Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The 1.70 Å X-ray crystal structure of *Mycobacterium tuberculosis* phosphoglycerate mutase

Peter Müller, Michael R. Sawaya, Inna Pashkov, Sum Chan, Chau Nguyen, Yim Wu, L. Jeanne Perry and David Eisenberg*

UCLA-DOE Institute for Genomics and Proteomics, Howard Hughes Medical Institute, Box 951570, Los Angeles, CA 90095-1570, USA

Correspondence e-mail: david@mbi.ucla.edu

Received 6 July 2004
Accepted 15 December 2004

PDB Reference: phosphoglycerate mutase, 1rii, r1riisf.

The single-crystal X-ray structure of phosphoglycerate mutase from *Mycobacterium tuberculosis* has been determined at a resolution of 1.70 Å. The C-terminal tail of each of the subunits is flexible and disordered; however, for one of the four chains (chain *A*) all but five residues of the chain could be modeled. Noteworthy features of the structure include the active site and a proline-rich segment in each monomer forming a short left-handed polyprolyl helix. These segments lie on the enzyme surface and could conceivably participate in protein–protein interactions.

1. Introduction

The protein with open reading frame code Rv0489 encoded by the genome of *Mycobacterium tuberculosis* is a cofactor-dependent phosphoglycerate mutase (dPGAM) and hence a member of a large family of highly conserved proteins that reversibly catalyze the transfer of a phosphate group from C3 to C2 of phosphoglycerate (as described by Fothergill-Gilmore & Watson, 1989). dPGAMs are found in all vertebrates, most invertebrates and some bacteria and fungi.

Several dPGAM structures have been determined. Among them are three structures of *Saccharomyces cerevisiae* dPGAM, two without the substrate at 2.3 and 2.1 Å resolution (PDB codes 4pgm and 5pgm; Rigden *et al.*, 1998, 1999) and one in complex with 3-phosphoglycerate at 1.7 Å (PDB code 1qhf; Crowhurst *et al.*, 1999). Another yeast dPGAM structure, an NMR-determined structure from *Schizosaccharomyces pombe*, has also been reported (PDB code 1ftz; Uhrinova *et al.*, 2001). In addition to these eukaryotic dPGAMs, structures of *Escherichia coli* dPGAM in the activated form at 1.25 Å (PDB code 1e58; Bond *et al.*, 2001) and in complex with the inhibitor vanadate at 1.3 Å (PDB code 1e59; Bond *et al.*, 2002) have been reported.

The proposed mechanism of dPGAM involves the formation of a phospho-histidine intermediate (Fothergill-Gilmore & Michels, 1993) and requires the presence of traces of the cofactor 2,3-bisphosphoglycerate. The C-terminal residues seem to be involved in the catalytic mechanism, regulating substrate release and shielding the active site (Garel *et al.*, 1989). This C-terminal tail is flexible and disordered and is not included in most dPGAM structures deposited in the PDB.

The structure of dPGAM from *M. tuberculosis* reported here was solved in the context of the structural genomics project on *M. tuberculosis* (Goulding *et al.*, 2002). With the ultimate goal of finding a cure for tuberculosis, a disease caused by a bacterium with which more than a third of the

world population is infected, the members of the *M. tuberculosis* consortium aim at solving as many structures of the *M. tuberculosis* genome as achievable.

2. Experiment

2.1. Cloning and expression

The Rv0489 gene was amplified by PCR using *M. tuberculosis* H37Rv genomic DNA as the template, forward primer GATC**CATATG**GCTGCAAACACTGGCAGCCTGGTGTTGCTGCGCCA, which introduced an *Nde*I site (bold), and reverse primer ATTT**GCGGCCGC**ACCTCGTGGTACACCCCCGCGGCCCTGGCCGGCCAC, which introduced a *Not*I site (bold). The forward primer also inserted an alanine codon (GCT) immediately following the start codon in order to enhance protein expression (Looman *et al.*, 1987), while the reverse primer introduced a thrombin-recognition sequence to the C-terminus. The PCR product was cloned into pCRBluntII-TOPO (Invitrogen). Following sequence confirmation, the gene was subcloned into pET22b (Stratagene), which added a hexahistidine tag to the expressed protein following the thrombin-recognition sequence.

The recombinant protein was expressed at approximately 10 mg l^{-1} according to the $OD_{600\,nm}$ in BL21-Gold(DE3) (Stratagene) *Escherichia coli* in enriched buffered LB medium (10 g NaCl, 40 g tryptone, 20 g yeast extract per litre of medium, 5% glycerol, 10 m*M* MOPS pH 7.0).

2.2. Purification and crystallization

In the purification, each gram of cell pellet was resuspended and lysed in 5 ml lysis buffer (20 m*M* Tris pH 8.0, 0.3 *M* NaCl, 10 m*M* imidazole, 2 m*M* β-mercaptoethanol, 2 µg ml^{-1} DNase I, 0.2 mg ml^{-1} lysozyme and 1:100 protease-inhibitor cocktail from Sigma). The lysate was clarified by centrifugation at 27 000*g* for 30 min. The soluble recombinant protein was initially purified using Ni–NTA Superflow resin (Qiagen). The peak elution fractions were pooled and further purified on a Superdex 75 column (Amersham Biosciences) equilibrated with 20 m*M* Tris pH 8.0, 0.1 *M* NaCl. The peak fractions were pooled and the buffer was exchanged to 20 m*M* Tris pH 8, 10 m*M* NaCl. The pure protein solution was concentrated to 33 mg ml^{-1} using a Centricon 10; the concentration was measured using the absorbance reading at 280 nm wavelength (extinction coefficient ε = 38 690 M^{-1} cm^{-1}).

