electronic reprint

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449
Editors: **E. N. Baker and Z. Dauter**

Conformational flexibility of *Mycobacterium tuberculosis* thioredoxin reductase: crystal structure and normal-mode analysis

Mohd Akif, Karsten Suhre, Chandra Verma and Shekhar C. Mande

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Conformational flexibility of *Mycobacterium tuberculosis* thioredoxin reductase: crystal structure and normal-mode analysis

Mohd Akif,[a] Karsten Suhre,[b] Chandra Verma[c] and Shekhar C. Mande[a]*

[a]Centre for DNA Fingerprinting and Diagnostics, ECIL Road, Nacharam, Hyderabad 500076, India, [b]Information Genomique et Structurale, IBSM/CNRS, 31 Chemin Joseph Aiguier, 13402 Marseille, France, and [c]Bioinformatics Institute, 30 Biopolis Way, #07-01 Matrix, Singapore 138671, Singapore

Correspondence e-mail: shekhar@cdfd.org.in

Received 23 June 2005
Accepted 23 September 2005

PDB Reference: thioredoxin reductase, 2a87, r2a87sf.

The thioredoxin system exists ubiquitously and participates in essential antioxidant and redox-regulation processes *via* a pair of conserved cysteine residues. In *Mycobacterium tuberculosis*, which lacks a genuine glutathione system, the thioredoxin system provides reducing equivalents inside the cell. The three-dimensional structure of thioredoxin reductase from *M. tuberculosis* has been determined at 3 Å resolution. TLS refinement reveals a large libration axis, showing that NADPH-binding domain has large anisotropic disorder. The relative rotation of the NADPH domain with respect to the FAD domain is necessary for the thioredoxin reduction cycle, as it brings the spatially distant reacting sites close together. Normal-mode analysis carried out based on the elastic network model shows that the motion required to bring about the functional conformational change can be accounted for by motion along one single mode. TLS refinement and normal-mode analysis thus enhance our understanding of the associated conformational changes.

1. Introduction

Thioredoxin (Trx), thioredoxin reductase (TrxR) and NADPH constitute the well known thioredoxin system that is responsible for maintaining reducing conditions inside the cell (Arner & Holmgren, 2000). The other major system required for maintaining intracellular reducing conditions, namely the glutathione system, consists of glutathione, glutaredoxin, glutathione reductase and NADPH. A crucial member of the thioredoxin system, TrxR, is a member of the family of dimeric pyridine nucleotide-disulfide oxidoreductase flavoproteins that catalyze the transfer of electrons between pyridine nucleotide and disulfide/dithiol compounds and promote catalysis *via* FAD and a redox-active disulfide. This family of oxidoreductases includes glutathione reductase, lipoamide dehydrogenase, mercuric reductase and trypanothione reductase (Williams *et al.*, 2000). TrxR catalyzes the reduction of Trx using NADPH which is available in the cell. The thioredoxin system is present in all prokaryotic and eukaryotic cells studied so far and operates *via* a redox-active disulfide, thereby providing electrons for different metabolic processes (Gromer *et al.*, 2004). Trx, a ubiquitous 12 kDa protein, in its dithiol form is the major disulfide reductase in the cells (Williams *et al.*, 2000). TrxR reduces not only Trx but also the glutaredoxin-like protein NrdH in *Escherichia coli* and non-disulfide substrates such as selenite, lipid hydroperoxides and H_2O_2 in the case of mammalian TrxR. Two distinct classes of TrxRs exist in nature. The high-molecular-weight type occurs in mammalian cells, *Drosophila melanogaster*, *Caenorhabditis elegans* and *Plasmodium falciparum*, whereas plants, fungi and

bacteria contain a low-molecular-weight type of TrxR (Gromer *et al.*, 2004).

The NADPH-dependent thioredoxin reductase and thioredoxin were first characterized in *E. coli* as the hydrogen-donor system for ribonucleotide reductase (Arner & Holmgren, 2000). The availability of high-resolution crystal structures of *E. coli* TrxR as a stand-alone structure (F_O conformation) and in complex with its substrate Trx and cofactor analogue (F_R conformation) has resulted in a detailed understanding of the structural basis of the reaction mechanism (Kuriyan *et al.*, 1991; Waksman *et al.*, 1994; Lennon *et al.*, 2000). It has been shown that a sizable conformational change is required during catalysis to allow the pyridine ring of NADPH to approach the isoalloxazine ring of FAD. The complex structure of Trx and TrxR indicates that there is a 66° rotation around the NADPH domain with respect to the FAD domain during catalysis and that this conformational change is necessary to bring the buried active site to the surface of the enzyme. Moreover, the reduced-form *E. coli* TrxR has shown conformational flexibility in the isoalloxazine ring of the FAD cofactor (Lennon *et al.*, 1999). Thus, conformational flexibility of TrxR and its cofactors is the key to its activity.

Mycobacterium tuberculosis, the causative agent of tuberculosis, is an intracellular pathogen that resides in mononuclear phagocytes. The mechanisms by which *M. tuberculosis* and other mycobacteria resist being killed by mononuclear phagocytes are poorly understood. Several mechanisms for resistance to intracellular killing have been proposed, one of which could be the scavenging of free radicals produced by mononuclear phagocytes from infection (Shinnick *et al.*, 1995). One of the proteins implicated in a free-radical scavenging activity is Trx, which has been shown to protect cells from oxidative stress. Furthermore, Trx can regenerate proteins that are inactivated as a result of oxidative stress (Fernando *et al.*, 1992). The genome of *M. tuberculosis* encodes three thioredoxins and a single copy of thioredoxin reductase. *M. tuberculosis* Trx and TrxR have been shown to be involved in the reduction of peroxides and dinitrobenzenes (Zhang *et al.*, 1999) and also detoxify hydroperoxides (Jaeger *et al.*, 2004). Unlike in other organisms, there is a total absence of the glutathione system in the *M. tuberculosis* genome sequence (Cole *et al.*, 1998). Dipteran insects have no glutathione reductase, but possess glutathione, which is reduced by reduced thioredoxin reductase (Bauer *et al.*, 2003). One of the other enzymes involved in the detoxification of peroxide radicals is the peroxiredoxin alkylhydroperoxidase AhpC. As the canonical reducing partner of AhpC, namely AhpF, is absent in the *M. tuberculosis* genome, it has been proposed that TrxR might perform the function of regenerating oxidized AhpC (Chauhan & Mande, 2002). Therefore, features of the thioredoxin system of *M. tuberculosis* and other organisms that lack a glutathione reductase (*e.g.* trypanosomes) are of great interest with regard to the mechanisms of redox reactions and their control (Friemann *et al.*, 2003).

