Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structures of native, substrate-bound and inhibited forms of *Mycobacterium tuberculosis* imidazoleglycerol-phosphate dehydratase

Mohammad Syed Ahangar, Rajan Vyas,‡ Nazia Nasir and Bichitra K. Biswal*

Protein Crystallography Laboratory, National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi 110 067, India

‡ Current address: Department of Biochemistry, Ohio State University, 484 West 12th Avenue, Columbus, OH 43210, USA.

Correspondence e-mail: bbiswal@nii.res.in

Received 18 July 2013
Accepted 12 August 2013

PDB References: IGPD, native, 4gqu; complex with IGP, 4lom; complex with 3-amino-1,2,4-triazole, 4lpf

Imidazoleglycerol-phosphate dehydratase (IGPD; HisB), which catalyses the conversion of imidazoleglycerol-phosphate (IGP) to imidazoleacetol-phosphate in the histidine biosynthesis pathway, is absent in mammals. This feature makes it an attractive target for herbicide discovery. Here, the crystal structure of *Mycobacterium tuberculosis* (*Mtb*) IGPD is reported together with the first crystal structures of substrate-bound and inhibited (by 3-amino-1,2,4-triazole; ATZ) forms of IGPD from any organism. The overall tertiary structure of *Mtb* IGPD, a four-helix-bundle sandwiched between two four-stranded mixed β-sheets, resembles the three-dimensional structures of IPGD from other organisms; however, *Mtb* IGPD possesses a unique structural feature: the insertion of a one-turn 3_{10}-helix followed by a loop ten residues in length. The functional form of IGPD is 24-meric, exhibiting 432 point-group symmetry. The structure of the IGPD–IGP complex revealed that the imidazole ring of the IGP is firmly anchored between the two Mn atoms, that the rest of the substrate interacts through hydrogen bonds mainly with residues Glu21, Arg99, Glu180, Arg121 and Lys184 which protrude from three separate protomers and that the 24-mer assembly contains 24 catalytic centres. Both the structural and the kinetic data demonstrate that the inhibitor 3-amino-1,2,4-triazole inhibits IGPD competitively.

1. Introduction

The histidine biosynthesis pathway, leading to the enzymatic synthesis of histidine from 5-phosphoribosyl-1-pyrophosphate in ten enzymatic steps carried out by 11 enzymes (Supplementary Fig. S1[1]), is conserved in bacteria, lower eukaryotes and plants, but is absent in mammals (Ames *et al.*, 1960; Alifano *et al.*, 1996; Stepansky & Leustek, 2006). Decades of genetic, molecular and biochemical research on this pathway have resulted in the unravelling of many fundamental mechanisms of biology, particularly in the area of operon theory (Alifano *et al.*, 1996). The enzymes of this pathway are conserved across many different organisms that synthesize histidine *de novo*. However, subtle differences exist in some of the catalytic steps, particularly the fifth and the sixth steps. The fifth step, the conversion of phosphoribosyl formimino-5-aminoimidazole-4-carboxamide ribonucleotide to imidazole-glycerol-phosphate (IGP) and aminoimidazole-carboxamide-ribonucleotide (Supplementary Fig. S1), requires two catalytic functions: transfer of an amide from glutamine followed by a cyclization reaction. In prokaryotes, these reactions are

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: MV5093). Services for accessing this material are described at the back of the journal.

catalyzed by two different monofunctional enzymes, HisH and HisF, respectively, which form a stable complex (Klem & Davisson, 1993), whereas in plants and yeasts this step is catalyzed by a bifunctional enzyme exhibiting both glutamine amidotransferase and cyclase activities (Fujimori & Ohta, 1998; Brilli & Fani, 2004). The scenario is very similar in the sixth step, the dehydration of IGP to imidazoleacetol-phosphate (IAP). In some prokaryotes such as *Escherichia coli* (Chiariotti *et al.*, 1986), *Salmonella typhimurium* (Staples & Houston, 1979) and *Azospirillum brasilense* (Bazzicalupo *et al.*, 1987) this step is catalysed by the dehydratase domain of a bifunctional enzyme (encoded by the *hisB* gene) that possesses both dehydratase and phosphatase activities. The phosphatase activity of this bifunctional enzyme is required to catalyse the eighth step in the pathway, the conversion of L-histidinol-phosphate to L-histidinol. On the other hand, in organisms such as *Mycobacterium tuberculosis* (Cole *et al.*, 1998), fungi (Parker *et al.*, 1994) and plants (Mano *et al.*, 1993; Tada *et al.*, 1994) the sixth and eighth steps are catalyzed by two separate monofunctional enzymes: imidazoleglycerol-phosphate dehydratase (IGPD) and histidinol-phosphate phosphatase, respectively.