Monoclinic crystals were obtained from this solution using 0.2 *M* $MgCl_2$, 20% PEG 3350 as precipitant in hanging-drop vapour-diffusion experiments at room temperature. The space group is $P2_1$ and the unit-cell parameters are a = 58.92 (6), b = 136.8 (14), c = 65.94 (7) Å, β = 97.78 (5)°. For transport and data collection, the crystals were cryogenically frozen in 1.5 *M* sodium formate solution containing 25% glycerol.

2.3. Data collection

A first in-house data set was collected to a resolution of 2.3 Å using a Cu rotating-anode source and a Rigaku R-AXIS IV image-plate detector. Later, X-ray diffraction data to 1.7 Å were collected at ALS beamline 8.2.2 in Berkeley using an ADSC Quantum 315 3 × 3 CCD array. 300 1.0° oscillation frames were collected at a wavelength of 1.0000 Å for both the high-resolution data and the inner data. Data reduction was performed using *DENZO* (Otwinowski & Minor, 1997) and scaling was performed with the program *SADABS* (Sheldrick, 2003).

2.4. Structure determination and refinement

The structure was solved by molecular replacement using the program *GLRF* (Tong & Rossmann, 1997) with one homotetramer of the *S. cerevisiae* phosphoglycerate mutase (PDB code 5pgm) as the search model. The *S. cerevisiae* structure was chosen as a search model rather than the *E. coli* structure as the similarity between the dPGAM amino-acid sequences of *S. cerevisiae* and *M. tuberculosis* is higher than between the two bacterial sequences; sequence alignments performed online with the program *CLUSTALW* (Thompson *et al.*, 1994)

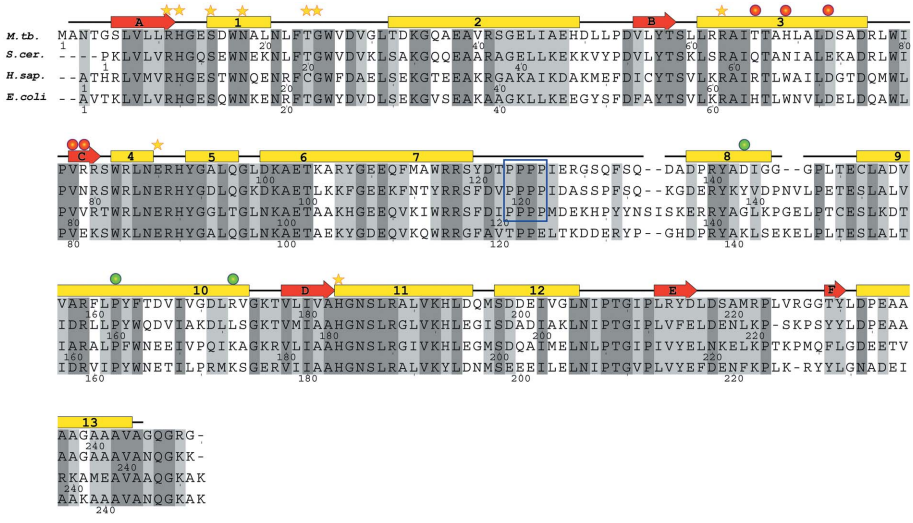

Figure 1
Sequence alignment (performed online with the program *CLUSTALW*; Thompson *et al.* 1994) comparing the phosphoglycerate mutases from *M. tuberculosis* with those from *S. cerevisiae*, *E. coli* and *H. sapiens*. Perfect conservation throughout the four organisms is indicated by a dark gray background, while similarity is indicated by a light gray background. The secondary-structure elements are given on top of the four sequences (α-helices in yellow, β-strands in red); the α-helices have been labeled from 1 to 13 and the β-strands from *A* to *F* (analogous to Fig. 2). The $(Pro)_4$ motif (as shown in Fig. 5) is marked with a box around the residues. Residues that are part of the active site (as shown in Fig. 4) are labeled with a yellow star, residues involved in the formation of the dimer interface (as shown in Figs. 3*a* and 3*b*) are marked with a red ball and those involved in the formation of the tetramer interface (as shown in Figs. 3*c* and 3*d*) are marked with a green ball. This figure was prepared using the program *SECSEQ* (D. E. Brodersen; http://xray.imsb.au.dk/~deb/secseq).

gave 63% identity between *S. cerevisiae* and *M. tuberculosis* and only 46% between *E. coli* and *M. tuberculosis*. Interestingly, the similarity between the sequence of the human dPGAM and the *M. tuberculosis* homolog is slightly higher than the similarity between the two bacterial sequences (48%). The sequence alignment is shown in Fig. 1.