The TrxR in *M. tuberculosis* is a low-molecular-weight dimeric protein containing FAD- and NADPH-binding domains in each subunit, as in *E. coli*. It shares more than 45% sequence identity with *E. coli* TrxR. However, within the mycobacterial species some variation exists in the organization of thioredoxins and thioredoxin reductase, the significance of which is not well understood. For example, in *M. leprae* the TrxR is linked to a Trx *via* an N-terminal peptide like a naturally occurring fusion peptide (Wieles *et al.*, 1995) and this hybrid protein is involved in modulating intracellular killing by phagocytes (Wieles *et al.*, 1997).

In order to determine whether the thioredoxin reductases from mycobacteria are structurally different, we have solved the crystal structure of *M. tuberculosis* TrxR. We find interesting evidence for domain flexibility through TLS refinement, which has been complemented by normal-mode calculations. These observations thus suggest the mechanism by which conformational changes are accomplished between the F_O and F_R states of TrxR.

2. Experimental

2.1. Biochemical activity of *M. tuberculosis* TrxR

5 µl enzyme stock solution (7.5 mg ml^{-1}) was mixed with 0.7 ml 2 m*M* EDTA, 2 µ*M* BSA, 100 m*M* sodium phosphate buffer pH 6.5 supplemented with 200 µ*M* NADPH and 200 µ*M* DTNB. The reduction of DTNB to TNB was followed at 298 K by measuring the increase in absorbance at 412 nm using a Perkin–Elmer Lambda 35 UV–Vis spectrometer. The measurements were carried out for 20 min until the activity reached a plateau.

2.2. Crystallization and data collection

The C-terminal $(His)_6$-tagged *M. tuberculosis* TrxR was purified and crystallized and data were processed as described in Akif *et al.* (2004). Prior to setting up crystallization experiments, *M. tuberculosis* TrxR was incubated in 10 m*M* freshly prepared DTT. Yellow-coloured crystals were obtained in well solution containing 15% PEG 3350 and 0.2 *M* sodium phosphate citrate pH 5. Details of data-collection and reduction statistics have been reported previously (Akif *et al.*, 2004). Briefly, a total of 194 456 reflections constituting 14 186 unique reflections formed the final reduced data. The overall completeness of the data was 97.8%, with the last resolution shell showing a completeness of 97.9%. The overall and the last shell average $I/\sigma(I)$ values were 16.2 and 4.1, respectively. The $R_{\rm merge}$ was 0.081 and 0.369 overall and for the last resolution shell, respectively.

2.3. Structure determination and refinement

The structure was determined by molecular replacement using the coordinates of oxidized TrxR from *E. coli* reported previously at 2 Å resolution (45% sequence identity with *M. tuberculosis* TrxR; PDB code 1tde) as the search model (Waksman *et al.*, 1994). Side-chain atoms were excluded from the model, making it a polyalanine sequence. Moreover, the coenzymes FAD and NADPH were also excluded from the search model. The rotation and translation searches were performed using the dimer as the search model in *AMoRe*

(Navaza, 1994) from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). The correctness of the molecular-replacement solution was confirmed from the electron density of FAD and NADPH in both chains and that of the side-chain atoms. Refinement was carried out using *REFMAC* v.5.0 (Murshudov *et al.*, 1997) interspersed with manual model building using *O* (Jones *et al.*, 1991). The progress of refinement was monitored throughout using 5% reflections set aside randomly for $R_{\rm free}$ calculations. Model rebuilding was performed by examining $\sigma_{\rm A}$-weighted $2F_{\rm o} - F_{\rm c}$ and $F_{\rm o} - F_{\rm c}$ electron-density maps. Refinement was performed using all reflections between 36.0 and 3.0 Å resolution. Tight non-crystallographic symmetry restraints for the main-chain atoms and loose restraints for the side-chain atoms were applied during refinement.

2.4. Modelling of the F_R conformation

The reduced conformation of *M. tuberculosis* TrxR (F_R) was modelled using the F_R (allows flavin reduction by pyridine nucleotide) coordinates of *E. coli* TrxR from the 3 Å resolution crystal structure of the complex between TrxR, Trx and an NADPH analogue (PBD code 1f6m). In this structure, the active-site Cys residues occur on the surface of the protein, unlike in the oxidized structure (F_O), where the Cys residues are buried at the interface of the FAD- and NADPH-binding domains. The FAD domains of both proteins were superimposed on one another using the least-squares procedure in *O*. The NADPH domain of the refined *M. tuberculosis* TrxR structure was then excised, followed by rotating this domain so that it matches that of the *E. coli* F_R conformation. The two domains were then once again joined together.

2.5. Domain-motion analysis

The refined coordinates of the crystal structure of *M. tuberculosis* TrxR and those of the modelled F_R conformation were submitted to the web-based *DynDom* program in order to determine the hinge axes and the residues involved in hinge bending (Hayward & Berendsen, 1997); the *DynDom* default settings were used.

2.6. TLS refinement

TLS refinement was performed with *REFMAC* v.5.0 using a standard script (Murshudov *et al.*, 1997). Residues 124–127 and 249–252 were involved in the hinge region in both chains, as suggested by *DynDom*. The geometric mean of these residues was therefore chosen as the TLS origin. Each domain, *i.e.* the NADPH-binding (residues 126–250) and FAD-binding domains (residues 10–125 and 251–318), from both chains was treated as a rigid TLS group. All temperature factors were set to a constant value of 87 Å^2 approximating the Wilson *B* factor prior to TLS refinement. Six cycles of TLS refinement were carried out followed by 15 cycles of restrained coordinate and residual isotropic temperature-factor refinement. The temperature factors between non-crystallographic symmetry (NCS) related atoms were tightly restrained owing to the poor observation-to-parameter ratio at this resolution. Tight positional non-crystallographic restraints for the main-chain atoms and loose restraints for the side-chain atoms were selected.