IGPD (also known as HisB), one of the most extensively studied enzymes of the histidine biosynthesis pathway, is a recognized target for herbicide discovery as mammals do not possess an IGPD. In fact, triazole and its derivatives as well as carboxamido phosphonates and diazafulvene mimics have been shown to be specific inhibitors of IGPD (Mori *et al.*, 1995; Tada *et al.*, 1995; Schweitzer *et al.*, 1999, 2002; Gohda *et al.*, 1998). Importantly, over the years the understanding of the structural aspects of this enzyme has greatly advanced with the availability of its crystal structure from various organisms such as *Filobasidiella neoformans* (Sinha *et al.*, 2004), *Arabidopsis thaliana* (Glynn *et al.*, 2005) and *Staphylococcus aureus* (PDB entry 2ae8; Midwest Center for Structural Genomics, unpublished work). However, to date no experimentally determined ligand-bound structure of IGPD is available. Understandably, for better understanding of the action of IGPD and to design potent inhibitors against it, the elucidation of the molecular interactions between IGPD and its substrate and IGPD and its inhibitor is important. In this respect, we chose IGPD from *M. tuberculosis* (*Mtb*), the organism that causes tuberculosis (TB) in humans, as a model and have carried out structural and biochemical studies of the native *Mtb* IGPD and its complexes with IGP (Fig. 1*a*) and the inhibitor 3-amino-1,2,4-triazole (ATZ; Fig. 1*b*). To the best of our knowledge, this is the first report of the structures of substrate-bound and inhibited complexes of IGPD from any organism. Taken together, this study would assist the design of new compounds with enhanced affinity and increased selectivity for IGPD.

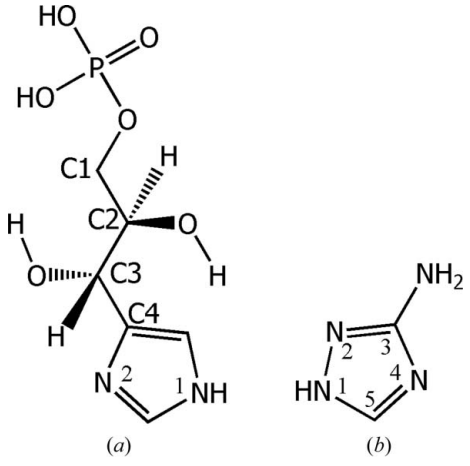

Figure 1
Schematic drawing showing the molecular structures of IGP (*a*) and ATZ (*b*).

2. Materials and methods

2.1. Enzyme preparation and crystallogenesis

Details of the cloning, overexpression in *M. smegmatis* (*Msg*), purification and preliminary X-ray studies of native IGPD have been reported elsewhere (Ahangar *et al.*, 2011). Briefly, IGPD was cloned in the *Msg–E. coli* shuttle expression vector pYUB1062 and overexpressed in *Msg* strain mc^{2}4517 in Difco Middlebrook 7H9 broth medium. The protein was purified to homogeneity by Ni–NTA affinity and gel-filtration chromatography. To determine the oligomeric state of IGPD in solution, we also performed dynamic light-scattering (DLS) experiments using a Zetasizer APS instrument (Malvern, UK). IGPD was crystallized in three different forms (Ahangar *et al.*, 2011). The IGPD–IGP complex was prepared by co-crystallization with 4 m*M* IGP in the native protein crystallization drop, whereas the IGPD–triazole inhibitor complex was prepared by soaking native crystals with 5 m*M* inhibitor solution in the reservoir solution for 12 h. X-ray diffraction data from the native and inhibitor-complex crystals were collected in-house, whereas data from the enzyme–substrate complex were collected on beamline BM14 at the European Synchrotron Radiation Facility, Grenoble, France. Data sets were processed using *HKL*-2000 (Otwinowski & Minor, 1997).

2.2. Structure solution and refinement

Owing to its high resolution, IGPD crystallized in a cubic crystal form was chosen for the final structure refinement and analysis. The structure was solved by the molecular-replacement phasing method using the program *Phaser* (McCoy *et al.*, 2007) from the *CCP*4 suite (Winn *et al.*, 2011). The structure of IGPD from *A. thaliana* (PDB entry 2f1d; Glynn *et al.*, 2005), which shares 47% sequence identity with that from *Mtb*, was used as the search model. The molecular-replacement solution yielded one molecule of IGPD in the asymmetric unit, consistent with expectation from the Matthews coefficient and solvent content. The structure was refined using the program *REFMAC*5 from the *CCP*4 suite. To start with, 100 cycles of restrained coordinate refinement were carried out using a maximum-likelihood target function. At this stage, the *Mtb* IGPD-specific amino acids were incorporated/substituted into the electron density using the model-building program *Coot* (Emsley & Cowtan, 2004). After every round of model building, positional and isotropic *B*-factor