Initial refinement steps were performed with *CNS* (Brünger *et al.*, 1998) against the in-house data to 2.3 Å. With the 1.7 Å synchrotron data, further refinement was performed with *SHELXL* (Sheldrick & Schneider, 1997) using all data from 58.37 to 1.70 Å and applying a bulk-solvent correction. No non-crystallographic symmetry restraints were used as the data-to-parameter ratio at 1.7 Å was good enough to refine the four chains independently. Automatic water divination was performed with *SHELXPRO* (Sheldrick & Schneider, 1997) and after each refinement step the model was visually inspected in *XtalView* (McRee, 1999) using both $2F_o - F_c$ and $F_o - F_c$ difference maps. For the discretely disordered side chains of 11 residues a second conformation was modeled. The relative occupancies were refined freely using one free variable per disordered residue. All H atoms connected to C atoms and backbone N atoms were included at their geometrically calculated positions and refined using a riding model. H atoms bound to N atoms on histidine side chains were included in the model after careful inspection of the hydrogen-bonding patterns around these residues. H atoms connected to O or side-chain N atoms other than those from histidine were not included in the model, as doing so can create ambiguities (two H atoms along the same hydrogen bond) and problems with the automatic application of anti-bumping restraints. The coordinates of the final model and the merged structure factors have been deposited with the Protein Data Bank (Bernstein *et al.*, 1977).

Table 1
Statistics of X-ray data collection and atomic refinement.

Values in parentheses refer to the outer shell of data.

Space group	$P2_1$
Unit-cell parameters (Å, °)	a = 58.92 (6), b = 136.8 (14), c = 65.94 (7), β = 97.78 (5)
Resolution range (Å)	58.37–1.70 (1.80–1.70)
R_{int}† (%)	6.99 (45.5)
R_σ‡ (%)	3.67 (32.0)
MoO§	6.48 (3.72)
No. of unique data	111293 (16353)
Completeness of data (%)	98.2 (91.8)
$I/\sigma(I)$	18.1 (3.2)
No. of residues (four chains per AU)	945
No. of protein atoms	7365
No. of solvent atoms	481
Matthews coefficient¶ (Å^3 Da^{-1})	2.51
R†† (%)	21.9
R_{free}‡‡ (%)	27.0
Test-set size (%), selection	2.5, random
R.m.s. deviations from target values	
Bond lengths (Å)	0.006
Angle distances (Å)	0.022
Planar groups (Å)	0.025
Chiral volumes (Å)	0.038
Anti-bumping distances (Å)	0.007
Average B factor for main-chain atoms (Å^2)	32.1
Average B factor for side-chain atoms (Å^2)	35.4
R.m.s. B for main-chain atoms (Å^2)	1.3
R.m.s. B for side-chain atoms (Å^2)	2.3

† $R_{int} = \sum|F_o^2 - \langle F_o^2\rangle|/\sum F_o^2$, where F_o is the observed structure factor. Both summations involve all input reflections for which more than one symmetry equivalent is averaged. ‡ $R_\sigma = \sum[\sigma(F_o^2)]/\sum F_o^2$ over all reflections in the merged list. § MoO: multiplicity of observations. This term was defined at the *SHELX* workshop in Göttingen in September 2003 to distinguish the MoO from redundancy or multiplicity, with which the MoO has frequently been confused in the past. In contrast to redundancy, which is repeated recording of the same reflection obtained from the same crystal orientation (performing scans that rotate the crystal by more than 360°), MoO, which is sometimes also referred to as the 'true redundancy', describes multiple measurements of the same reflection obtained from different crystal orientations (*i.e.* measured at different Ψ angles). ¶ Matthews coefficient as defined by Matthews (1968). †† $R = \sum\big||F_o| - |F_c|\big|/\sum|F_o|$. ‡‡ R_{free} as defined by Brünger (1992).

3. Results and discussion

3.1. The model

The model of *M. tuberculosis* dPGAM was refined to a crystallographic R factor of 21.9% (free R factor = 27.0%). The refined model consists of four dPGAM subunits (chains *A*–*D*), four fully occupied glycerol molecules, one 50% occupied glycerol molecule and 451 fully occupied waters per asymmetric unit. Glycerol is structurally similar to both the substrate (3-phosphoglycerate) and the cofactor (2,3-bisphosphoglycerate) of dPGAM and three of the glycerol molecules are actually bound to the active sites of chains *A*, *B* and *D*. Why there is no glycerol located in the active site of chain *C* is unclear.

Table 1 provides statistical information about the quality of data, structure and refinement. 91% of the residues are in the most favored regions of the Ramachandran plot, 8.3% and 0.7% fall into the additionally allowed and generously allowed regions, respectively, and none fall into the disallowed region (data obtained using *PROCHECK*; Laskowski *et al.*, 1993). Analysis with the program *ERRAT* (Colovos & Yeates, 1993) gives an overall quality factor of 94.6%, which is good for a 1.7 Å resolution structure.