The stereochemical quality of the protein model was checked using *PROCHECK* (Laskowski *et al.*, 1993) and *MolProbity* (Lovell *et al.*, 2003). *TLSANL* was used to analyze the TLS parameters. The relative contribution of *B* factors, *i.e.* residual isotropic and TLS, was also separated using this program.

2.7. Conformational dynamics of *M. tuberculosis* TrxR through normal-mode analysis

The conformational flexibility of the *M. tuberculosis* TrxR molecule was analyzed by normal-mode analysis (NMA) using the elastic network model as embedded in the *ElNémo* webserver (Suhre & Sanejouand, 2004). Three residues at the N-terminus of *M. tuberculosis* TrxR were found to be highly flexible and were deleted prior to the NMA calculations. The coordinates were realigned manually using the *LSQMAN* option in *O* before the computation. The maximum amplitude was set to 210 arbitrary units in the computation of mode 8 (Fig. 7), while all other parameters were set to their default values.

3. Results and discussion

3.1. Biochemical activity

The recombinant *M. tuberculosis* TrxR was overexpressed in *E. coli* and purified to 95% homogeneity as observed on SDS–PAGE (data not shown). The typical yield of protein was 60 mg per litre of *E. coli* culture. All protein was obtained in soluble form with a distinct yellow colour owing to the presence of FAD. The measured activity showed that the recombinant protein was fully functional (data not shown). Reduction of DTNB under the conditions tested was complete in approximately 20 min.

3.2. Structure solution and refinement

Crystals of *M. tuberculosis* TrxR were only obtained when the protein was incubated with 10 m*M* DTT prior to crystallization. This suggested that an appropriate redox balance in the crystallization buffer was necessary for successful crystallization. The crystals diffracted to a modest resolution of 3 Å at the Elettra synchrotron source. Analysis of intensities showed a high Wilson *B* value of 87 Å^2. The asymmetric unit of the $P4_12_12$ crystals contained two monomers. At the end of restrained refinement, *R* and $R_{\rm free}$ were 0.25 and 0.345, respectively. At this stage, further improvement in *R* and $R_{\rm free}$ could not be obtained despite extensive inspection of difference electron-density maps and the use of simulated-annealing procedures. It has been suggested that the relatively high *R* value might be a consequence of unmodelled anisotropic displacement (Isupov *et al.*, 1999). We therefore attempted to model anisotropic displacement with TLS refinement. A dramatic improvement in *R* and $R_{\rm free}$ was observed when TLS parameters using four rigid groups were

Table 1
Refinement statistics.

Resolution (Å)	36.0–3.0
No. of reflections	13357
No. of atoms	4813
R factor (%)	21.1
R_{free} (%)	29.0
No. of reflections used for R_{free} calculations	668
Overall *B* factor (Å^2)	87.0
Average *B* values	
Main chain (*A* subunit)	89.0
Side chain (*A* subunit)	89.56
Main chain (*B* subunit)	88.61
Side chain (*B* subunit)	92.65
Water	62
No. of ligands	4
No. of waters	16
R.m.s. deviation, bonds (Å)	0.021
R.m.s. deviation, angles (°)	1.990
Ramachandran plot, non-glycine residue in	
Most favourable region (%)	83.3
Additional allowed region (%)	15.2
Disallowed region (%)	0.8
PDB code	2a87

refined. At the end of TLS refinement, *R* and R_{free} were 0.211 and 0.29, respectively.

With a few exceptions, all the residues of the polypeptide chain were well defined in the electron-density maps. At least ten residues from the N-terminus and 15 residues from the C-terminus were not observed in the electron density. Less well defined regions included disordered residues in two long loops (214–235) in both chains of the protein. The final model includes residues 10–322 of the *A* subunit and residues 15–318 of the *B* subunit, two FAD molecules, two $NADP^+$ molecules and 16 water molecules. The electron density for FAD was well defined in the *A* chain but poorly defined in the *B* chain. The electron density of $NADP^+$ in both chains was not well defined. The diphosphates of the *B*-chain FAD as well as that of the $NADP^+$ were, however, clearly seen in the density. $NADP^+$ was included for refinement as its inclusion led to a marginal drop in R_{free}. The final refinement statistics and the quality parameters are listed in Table 1.

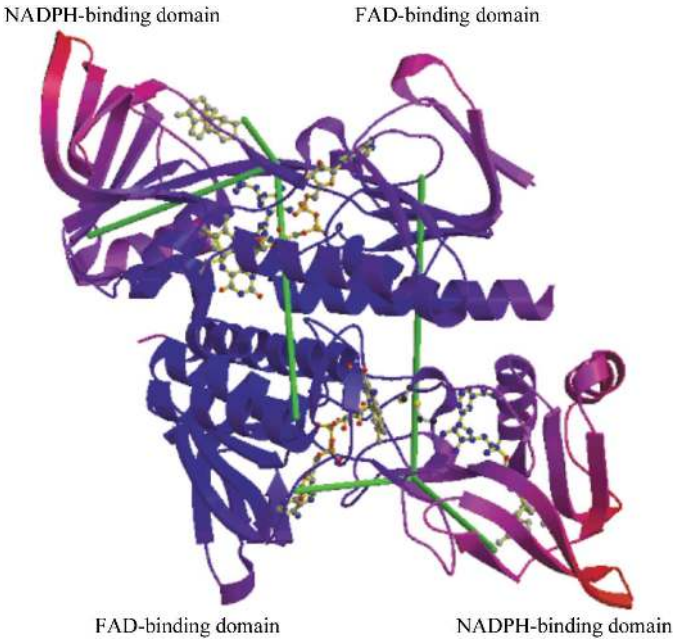

Figure 1
Overall structure of *M. tuberculosis* TrxR dimer. The molecule has been coloured according to the distribution of *B* factors, with the lowest *B* values being depicted in blue and the high values in red. The cofactors FAD and NADPH are shown in ball-and-stick representation. The principal axes of the libration tensor for the NADPH domains of *M. tuberculosis* TrxR are shown in green. Figs. 1, 3, 5 and 7 were generated using *MOLSCRIPT* (Kraulis, 1991) and *RASTER3D* (Merritt & Bacon, 1997).

