binding (Orengo *et al.*, 1997), while the catalytic domain is an α/β-structure and the C-terminal domain is a four-helix bundle. The flexibility of the enzyme was revealed by the different conformations observed in the earlier substrate and inhibitor complexes, such that the C-terminal domain pushes the active site toward the N-terminal domain where the NADPH-binding site resides. A more closed form of the enzyme represents the active form, which is observed only after both substrates have bound. The structure reported here represents a conformational intermediate between this closed form and the more open one where no or only one substrate is bound; such an intermediate conformation has not been observed previously.

In the complete *E. coli* NADPH–DXP complex the flap is well ordered, covers the active site and interacts with NADPH; in the apo structures the flap is either disordered or in a wide open position. Interestingly, in the *Mt*DXR–SO_4^{2-} complex structure the flap occupies essentially the same position as in the complete complex and is only slightly less ordered. The presence of a sulfate or a phosphate, either alone or as part of the substrate or inhibitor, thus seems to have effects on both conformation and order of this flap, despite the fact that there is no direct interaction between them.

The positively charged pocket of the substrate-binding cavity serves to anchor the substrate in the correct position for catalysis. The high affinity of the site is illustrated by the fact that it can trap a sulfate ion when $MgSO_4$ was included at a concentration of only 1 m*M* in the crystallization mixture. Surprisingly, even though fosmidomycin was included at similar concentrations, the enzyme instead formed a complex with the sulfate. The observed preference could be dependent on the relatively low pH (∼5.1) used in the crystallization experiment, combined with the lower pK_a of the sulfate group. The residues predicted to participate in DXP interactions are strictly conserved in the known structures, indicating a conserved reaction mechanism for these enzymes. Less strict conservation seems to be acceptable for the NADPH interactions and indeed for the N- and C-terminal domains as a whole.

There is currently no structural or functional information available for the C-terminal 22-residue segment that is unique to the mycobacterial DXR.