3.2. The fold

The crystal structure of *M. tuberculosis* dPGAM reveals essentially the same α/β-fold as the dPGAM structures from *S. cerevisiae* (Rigden *et al.*, 1998, 1999; Crowhurst *et al.*, 1999), *E. coli* (Bond *et al.*, 2001, 2002) and *S. pombe* (Uhrinova *et al.*, 2001), as could be anticipated from their high sequence similarity (Fig. 1). The 265-residue monomer (including a C-terminal His tag, which is not visible in the electron density) consists of a six-stranded β-sheet, *C*–*B*–*D*–*A*–*E*–*F*, with all strands but *E* being parallel. It is flanked by five α-helices: helices 2 and 3 on one side and helices 9, 10 and 11 on the other (Fig. 2). Helices 9 and 10 are fused and can also be seen as one large slightly bent helix. Helices 5–8 compose a smaller domain (residues 91–150), part of which is structurally divergent (residues 127–147) and appears to be involved in tetramerization, a property not shared by all dPGAMs (see below).

3.2.1. Structural comparisons. Least-squares superpositions of the C$^\alpha$ coordinates of the dPGAM monomers reveal a closer resemblance between the *M. tuberculosis* and

S. cerevisiae structures than between the *M. tuberculosis* and *E. coli* structures, with r.m.s.d. values of 0.8 Å (228 C^{α} atoms) and 1.1 Å (232 C^{α} atoms), respectively. This result could be anticipated from the sequence identity; however, from a phylogenetic standpoint it is surprising that an *M. tuberculosis* structure would resemble a eukaryotic homolog more closely than a eubacterial homolog. A similar trend has been observed for several other *M. tuberculosis* proteins (Chan *et al.*, 2004).

A somewhat more distant homolog of *M. tuberculosis* dPGAM, human bisphosphoglycerate mutase (bPGAM), shares 44% sequence identity (Wang *et al.*, 2004), but has synthase and phosphatase activity in addition to mutase activity. *M. tuberculosis* dPGAM superimposes on human bPGAM with an r.m.s.d. of 1.2 Å (196 C^{α} atoms).

3.2.2. The quaternary structure. The asymmetric unit contains four subunits. Structural superpositioning of chains *A*, *B*, *C* and *D* in all pairwise combinations shows that the secondary structure of the four chains is practically identical apart from a helix–loop–helix segment (residues 99–120) which appears in two different conformations, one found in chains *A* and *B* and the other in chains *C* and *D*. All results of the superpositions are given in Table 2.

Table 2
All-C^{α} superposition of the four chains of the final model onto one another.

Chains fitted	Fitting all residues (Å)	Excluding residues 99–120 (Å)
A onto *B*	0.43	0.42
A onto *C*	1.06	0.39
A onto *D*	1.15	0.57
B onto *C*	1.14	0.44
B onto *D*	1.17	0.50
C onto *D*	0.51	0.50

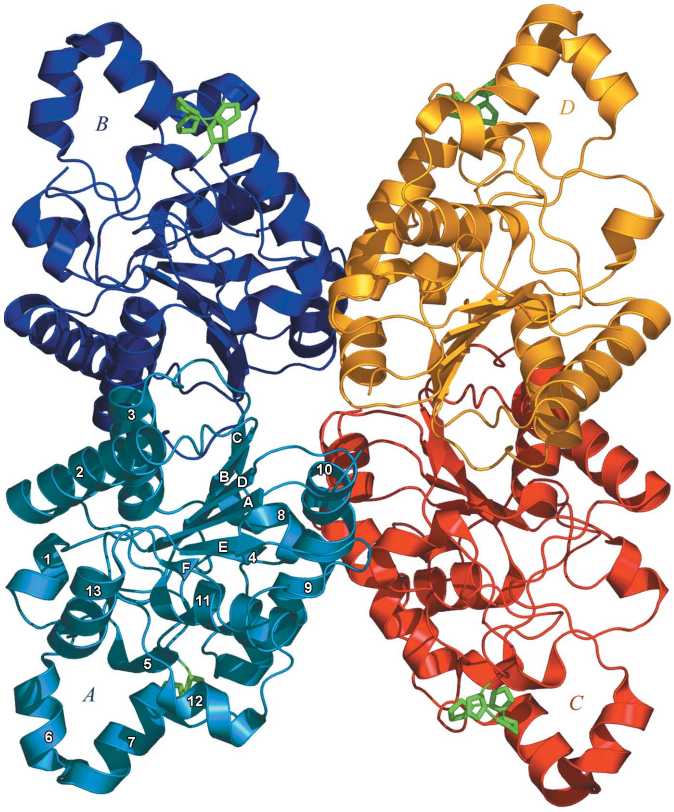

Figure 2
Homotetramer of *M. tuberculosis* phosphoglycerate mutase, possessing pseudo-222 symmetry. The view is along one of the pseudo-twofold axes; the other two run vertically and horizontally. Chains *A* and *B* (light and dark blue) form one dimeric subunit and chains *C* and *D* (red and orange) form the other dimer. The interfaces are shown in detail in Fig. 3. The four Pro residues forming a short left-handed poly-Pro helix are shown in green (see also Fig. 5). For chain *A* the α-helices have been labeled from 1 to 13 and the β-strands from *A* to *F*. This figure, as well as all molecular graphics in Figs. 3, 4 and 5, was prepared using the program *PyMol* (DeLano, 2002).