3.3. TLS refinement and domain flexibility

During refinement of the translation–libration–screw (TLS) parameters, the two domains of each monomer were treated as independent TLS groups. The TLS definition that was used for TLS refinement was very similar to that calculated by *DynDom*. *DynDom* identifies a pronounced hinge-bending motion that relates two domains of two structures and provides information about the location of hinge regions. The putative hinge regions are pivotal points for functionally relevant motion of the NADPH domain of TrxR. These hinge

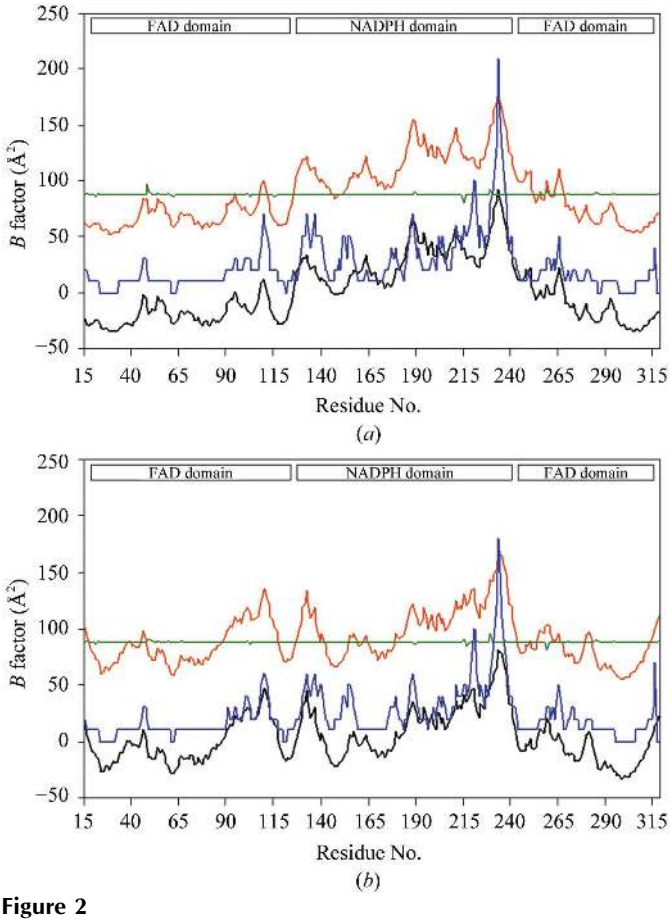

Figure 2
The plot of temperature factor along the polypeptide chain in (*a*) subunit *A* and (*b*) subunit *B*. The total *B* factor (red), the TLS component (black), the residual *B* factor (green) and the *B* factor from NMA (blue) are plotted. As can be seen clearly, there is a close correlation between the NMA-derived and the experimental *B* factors.

regions were therefore chosen as the TLS origin for both the NADPH and FAD domains during TLS refinement.

The axes of libration of both the NADPH domains are nearly parallel to one another (Fig. 1). Of particular interest is the libration of NADPH domain, where the principal axis of libration is nearly perpendicular to the vector joining the centroids of the FAD- and NADPH-binding domains and bisects the vector joining the isoalloxazine ring of FAD and the nicotinamide of NADPH. The mean magnitude of libration of the NADPH domain is 6.5 deg^2. Inspection of the molecule clearly suggests that libration around this axis brings the nicotinamide of NADPH into close vicinity of the isoalloxazine ring of FAD for electron transfer. Moreover, a comparison of different contributions to the *B* factor from TLS refinement (Fig. 2) showed that the residual *B* factors are largely constant over the protein chains. A localized peak at the NADPH domain was seen in the total *B* factor. The high temperature factors of the NADPH domain denote its flexibility compared with the FAD domain.

3.4. Overall structure

The *A*-chain subunit of *M. tuberculosis* TrxR consists of an FAD-binding domain (residues 10–125 and 251–318) and an NADPH-binding domain (residues 126–250) (Fig. 3). The FAD- and NADPH-binding domains have similar folds. The root-mean-square deviation between the FAD and NADPH domains of the *A* subunit was found to be 1.76 Å for 90 equivalent atoms. The secondary-structure folds are almost identical to those of *E. coli* TrxR.

Although NADPH was not added to the crystallization solution, bound nucleotide (NADPH) was observed in the final structure. Arg186, Arg187 and Arg191 bind to the 2′-phosphate of NADPH. The adenine and ribose bind in approximately the same pocket as in *E. coli* TrxR. The NADPH is disordered as evidenced by high temperature factors (>75 Å^2). There is a strong $F_o - F_c$ density up to the 3σ level in the vicinity of the nicotinamide and ribose fragment of NADPH in both chains in the final maps. However, when water molecules were modelled in this density, they showed short contacts with the NADPH. This density has not been ascribed to anything and, therefore, remains unaccounted for.

Overall, the FAD domains appear to be more ordered than the NADPH domains. The average temperature factors of the NADPH domains (119 and 103 Å^2 for the *A* and *B* chains, respectively) are higher that those of the FAD domains (68 and 84 Å^2 for the *A* and *B* chains, respectively). However, one helix at the C-terminus and one strand at the N-terminus of the FAD domains are more disordered (Fig. 1). FAD domains form most of the intersubunit interactions in the physiological dimer, while the NADPH domains are not involved in such interactions. Thus, lack of intersubunit interactions might make the NADPH domain more disordered and possibly more flexible compared with the FAD domain. A total of 27 residues in the FAD domain and only six residues in the NADPH domain of the *A* chain are involved in intermolecular contacts. Similarly, a total of 12 residues in the FAD domain and 11 residues in the NADPH domain of the *B* chain are involved in intermolecular contacts. This is reflected in the non-participation of the NADPH domain in crystal contacts. Thus, owing to a lack of either intersubunit or intermolecular contacts, the NADPH domains appear to be more flexible, which permits large conformational changes during the redox cycle.

The overall structure of *M. tuberculosis* TrxR is quite similar to that of *E. coli* TrxR. The root-mean-square deviation (r.m.s.d.) when all the C$^\alpha$ atoms of the two structures were superimposed was found to be 0.91 Å. The relative positions of the domains in *M. tuberculosis* TrxR differ from those of

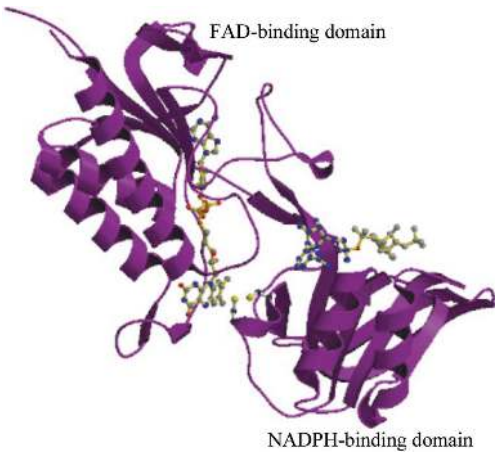

Figure 3
Overall structure of one subunit of the *M. tuberculosis* TrxR structure. The structure is seen to be distinctly divided into two domains: the FAD-binding and NADPH-binding domain.