refinements were carried out. Water molecules were incorporated into the model based on the peak heights ($2|F_o| - |F_c|$ at the 1σ contour level and $|F_o| - |F_c|$ at the 3σ contour level) in the electron-density maps. The final model, a single polypeptide chain comprising residues 10–200 of the total 210, was refined at 2.0 Å resolution to R_{work} and R_{free} values of 19.5 and 24.0%, respectively. The stereochemical acceptability of the structure was validated using *PROCHECK* (Laskowski *et al.*, 1993). The structures of the IGPD–IGP and the IGPD–ATZ complexes were determined by the molecular-replacement method using the refined *Mtb* IGPD structure as the template. The IGP and ATZ molecules were incorporated into the respective difference electron-density maps ($|F_o| - |F_c|$ contoured at the 3σ level). Subsequently, both of the complex structures were refined in a similar manner to that employed for the native structure. The data-collection, data-processing and refinement statistics are given in Table 1. The figures were prepared using *PyMOL* (DeLano, 2002).

2.3. Enzyme kinetics

The activity of the enzyme was determined using a previously described stopped-assay protocol (Martin & Goldberger, 1967) with minor modifications. The reaction mixture consisted of 40 m*M* triethanolamine (TEA) buffer pH 7.7, 50 µ*M* $MnCl_2$, 25 m*M* β-mercaptoethanol and 2 µg (1.16 µ*M*) of enzyme in a reaction volume of 75 µl. The reactions were carried out at 310 K using different IGP (Santa Cruz Biotechnology, USA) concentrations (0.035–0.861 m*M*). The reactions were stopped by adding 250 µl 1.43 *M* sodium hydroxide at different time points with an interval of 30 s. The reaction mixture was then incubated at 318 K for 20 min to convert the product imidazoleacetol-phosphate (IAP) into an enolized form, the absorbance of which at 280 nm was read in a Shimadzu UV spectrophotometer against a blank. The extinction coefficient of IAP formed under these conditions is 5310 M^{-1} cm^{-1}, as reported previously (Martin & Goldberger, 1967), and this value was used for the calculation of kinetic parameters using *GraphPad Prism* 5.0 (GraphPad Software). For inhibition studies, the activity of the enzyme in the presence of 0.4 and 0.8 m*M* ATZ with different substrate concentrations (0.036–1.6 m*M*) was determined.

3. Results

3.1. Three-dimensional structure of IGPD

The overall tertiary structure of *Mtb* IGPD, consisting of a four-helix bundle sandwiched between two four-stranded mixed β-sheets, resembles those known from other organisms such as *A. thaliana*, *F. neoformans* and *S. aureus*, as evident from the values of the root-mean-square deviations and from structural alignment (Table 2 and Supplementary Fig. S2); however, unlike the others, *Mtb* IGPD possesses a unique structural feature: the insertion of a one-turn 3_{10}-helix followed by a loop ten residues in length (Fig. 2*a*). The loop is mainly stabilized by hydrogen-bonding interactions (Fig. 2*a*). Two serine residues constituting the tip of this loop interact with water molecules. This region does not appear to possess a functional role, as none of the residues in this region are

Table 1
X-ray diffraction data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