The appearance of four monomers in the asymmetric unit suggests that the biologically relevant form of *M. tuberculosis* dPGAM is a tetramer. The tetramer is organized as a dimer of dimers with approximate 222 symmetry (Fig. 2), very similar to that seen in the *S. cerevisiae* dPGAM structure. Mutases from other sources appear to be subsets of a tetramer; *E. coli* dPGAM and human bPGAM are dimers resembling half of the *S. cerevisiae* tetramer and *S. pombe* dPGAM is a monomer.

There are two types of interfaces in the tetramer; one type is larger than the other. The larger type is called the dimer interface (*i.e.* the interface between chains *A* and *B* and the interface formed between chains *C* and *D*) and the smaller type is called the tetramer interface (*i.e.* the interface between chains *A* and *C* and the interface between chains *B* and *D*). The dimer interface shares strong structural similarity to the dimer interface found in the *E. coli* (Bond *et al.*, 2001) and the *S. cerevisiae* structures (Rigden *et al.*, 1999), even though four of the five interacting residues are not conserved (see Fig. 1). This interface is formed by two twofold-symmetric contacts between subunits: one between strands *C* (Fig. 3*a*) and the other between helices 3 (Fig. 3*b*).

The tetramer interface is also formed by twofold-symmetric interactions. The core of the interface involves residues Asp166, Arg87 and Pro162, all of which are identical in *S. cerevisiae* (Fig. 3*c*). These residues are not conserved in the *E. coli* enzyme and this lack of conservation has been cited as the reason why the *E. coli* enzyme does not form a tetramer (Bond *et al.*, 2001).

Although the *M. tuberculosis* and *S. cerevisiae* tetramers are assembled similarly, the periphery of the tetramer interface is not as well conserved as the core. A two-residue deletion at the end of helix 8 in the *M. tuberculosis* enzyme changes its conformation relative to the *S. cerevisiae* enzyme. Thus, the salt bridge between Asp143 on helix 8 and Arg173 of helix 10 (from the second dimer) has no equivalent in the *S. cerevisiae* enzyme. Instead, this segment of the interface is mediated by van der Waals contacts between Tyr139 and Leu171.

3.3. The active site

As shown in Fig. 1, the residues forming the catalytic center are highly conserved in other phosphoglycerate mutases and they also possess very similar geometry, at least when compared with the *S. cerevisiae* structure. The active site is

situated at the C-terminal edge of the β-sheet and is composed of residues throughout the sequence (as marked in Fig. 1). A web of hydrogen bonds is found around the central catalytic histidine residues His12 and His183 involving other residues (Arg11, Ser15, Asn18, Thr24, Gly25, Arg63 and Glu90) that have previously been suggested to be involved in the catalytic mechanism (Rigden *et al.*, 1998 and references therein), as shown in Fig. 4.

A glycerol molecule bound in the active site elicits some side-chain conformations that resemble the phospho-histidine (activated) enzyme and other side-chain conformations which resemble the vanadate (inhibited) enzyme structure (PDB codes 1e58 and 1e59; Bond *et al.*, 2001, 2002). Glycerol shares structural similarity with the substrate, cofactor and product of the catalytic reaction and so might be expected to elicit some of the same conformational changes as these ligands. Although this similarity is limited by the absence of a carboxylate or phosphate group, it is evidently sufficient to permit direct interactions with four active-site residues (Fig. 4). Activated and inhibited forms of the *E. coli* homolog can be distinguished by different conformations in three residues: Arg11, His12 and Asn18. Only one of these residues, Asn18, is a different rotamer in the activated structure; the other two residues are only slightly displaced. The active-site residues in the dPGAM structure of *M. tuberculosis* superimpose very well with the equivalent residues in the *E. coli* structures. While residues 11 and 12 of the *M. tuberculosis* structure appear to be in the inactive conformation, Asn18 is clearly the same rotamer as observed in the structure of the activated form of *E. coli* dPGAM. Asn18 is held in place by two water molecules and the glycerol molecule, as shown in Fig. 4. This finding suggests that the glycerol molecule is in some way mimicking the activating cofactor. The glycerol-binding geometry might be more credible than the 3-phosphoglycerate ligand observed in the *S. cerevisiae* structure (PDB code 1qhf; Crowhurst *et al.*, 1999), which is 4 Å from the glycerol site and appears to be bound with less specificity.