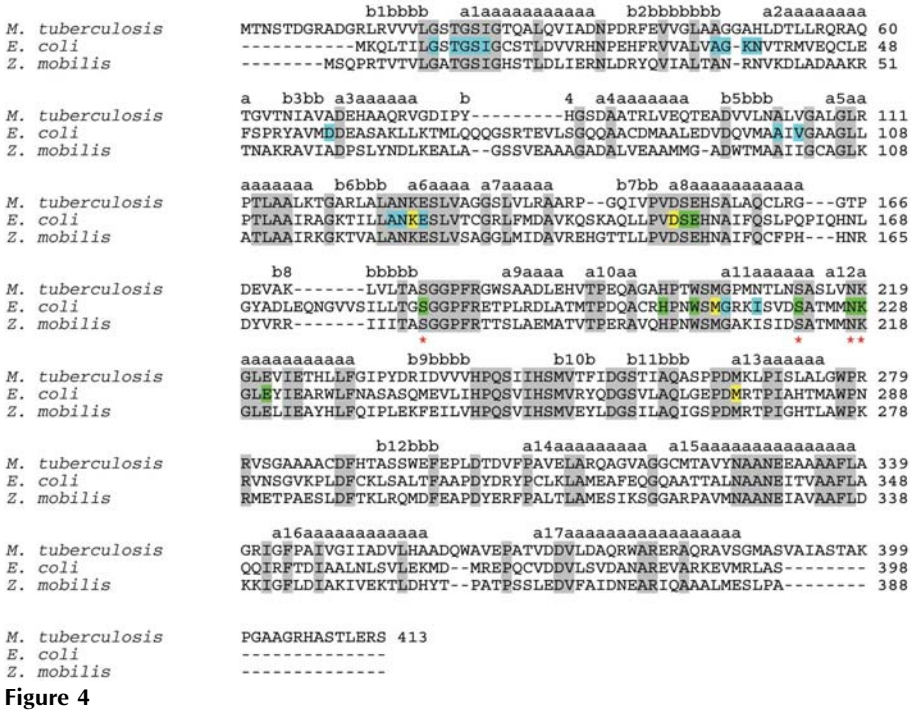

Figure 4
Sequence alignment. DXR sequences correspond to the following entries in GenBank (Benson *et al.*, 2003): *M. tuberculosis* H37Rv, gi:54040750; *E. coli*, gi:2506592; *Z. mobilis*, gi:59802559. Residues in the *E. coli* structure (PDB code 1q0q; MacSweeney *et al.*, 2005) interacting with the substrate DXP are shaded in green and those interacting with NADPH are shaded in cyan. Residues interacting with both DXP and NADPH are shaded in yellow. The red stars indicate residues in *Mt*DXR that are involved in binding to the sulfate ion. Identical residues in the three sequences are shaded in grey. Secondary-structure elements for *Mt*DXR are indicated by a for α-helices and by b for β-strands and naming conventions are indicated as appropriate.

Secondary-structure and other predictions, together with observed sensitivity to proteases, suggested there is no ordered structure in this segment. Searches of the sequence databases indicate that the sequence is peculiar to *M. tuberculosis*. A related sequence is found in the form of a repeated 19-residue insert in the mycobacterial isopropylmalate synthases (LeuA), with the number of repeats varying among species and strains (Chanchaem & Palittapongarnpim, 2002). However, no electron density is observed in the relevant LeuA structure from *M. tuberculosis* strain H37Rv (Koon *et al.*, 2004). Very similar sequences are furthermore seen at the C-terminal end of *M. avium* FadE28, as well as embedded multiple times in the sequences of some hypothetical proteins.

The DOXP/MEP pathway, in which DXR performs the second reaction step, is essential to *M. tuberculosis* but missing in humans, where the mevalonate pathway is instead required for the formation of IPP, the precursor for isoprenoids. These differences suggest the potential usefulness of *Mt*DXR as a drug target. Therefore, structural studies of inhibitor compounds in complex with the enzyme will be a natural next step. The lower sequence conservation in the N-terminal NADPH-binding domain suggests that it is the best choice for designing novel drugs that are specific for *M. tuberculosis*. Further studies will also be needed to determine whether the new conformation we observe in the present structure also represents a species difference and thus an opportunity that can be utilized in drug design. Certainly, this conformation offers insight into the route by which the previously observed open and closed structures are interconverted.

The authors are grateful to AstraZeneca India Pvt. Ltd for providing the template plasmid, the inhibitor fosmidomycin and the substrate 1-deoxy-D-xylulose 5-phosphate. They also thank Patrik Johansson as well as Alina Castell, Evalena Andersson and Kristina Bäckbro for their assistance in data collection. Our tuberculosis-related work is supported by funding from the Foundation for Strategic Research (SSF), the Swedish Research Council (VR), the EU Sixth Framework Programme NM4TB CT:018923 and Uppsala University.