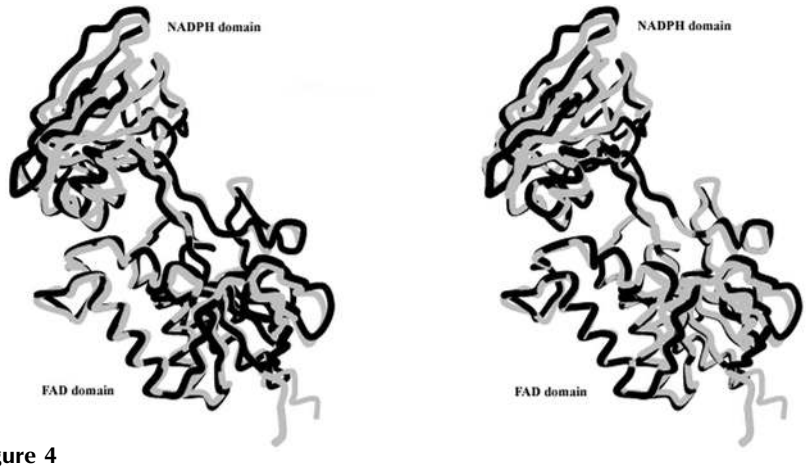

Figure 4
Subunit superimposition of *M. tuberculosis* and oxidized *E. coli* TrxR in stereo. The superposition operation was carried out using only the FAD domain. Using this superimposition, the NADPH domains are observed to be rotated by 11° with respect to each other, as mentioned in the text. The *M. tuberculosis* TrxR is shown in black and the *E. coli* TrxR in grey.

E. coli TrxR. When the FAD domains were superimposed, the NADPH domain of *M. tuberculosis* TrxR was found to be rotated by approximately 11° compared with that of *E. coli* TrxR (Fig. 4).

The active-site disulfide is located in the same position in the NADPH-binding domain as in *E. coli* TrxR. The presence of a disulfide bond between Cys145 and Cys148 was indicated by a distance between the S atoms of these two cysteine residues of 2.03 Å. The isoalloxazine ring of FAD packs against the disulfide and the S atom of Cys148 is close to the isoalloxazine ring (4.23 Å), while Cys145 is farther away from the ring (Fig. 5). This conformation is referred to as the F_O conformation because it allows flavin oxidation by enzyme-disulfide (Waksman *et al.*, 1994). The active-site disulfide, however, is buried and is not available for reaction with thioredoxin. The nicotinamide ring of NADPH is 15.2 Å away from the isoalloxazine of the FAD cofactor in a position that precludes hydride transfer between the pyridine nucleotide and the isoalloxazine ring of flavin. It is interesting to note that although the isoalloxazine ring of the flavin moiety is in a reduced form, possibly owing to the presence of 10 m*M* DTT, the disulfide bond is still observed in the oxidized conformation.

3.5. Dynamic behaviour of *M. tuberculosis* TrxR

To explore the conformational dynamics of *M. tuberculosis* TrxR, normal-mode analysis was carried out. During catalysis, *E. coli* TrxR alternates between F_O and F_R conformations. The crystal structure of *M. tuberculosis* TrxR has been refined in the F_O conformation in this study. The F_R conformation of *M. tuberculosis* TrxR has been modelled based on the corresponding conformation in *E. coli* TrxR. Large-scale conformational changes in proteins, such as that invoked in the TrxR system, often involve coupled motions of several atoms (Gerstein *et al.*, 1994) and in order to examine these, either traditional molecular-dynamics calculations or normal-mode analyses are the best tools available. Normal-mode analysis is a method to express the motions of a protein in terms of collective motions of a group of atoms (Tirion, 1996). It is based on the harmonic approximation of the potential energy function around a minimum energy conformation. This enables an analytical solution to the equations of motion by diagonalizing the mass-weighted second derivatives of the potential energy matrix.

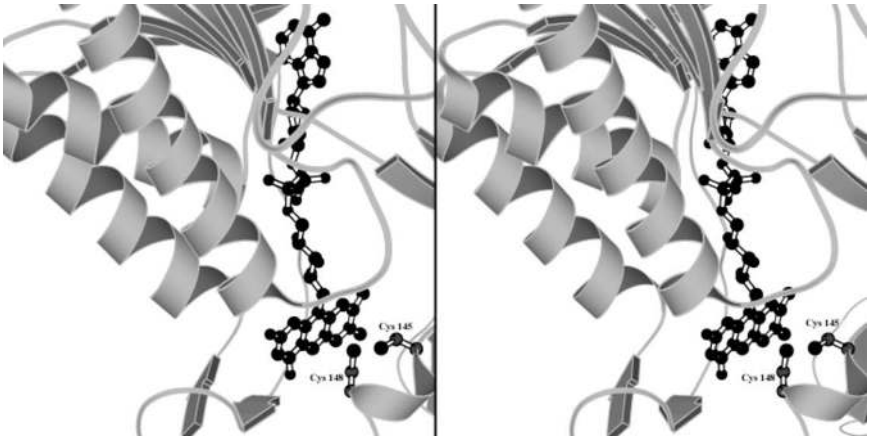

Figure 5
Stereoview of the active site of *M. tuberculosis* TrxR with the location of the redox-active disulfide close to the isoalloxazine moiety of FAD.