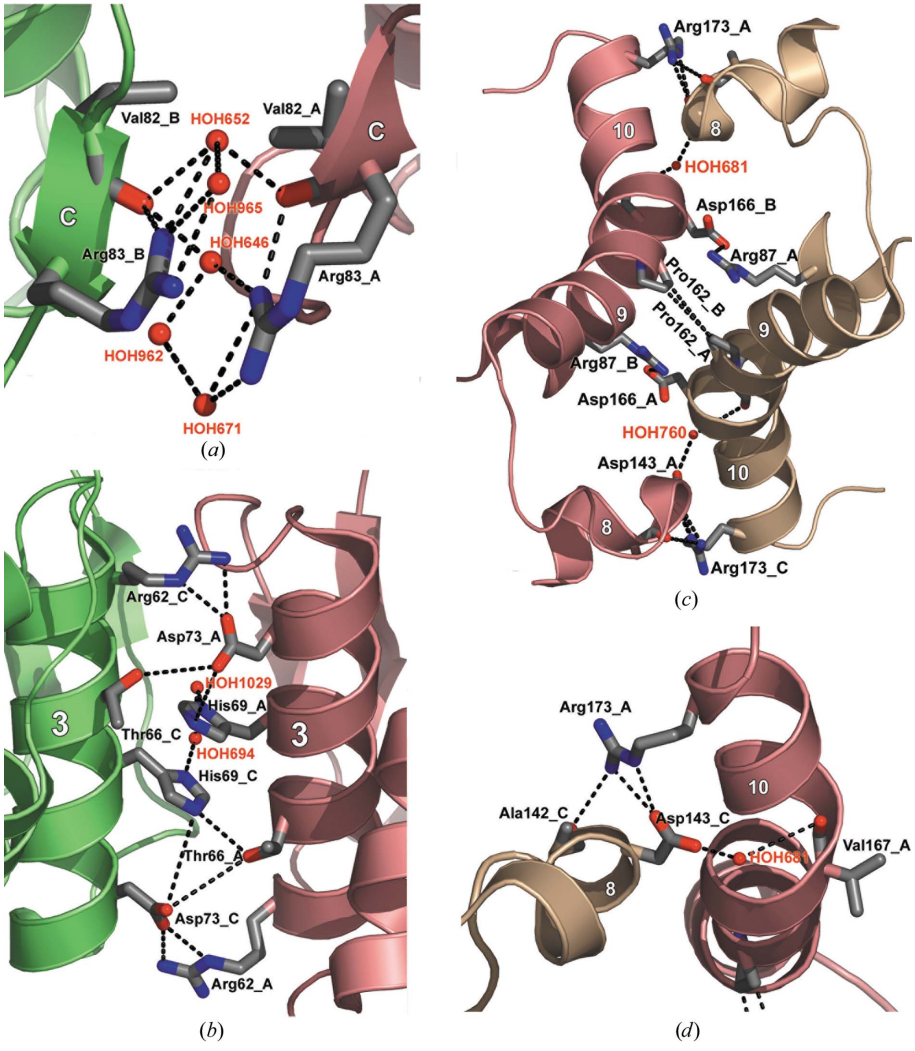

Figure 3
Hydrogen bonds and van der Waals interactions forming the oligomerization interfaces of the *M. tuberculosis* dPGAM tetramer. Chain *A* is shown in pink, chain *B* in green and chain *C* in gold. The dimer interface, as shown in (*a*) and (*b*), involves 12 amino acids and seven water molecules. It is more extensive than the tetramer interface (*c* and *d*), which involves ten amino acids and two water molecules. All amino-acid residues contributing to the interfaces are also marked in the sequence alignment (Fig. 1).

3.4. The C-terminus

The C-terminal tails of the subunits are disordered; only for chain *A* could most of the C-terminal residues be modelled (in our model chain *A* ends with Ala244, chain *B* with Gly239, chain *C* with Ala238 and chain *D* with Pro233). The additional C-terminal residues in chain *A* extend the length of the C-terminal helix (helix 13) an additional two turns beyond that observed in the dPGAM structure of *S. cerevisiae* (*e.g.* Rigden *et al.*, 1998). The five C-terminal residues, however, show rather high *B* values (over 60 Å^2) and it can be assumed that the C-terminus is only partially ordered. For this reason, additional residues were added to the model only if doing so improved both the working *R* factor and $R_{\rm free}$. Fig. 5 shows the C-terminal region of chain *A* superimposed onto an $F_{\rm o} - F_{\rm c}$ σ_A-weighted omit map.

Mutational studies on the *S. cerevisiae* homolog (Walter & Fothergill-Gilmore, 1995) suggest the involvement of the C-terminal residues in the catalytic mechanism and, indeed, the C-terminus points towards the entry of the catalytic site (Fig. 4). It has been hypothesized that helix 13 and the following C-terminal residues could close the active site like a lid to support

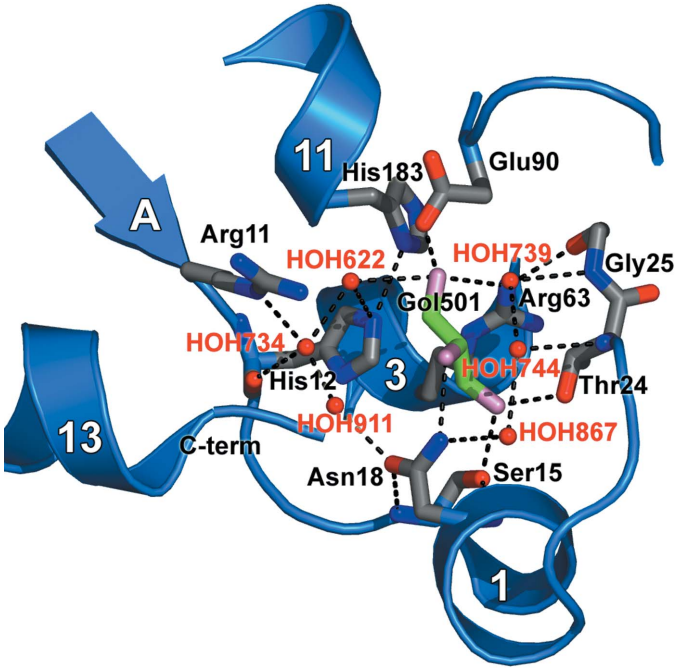

Figure 4
View of the active site of chain *A*. Hydrogen bonds in the active site of *M. tuberculosis* dPGAM, involving one glycerol molecule and four waters. The C-terminal helix, helix 13, points towards the entrance of the active site.