References

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). *Nucleic Acids Res.* **25**, 3389–3402.
Argyrou, A. & Blanchard, J. S. (2004). *Biochemistry*, **43**, 4375–4384.
Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. & Wheeler, D. L. (2003). *Nucleic Acids Res.* **31**, 23–27.
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.
Boucher, Y. & Doolittle, W. F. (2000). *Mol. Microbiol.* **37**, 703–716.
Chanchaem, W. & Palittapongarnpim, P. (2002). *Tuberculosis*, **82**, 1–6.
Cole, S. T. *et al.* (1998). *Nature (London)*, **393**, 537–544.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.
Engh, R. & 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.
Harris, M. & Jones, T. A. (2001). *Acta Cryst.* D**57**, 1201–1203.
Jomaa, H., Wiesner, J., Sanderbrand, S., Altincicek, B., Weidemeyer, C., Hintz, M., Turbachova, I., Eberl, M., Zeidler, J., Lichtenthaler, H. K., Soldati, D. & Beck, E. (1999). *Science*, **285**, 1573–1576.
Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kleywegt, G., 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.
Kleywegt, G. J. & Jones, T. A. (1996). *Structure*, **4**, 1395–1400.
Kleywegt, G. J. & Jones, T. A. (1997). *Methods Enzymol.* **277**, 203–230.
Koon, N., Squire, C. J. & Baker, E. N. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 8295–8300.
Kuzuyama, T., Shimizu, T., Takahashi, S. & Seto, H. (1998). *Tetrahedron Lett.* **39**, 7913–7916.
Kuzuyama, T., Takahashi, S., Watanabe, H. & Seto, H. (1998). *Tetrahedron Lett.* **39**, 4509–4512.
Lamzin, V. S. & Wilson, K. S. (1993). *Acta Cryst.* D**49**, 129–147.
Lell, B., Ruangweerayut, R., Wiesner, J., Missinou, M. A., Schindler, A., Baranek, T., Hintz, M., Hutchinson, D., Jomaa, H. & Kremsner, P. G. (2003). *Antimicrob. Agents Chemother.* **47**, 735–738.
Leslie, A. G. W. (1999). *Acta Cryst.* D**55**, 1696–1702.
Lichtenthaler, H. K. (1999). *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**, 47–65.
McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* D**61**, 458–464.
MacSweeney, A., Lange, R., Fernandes, R. P., Schulz, H., Dale, G. E., Douangamath, A., Proteau, P. J. & Oefner, C. (2005). *J. Mol. Biol.* **345**, 115–127.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Missinou, M. A., Borrmann, S., Schindler, A., Issifou, S., Adegnika, A. A., Matsiegui, P. B., Binder, R., Lell, B., Wiesner, J., Baranek, T., Jomaa, H. & Kremsner, P. G. (2002). *Lancet*, **360**, 1941–1942.
Murshudov, G., Vagin, A. & Dodson, E. (1997). *Acta Cryst.* D**53**, 240–255.
Orengo, C. A., Michie, A. D., Jones, S., Jones, D. T., Swindells, M. B. & Thornton, J. M. (1997). *Structure*, **5**, 1093–1108.
Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.
Reuter, K., Sanderbrand, S., Jomaa, H., Wiesner, J., Steinbrecher, I., Beck, E., Hintz, M., Klebe, G. & Stubbs, M. T. (2002). *J. Biol. Chem.* **277**, 5378–5384.
Ricagno, S., Grolle, S., Bringer-Meyer, S., Sahm, H., Lindqvist, Y. & Schneider, G. (2004). *Biochim. Biophys. Acta*, **1698**, 37–44.
Rohmer, M. (1999). *Nature Prod. Rep.* **16**, 565–574.
Rohmer, M., Knani, M., Simonin, P., Sutter, B. & Sahm, H. (1993). *Biochem. J.* **295**, 517–524.
Sacchettini, J. C. & Poulter, C. D. (1997). *Science*, **277**, 1788–1789.
Shigi, Y. (1989). *J. Antimicrob. Chemother.* **24**, 131–145.
Spurgeon, S. & Porter, J. (1981). *Biosynthesis of Isoprenoid Compounds*, Vol. 1, edited by J. Porter & S. Spurgeon, pp. 1–46. New York: Wiley.
Steinbacher, S., Kaiser, J., Eisenreich, W., Huber, R., Bacher, A. & Rohdich, F. (2003). *J. Biol. Chem.* **278**, 18401–18407.
Takahashi, S., Kuzuyama, T., Watanabe, H. & Seto, H. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 9879–9884.
Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
Yajima, S., Hara, K., Sanders, J. M., Yin, F., Ohsawa, K., Wiesner, J., Jomaa, H. & Oldfield, E. (2004). *J. Am. Chem. Soc.* **126**, 10824–10825.
Yajima, S., Nonaka, T., Kuzuyama, T., Seto, H. & Ohsawa, K. (2002). *J. Biochem. (Tokyo)*, **131**, 313–317.
Zeidler, J., Schwender, J., Muller, C., Wiesner, J., Weidemeyer, C., Beck, E., Jomaa, H. & Lichtenthaler, H. K. (1998). *Z. Naturforsch. C*, **53**, 980–986.