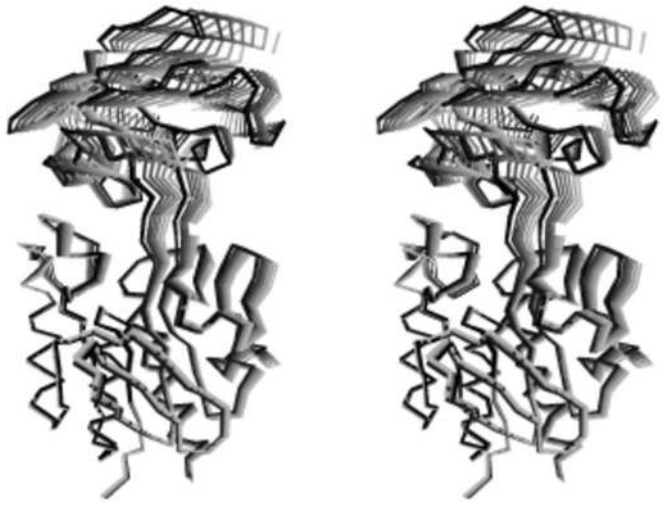

Figure 6
Stereoview of the superposed FAD domains of normal mode 8. The C^{α} trace of the superposed coordinates in stereo showing the movement of NADPH domain from the F_O to the F_R conformation. The different structures shown represent seven intermediates between the F_O and the F_R conformations.

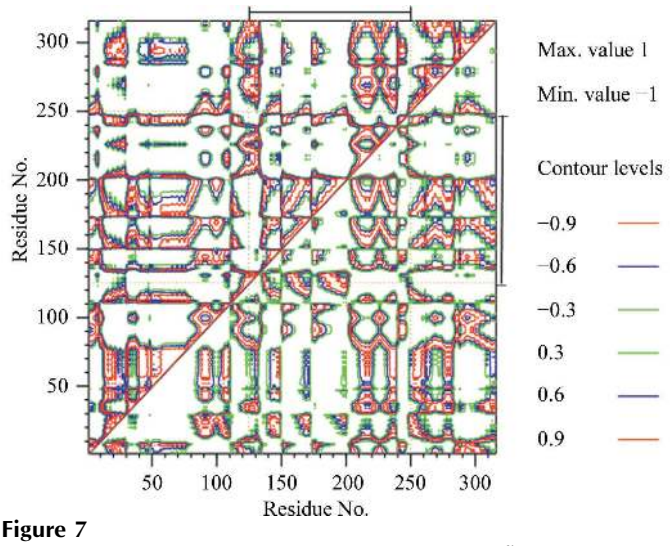

Figure 7
The covariance map showing motions between the C^{α} atoms. Negative covariance refers to atoms that move in opposite directions, while positive covariance refers to atoms that move along the same direction. The horizontal line shows the NADPH domain and the rest of the structure refers to the FAD domain. Covariances above the diagonal are positive, while those below the diagonal are negative.

The eigenvectors of this matrix are the normal modes and the eigenvalues are the squares of the associated frequencies. Each mode represents fluctuations of the protein around a minimum energy conformation. For proteins, the normal modes responsible for most of the amplitude of the atomic displacement are those with the lowest frequencies (Suhre & Sanejouand, 2004) and have been found to be the motions that are functional (Yang & Bahar, 2005). In order to construct the conformations that represent the motions along each mode, the coordinates of the minimum energy structure are deformed along that vector in small units. The number of structures generated can be decided by the user. Normal-mode studies of lysozymes (Brooks & Karplus, 1985) and citrate synthase (Marques & Sanejouand, 1995), as well as of multisubunit proteins such as haemoglobin (Monawad & Perahia, 1996) and ATCase (Thomas, Field, Mouawad *et al.*, 1996; Thomas, Field & Perahia, 1996), have been used to interpret the structural changes that occur on ligand binding. These results indicate that the structures of certain proteins have evolved so that their intrinsic flexibility facilitates the conformational changes required for their function. One such development for the rapid analysis of such large-scale motions is based on the elastic network model and has been now incorporated in a few web-based servers for easy usage (Suhre & Sanejouand, 2004). Moreover, when the collective character of the protein motion is obviously the cause of a conformational change, then usually this is encompassed in a single low-frequency normal mode (Tama & Sanejouand, 2001; Delarue & Sanejouand, 2002; Krebs *et al.*, 2002).