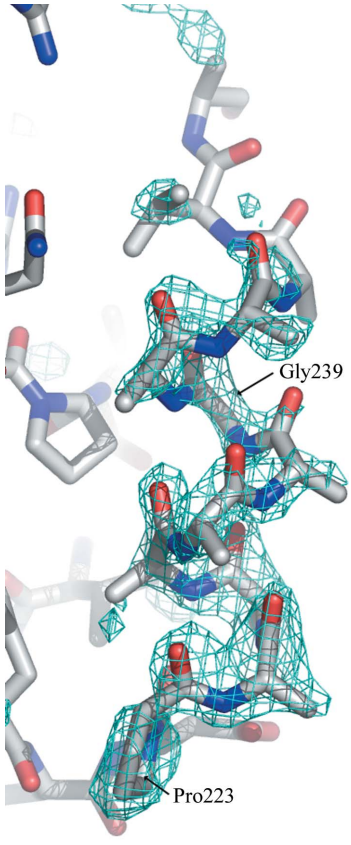

Figure 5
A σ_A-weighted $F_o - F_c$ omit map superimposed onto the C-terminal residues of chain *A*. The α-helical arrangement of the amino acids in the final model is supported by the electron density.

and shield the catalytic reaction when a substrate molecule is bound. This hypothesis is in agreement with the location of the C-terminal residues observed in the structure of the phosphohistidine activated dPGAM of *E. coli* (Bond *et al.*, 2001; PDB code 1e58), the only structure in the PDB that shows the complete C-terminus. In this activated dPGAM structure, a hydrogen bond between Asn16 and the phosphate moiety is proposed to trigger a conformational change in loop residues 9–21 that enables interaction with the C-terminal tail, thus ordering it. By analogy, partial ordering of the C-terminal tail in the *M. tuberculosis* structure might be the result of the participation of Asn18 in a hydrogen bond with glycerol.

3.5. The $(Pro)_4$ motif

dPGAM contains a PPPP motif from residues 123–126 in each monomer, forming a short left-handed poly-prolyl helix which is solvent-exposed at the surface of the molecule (Figs. 1, 2 and 6). Because prolyl residues are found in the primary structures of many protein–protein interfaces, we need to ask whether this segment is involved in a protein–protein inter-

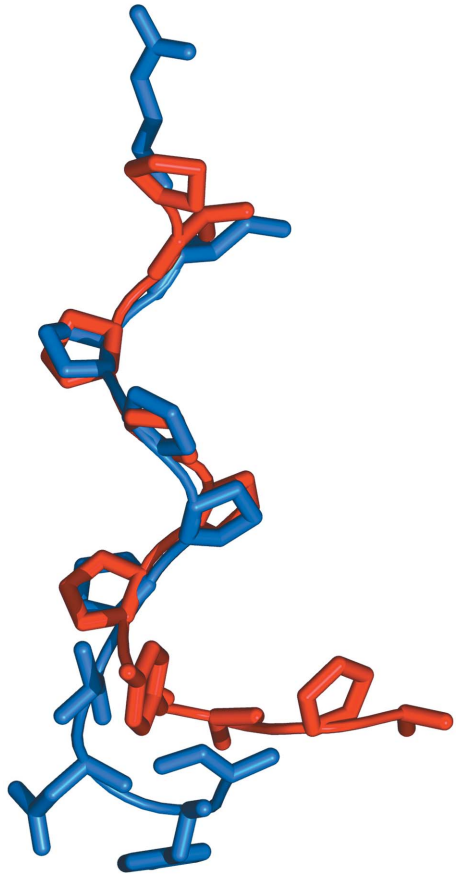

Figure 6
Least-squares fit of the proline-rich stretch of *M. tuberculosis* dPGAM (residues 121–130) in blue onto a proline-rich peptide known to bind SH3 domains (taken from the structure with PDB code 1fyn; Musacchio *et al.*, 1994) in red. The two $(Pro)_4$ motifs overlap very well (r.m.s.d. = 0.73 Å for all atoms). This similarity supports the assumption that the four consecutive proline residues may be involved in protein–protein interactions.

action. SH3 and WW domains tend to interact with ligands that are proline-rich (Kay *et al.*, 2000).

A *DASEY* search (Mallick *et al.*, 2002) for WW and SH3 domains within the genome of *M. tuberculosis* did not result in any significant assignments. For *S. cerevisiae*, however, a physical interaction between dPGAM and fructose-bisphosphate aldolase has been described previously (Tong *et al.*, 2002). As both proteins are in the same pathway and the sequence similarities between the two *S. cerevisiae* proteins and their *M. tuberculosis* homologs are significant (63 and 38%, respectively), it is conceivable that the homologous proteins in *M. tuberculosis*, Rv0489 and Rv0363c, also form a complex.

It should be pointed out that although the sequences of several dPGAMs (such as *M. tuberculosis, Homo sapiens* and *S. cerevisiae*) contain a PPPP motif, the sequences of many other dPGAMs lack this motif. *E. coli*, for example, has only two consecutive proline residues in its dPGAM sequence.

The authors thank Dr John T. Belisle, Colorado State University and NIH, NIAID Contract No. 1 AI-75320 for the generous supply of *M. tuberculosis* H37Rv genomic DNA, Dr Daniel Anderson for proofreading and constructive criticism, Sul-Min Kim, Robert Riley and Dr Matteo Pellegrini for their expertise in bioinformatics, and Dr Duilio Cascio and the 8.2.2 staff for assistance at the beamline. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences Division of the US Department of Energy under Contract No. DE-AC03-76SF00098 at Lawrence Berkeley National Laboratory. This work was supported by the NIH, the DOE and the German Research Association (DFG Forschungsstipendium No. 227812).

References

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.

Bond, C. S., White, M. F. & Hunter, W. N. (2001). *J. Biol. Chem.* **276**, 3247–3253.

Bond, C. S., White, M. F. & Hunter, W. N. (2002). *J. Mol. Biol.* **316**, 1071–1081.

Brünger, A. T. (1992). *Nature (London)*, **35**5, 472–474.

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.

Chan, S., Segelke, B., Lekin, T., Krupka, H., Cho, U. S., Kim, M. Y., So, M., Kim, C. Y., Naranjo, C. M., Rogers, Y. C., Park, M. S., Waldo, G. S., Pashkov, I., Cascio, D., Perry, J. L. & Sawaya, M. R. (2004). *J. Mol. Biol.* **341**, 503–517.

Colovos, C. & Yeates, T. O. (1993). *Protein Sci.* **2**, 1511–1519.

Crowhurst, G. S., Dalby, A. R., Isupov, M. N., Campbell, J. W. & Littlechild, J. A. (1999). *Acta Cryst.* D**55**, 1822–1826.

DeLano, W. L. (2002). *PyMOL.* DeLano Scientific, San Carlos, CA, USA.

Fothergill-Gilmore, L. A. & Michels, P. A. M. (1993). *Prog. Biophys. Mol. Biol.* **59**, 105–235.

Fothergill-Gilmore, L. A. & Watson, H. C. (1989). *Adv. Enzymol. Related Areas Mol. Biol.* **62**, 227–313.

Garel, M. C., Joulin, V., Le Boulch, P., Calvin, M. C., Prehu, M. O., Arous, N., Longin, R., Rosa, R., Rosa, J. & Cohen-Solal, M. (1989). *J. Biol. Chem.* **264**, 18966–18972.

Goulding, C. W. *et al.* (2002). *Curr. Drug Targets Infect. Disord.* **2**, 121–141.

Kay, B. K., Williamson, M. P. & Sudol, M. (2000). *FASEB J.* **14**, 231–241.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Looman, A. C., Bodlaender, J., Comstock, L. J., Eaton, D., Jhurani, P., de Boer, H. A. & Knippenberg, P. H. (1987). *EMBO J.* **6**, 2489–2492.

McRee, D. E. (1999). *J. Struct. Biol.* **125**, 156–165.

Mallick, P., Weiss, R. & Eisenberg, D. S. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 16041–16046.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Musacchio, A., Saraste, M. & Wilmanns, M. (1994). *Nature Struct. Biol.* **1**, 546–551.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Rigden, D. J., Alexeev, D., Phillips, S. E. V. & Fothergill-Gilmore, L. A. (1998). *J. Mol. Biol.* **276**, 449–459.

Rigden, D. J., Walter, R. A., Phillips, S. E. V. & Fothergill-Gilmore, L. A. (1999). *J. Mol. Biol.* **286**, 1507–1517.

Sheldrick, G. M. (2003). *SADABS.* University of Göttingen, Germany.

Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.

Thompson, J., Higgins, D. & Gibson, T. (1994). *Nucleic Acids Res.* **22**, 4673–4680.

Tong, A. H., Drees, B., Nardelli, G., Bader, G. D., Brannetti, B., Castagnoli, L., Evangelista, M., Ferracuti, S., Nelson, B., Paoluzi, S., Quondam, M., Zucconi, A., Hogue, C. W. V., Fields, S., Boone, C. & Cesareni, G. (2002). *Science*, **295**, 321–324.

Tong, L. & Rossmann, M. G. (1997). *Methods Enzymol.* **276**, 594–611.

Uhrinova, S., Uhrin, D., Nairn, J., Price, N. C., Fothergill-Gilmore, L. A. & Barlow, P. N. (2001). *J. Mol. Biol.* **306**, 275–290.

Walter, G. A. & Fothergill-Gilmore, L. A. (1995). *Biopolymers and Bioproducts: Structure, Function and Applications*, edited by M. R. J. Svasti, pp. 369–375. Bangkok: Samakkhisan Co. Ltd.

Wang, Y., Wei, Z., Bian, Q., Cheng, Z. Wan, M., Liu, L. & Gong, W. (2004). *J. Biol. Chem.* **279**, 39132–39138.