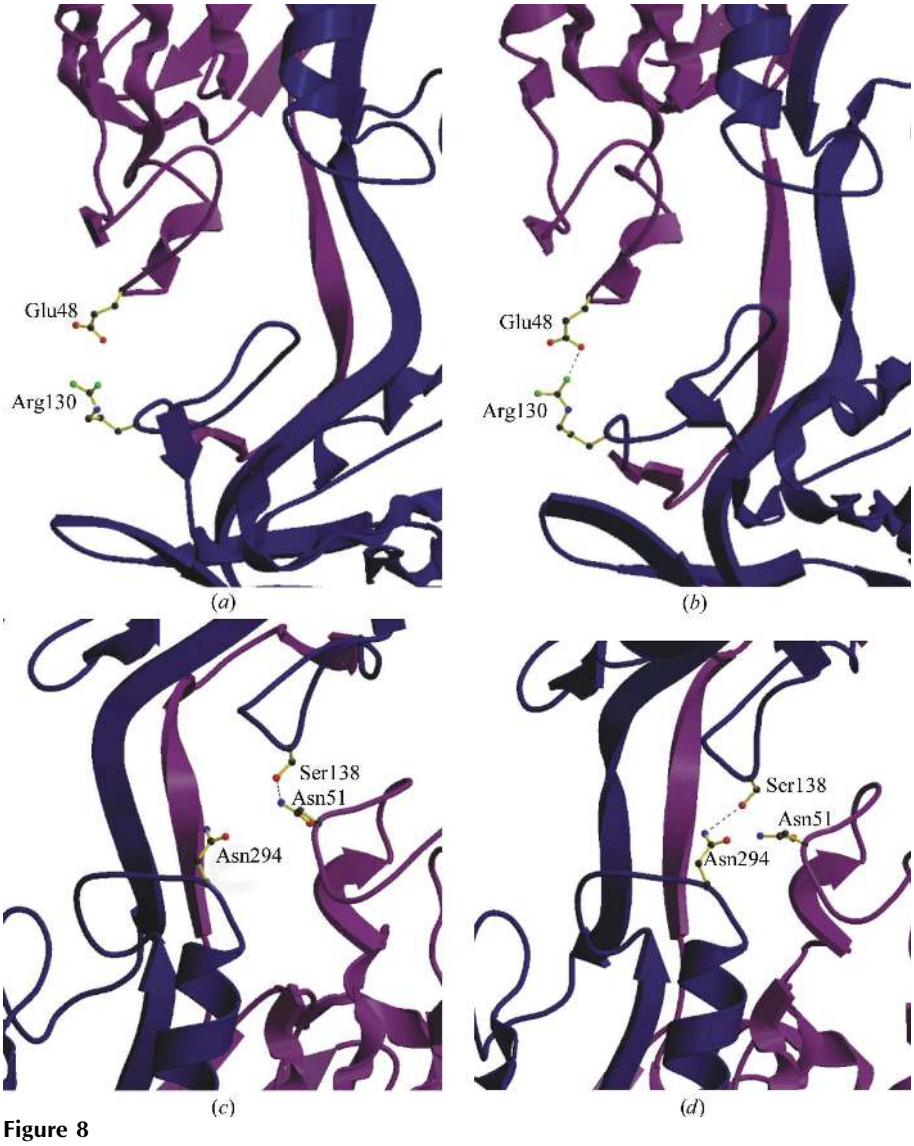

Figure 8
The interdomain interactions in *M. tuberculosis* TrxR during the large conformational change. The modelled interactions of the first structure of low-frequency mode 8 (*a*) do not show a salt bridge between the side chains Arg130 and Glu48, while the seventh structure of mode 8 (*b*) has a salt bridge. The first structure of low-frequency mode 8 (*c*) shows a hydrogen-bonding interaction between Ser138 and Asn51. This bond is lost and a new hydrogen bond is formed between Ser138 and Asn294 in the seventh structure of mode 8 (*d*). The two domains have been coloured differently. The figure was generated using *MOLSCRIPT* (Kraulis, 1991).

In order to examine the nature of the collective motions in F_O that might lead to the changes required for transition to F_R, the *ElNémo* server (http://igs-server.cnrs-mrs.fr/elnemo/) was used (Suhre & Sanejouand, 2004) to compute the 100 lowest frequency modes for TrxR. The low normal modes generally describe the most collective movements of a protein. When the structure of a protein is known in two different conformations, the overlap between the difference vectors between both conformations (in our case the difference vectors between the F_O and F_R conformations) and the normal modes can be used to identify which particular mode(s) contributes effectively in bringing about a conformational change (Tama & Sanejouand, 2001). We found that three dominating modes described most of the observed conformational change between both the forms (modes 8, 10 and 11). The largest contribution arises from mode 8, which alone contributes to 35% of the motions. Note that the first six modes in NMA are 'trivial' translation-rotation modes, so that mode 8 is the second non-trivial lowest frequency mode. When we examined the nature of the deformations along this mode, we observed that the NADPH domain rotates and brings the active site towards the surface of the molecule (Fig. 6). This suggests that the nature of the functional motions that bring about the essential transition in the molecule causing the F_O to F_R conformational change can be largely described by mode 8.

The dynamics and flexibility of the NADPH domain can further be inferred

by the *B* factor that can be derived from NMA. These are computed from the mean-square displacements of the 100 lowest frequency normal modes (Tama *et al.*, 2000). There is a close correlation between the computed and the observed (crystallographic) *B* factors (Fig. 2), which indicates that the normal modes are a realistic description of the protein flexibility. It also suggests that crystal-packing forces have little effect on the overall flexibility of the molecule. In particular, the NADPH domains possess higher *B* factors in both experiments and computations.

The covariance of motion between the C^{α} atoms shows three types of motions in the protein: interdomain, intradomain and a hinge motion (Fig. 7). It is clear that the movement is of whole domains as a coupled unit. They are largely positively correlated within themselves and anticorrelated relative to each other. However, there are also regions within each domain that move in an anticorrelated manner. Both types of motions contain contacts that are formed and broken. In this study of *M. tuberculosis* TrxR, the F_O to F_R transition involved anticorrelated motions that are associated with relative motions between the NADPH and FAD domain. This relative motion is associated with the formation of a salt bridge between the side chains of Arg130 and of Glu48 and of hydrogen-bond breakage between the main-chain carbonyl of Ile243 and the guanidine group of Arg117. Apart from this, the hydrogen bond between the side chain of Ser138 and the side chain of Asn51 was broken and a new bond was formed between the side chain of Ser138 and the side chain of Asn294. These were observed in the seven structures of mode 8 (Fig. 8). The formation of a hydrogen bond between the side chains of Asn51 and of Asn294 is associated with anticorrelated motion within the FAD domain, while anticorrelated motion within the NADPH domain resulted in the breakage of the hydrogen bond between the side chains of Glu160 and Thr137. These anticorrelated motions within domains that lead to the formation/breakage of hydrogen bonds are more like an unzipping motion.

The hinge region that is formed by a small β-sheet-like structure consists of hydrogen bonds between the main-chain amide of Ser115 and the main-chain carbonyl of Ser246, the carbonyl of Ser115 and the amide of Ser246, and the main-chain amide of Arg117 and the carbonyl of Gly247. These remain conserved and are quite stable to small perturbations that enable the movements between the two domains. In addition, there are two hydrogen bonds that form and break in this region during domain motion. Thus, the interdomain conformational changes present a complex picture, although the interactions at the hinge remain essentially the same.

4. Conclusions

The structure of *M. tuberculosis* TrxR reported in this study reveals the conformational heterogeneity of the molecule in the crystals. The anisotropy of the NADPH domain is clearly observed by TLS refinement, where the large libration axes are seen to be bisecting the two electron-transfer centres. Since the NADPH domains in both subunits are not involved in crystal contacts, the *B* factors possibly reflect the dynamic disorder of this domain. The lack of involvement of the NADPH domains in crystal contacts also suggests that the crystal-packing effects have little influence in representing the inherent disorder of the molecule. Libration of the NADPH domain is likely to bring the electron-transfer centres into close proximity to each other. The drastic reduction in *R* and $R_{\rm free}$ during TLS refinement indicates that modelling of the anisotropy in the domain can indeed reflect physiologically relevant conformational disorder. Normal-mode analysis confirms the results obtained by TLS refinement, in which the majority of changes are accounted for by one single mode. Normal-mode analysis thus allows us to track intermediate conformational changes that accompany transition between the F_O and the F_R conformational states. We have been able to obtain a glimpse of the interactions that are either made or broken during the conformational transition. Thus, TLS refinement combined with normal-mode analysis can reveal the conformational changes that have physiological relevance.

Recently, it has been reported that thioredoxin reductase is an important drug target (de Bernardo *et al.*, 2005). Since *M. tuberculosis* possesses only one copy of TrxR, its inhibition might lead to impaired growth of the bacillus. Thus, by understanding the conformational changes in *M. tuberculosis* TrxR, we believe that novel inhibitors of this enzyme may be designed.

We thank Stewart Cole and colleagues for providing a cosmid DNA library of *M. tuberculosis* H37Rv. We acknowledge the overall support of S. E. Hasnain and the support of the beamline staff at Elettra synchrotron, Trieste, Italy during diffraction and data collection. We gratefully acknowledge help from Nagasuma Chandra and the Bioinformatics facility of IISc Bangalore for analysis of flavoproteins. Financial support for this work was provided by the Department of Biotechnology. MA is a Council of Scientific and Industrial Research Senior Research Fellow and SCM is a Wellcome Trust International Senior Research Fellow. The authors declare that there are no competing financial interests.

References

Akif, M., Chauhan, R. & Mande, S. C. (2004). *Acta Cryst.* D**60**, 777–779.
Arner, E. S. J. & Holmgren, A. (2000). *Eur. J. Biochem.* **267**, 6102–6109.
Bauer, H., Massey, V., Arscott, L. D., Schirmer, R. H., Ballou, D. P. & Williams, C. H. Jr (2003). *J. Biol. Chem.* **278**, 33020–33028.
Bernardo, D. di, Thompson, M. J., Gardner, T. S., Chobot, S. E., Eastwood, E. L., Wojtovich, A. P., Elliot, S. J., Schaus, S. E. & Collins, J. J. (2005). *Nature Biotechnol.* **23**, 377–383.
Brooks, B. & Karplus, M. (1985). *Proc. Natl Acad. Sci. USA*, **82**, 4995–4999.
Chauhan, R. & Mande, S. C. (2002). *Biochem. J.* **367**, 255–261.
Cole, S. T. *et al.* (1998). *Nature (London)*, **393**, 537–544.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Delarue, M. & Sanejouand, Y. H. (2002). *J. Mol. Biol.* **320**, 1011–1024.
Fernando, M. R., Nanri, H., Yoshitake, S., Nagata-kuno, K. & Minikami, S. (1992). *Eur. J. Biochem.* **209**, 917–922.

Friemann, R., Schmidt, H., Ramsawamy, S., Frostner, M., Krauth-Siegel, R. L. & Eklund, H. (2003). *FEBS Lett.* **544**, 301–305.
Gerstein, M., Lesk, A. M. & Chothia, C. (1994). *Biochemistry*, **33**, 6739–6749.
Gromer, S., Urig, S. & Becker, K. (2004). *Med. Res. Rev.* **24**, 40–89.
Hayward, S. & Berendsen, H. J. C. (1997). *Proteins*, **27**, 425–437.
Isupov, M. N., Fleming, T. M., Dalby, A. R., Crowhurst, G. S., Bourne, P. C. & Littlechild, J. A. (1999). *J. Mol. Biol.* **291**, 651–660.
Jaeger, T., Budde, H., Flohe, L., Menge, U., Singh, M., Trujillo, M. & Radi, R. (2004). *Arch. Biochem. Biophys.* **423**, 182–191.
Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Krebs, W. G., Alexandrov, V., Wilson, C. A., Echols, N., Yu, H. & Gerstein, M. (2002). *Proteins*, **48**, 682–695.
Kuriyan, J., Krishna, T. S., Wong, L., Guenther, B., Pahler, A. & Williams, C. H. Jr (1991). *Nature (London)*, **352**, 172–174.
Laskowski, R. A., MacArthur, M. W., Moss, D. C. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lennon, B. W., Williams, C. H. Jr & Ludwig, M. L. (1999). *Protein Sci.* **8**, 2366–2379.
Lennon, B. W., Williams, C. H. Jr & Ludwig, M. L. (2000). *Science*, **289**, 1190–1194.
Lovell, S. C., Davis, I. W., Arendall, W. B. III, de Bakker, P. L., Word, J. M., Prisant, M. G., Richardson, J. S. & Richardson, D. C. (2003). *Proteins*, **50**, 437–450.
Marques, O. & Sanejouand, Y. H. (1995). *Proteins*, **23**, 557–560.
Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.
Monawad, L. & Perahia, D. (1996). *J. Mol. Biol.* **258**, 393–410.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Shinnick, T. M., King, H. & Quinn, F. D. (1995). *Am. J. Med. Sci.* **309**, 92–98.
Suhre, K. & Sanejouand, Y. H. (2004). *Nucleic Acids Res.* **32,** 10–14.
Tama, F., Gadea, F. X., Marques, O. & Sanejouand, Y. H. (2000). *Proteins*, **41**, 1–7.
Tama, F. & Sanejouand, Y. H. (2001). *Protein Eng.* **4**, 1–6.
Thomas, A., Field, M. J., Mouawad, L. & Perahia, D. (1996). *J. Mol. Biol.* **257**, 1070–1087.
Thomas, A., Field, M. J. & Perahia, D. (1996). *J. Mol. Biol.* **261**, 490–506.
Tirion, M. M. (1996). *Phys. Rev. Lett.* **77**, 1905- 1908.
Waksman, G., Krishna, T. S. R., Williams, C. H. Jr & Kuriyan, J. (1994). *J. Mol. Biol.* **236**, 800–816.
Wieles, B., Ottenhoff, T. H. M., Steenwijk, T. M., Franken, K. L. M. C., de Vries, R. R. P. & Langermans, J. A. M. (1997). *Infect. Immun.* **65**, 2537–2541.
Wieles, B., Soolingen, D., Holmgren, A., Offringa, R., Ottenhoff, T. & Thole, J. (1995). *Mol. Microbiol.* **16**, 921–929.
Williams, C. H., Arscott, L. D., Muller, S., Lennon, B. W., Ludwig, M. L., Wang, P. F., Veine, D. M., Becker, K. & Schirmer, R. H. (2000). *Eur. J. Biochem.* **267**, 6110–6117.
Yang, L. W. & Bahar, I. (2005). *Structure*, **13**, 893–904.
Zhang, Z., Hillas, P. J. & Ortiz de Montellano, P. R. (1999). *Arch. Biochem. Biophys.* **363**, 19–26.