	Native IGPD	IGPD–IGP	IGPD–triazole
Data collection			
Space group	Cubic *P*432	Cubic *P*432	Cubic *P*432
Unit-cell parameters (Å)	$a = b = c = 112.54$	$a = b = c = 112.51$	$a = b = c = 112.21$
No. of molecules per asymmetric unit	1	1	1
Matthews coefficient (Å^3 Da^{-1})	2.52	2.52	2.50
Solvent content (%)	51.1	51.1	50.7
Temperature (K)	100	100	100
Detector	R-AXIS IV^{++}	MAR225 CCD	R-AXIS IV^{++}
Wavelength (Å)	1.5418	0.97625	1.5418
Resolution (Å)	50.00–2.00 (2.07–2.00)	50.00–2.10 (2.18–2.10)	50.00–2.30 (2.38–2.30)
Unique reflections	16296 (1096)	14776 (1400)	11018 (973)
Multiplicity	9.5 (9.6)	13.4 (5.5)	6.4 (3.1)
$\langle I/\sigma(I)\rangle$	19.8 (3.2)	15.7 (2.3)	11.9 (3.0)
Completeness (%)	95.0 (65.8)	99.7 (97.5)	97.5 (89.4)
R_{merge}† (%)	11.5 (71.0)	13.4 (46.2)	15.7 (34.4)
$CC_{1/2}$‡	0.99 (0.82)	0.99 (0.26)	0.99 (0.66)
Refinement			
Unique reflections (working/test)	15436/823	14028/746	10495/523
R_{work}§ (%)	19.5	19.6	18.6
R_{free}§ (%)	24.0	24.9	22.1
No. of atoms			
Protein	1476	1476	1469
Ligand	4 [ethylene glycol]	15 [IGP]	6 [inhibitor]
Water	109	73	71
Manganese	4	3	3
Average *B* factor (Å^2)			
Protein atoms	34.3	27.7	24.4
Ligand	56.1	49.8	31.8
Water molecules	39.8	27.3	22.9
Manganese	48.8	29.2	29.1
R.m.s. deviations from ideal			
Bond lengths (Å)	0.005	0.006	0.006
Bond angles (°)	0.997	1.192	1.236
Ramachandran plot analysis, residues in (%)			
Most favoured regions	89.8	86.8	88.0
Additional allowed regions	8.4	11.4	10.2
Generously allowed regions	0.6 [Glu71]	0.6 [Glu71]	1.2 [Asp108, Glu71]
Disallowed regions	1.2 [Arg99, Asp108]	1.2 [Arg99, Asp108]	0.6 [Arg99]

† $R_{merge}(I) = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the average intensity of the *i* observations of reflection *hkl*. ‡ $CC_{1/2}$, the correlation coefficient between two half data sets (Karplus & Diederichs, 2012), was calculated using *PHENIX* (Adams *et al.*, 2010). § R_{work} and $R_{free} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where $|F_{obs}|$ and $|F_{calc}|$ are the observed and calculated structure-factor amplitudes, respectively. R_{free} was calculated using 5% of the data.

involved in substrate binding. Like the structures of IGPD from *A. thaliana, F. neoformans* and *S. aureus*, *Mtb* IGPD is pseudo-symmetric, which may have been arisen from a gene-duplication event. The two symmetry-related halves share 23% amino-acid sequence identity and 44% similarity. The overall dimensions of the molecule are approximately 50 × 40 × 25 Å. The pseudo-twofold symmetry relates strands $\beta1$, $\beta2$, $\beta3$ and $\beta4$ of the N-terminal β-sheet to the respective strands $\beta5$, $\beta6$, $\beta7$ and $\beta8$ of the C-terminal β-sheet. Similarly, helices $\alpha1$ and $\alpha2$ are symmetry mates of helices $\alpha3$ and $\alpha4$, respectively (Fig. 2*a*). *Mtb* IGPD contains only a single cysteine residue; hence, no disulfide bonds are present in the molecule.

3.2. Quaternary structure and kinetics of IGPD

The crystal asymmetric unit contains one IGPD molecule; however, analysis of the crystal packing shows that the 432 space-group symmetry generates a 24-meric hollow spherical structure with a diameter of approximately 130 Å and a molecular mass of 570 kDa (Fig. 2*b*). The results of the DLS and gel-filtration studies suggest that, as in the crystals, IGPD forms a 24-mer in solution. The value of the average hydrodynamic diameter of the molecules as measured using dynamic light scattering was about 150 ± 5.1 Å (Supplementary Fig. S3). The solvent-accessible and buried surface areas of the complex were computed to be 129 200 and 107 500 Å^2, respectively.

To establish that IGPD indeed possesses enzymatic activity as a dehydratase, we tested the ability of the purified IGPD to dehydrate its substrate IGP. The enzyme obeys Michaelis–Menten kinetics and exhibits dehydratase activity with K_m and K_{cat} values of 122 ± 15 µ*M* and (1.4 ± 0.05) × 10^3 s^{-1}, respectively (Fig. 3). An inhibition study showed that ATZ inhibits IGPD competitively with a K_i value of 310 ± 30 µ*M* (Fig. 3).

3.3. Structures of the substrate-bound and inhibited forms

To delineate the precise location of the IGP and its interactions with IGPD, the structure of an IGPD–IGP complex was determined. The strong electron-density peaks observed in the ($2|F_o| - |F_c|$) map contoured at 1σ and in the ($|F_o| - |F_c|$) map contoured at 3σ in the active-site region clearly show the binding of IGP in a cleft comprised of mainly charged residues such as histidine, lysine, arginine, aspartate and glutamate protruding from three different subunits (Figs. 4 and 5*a*). The imidazole ring of the IGP is positioned between the two active-site manganese ions (Mn1

Table 2
Pairwise root-mean-square deviations (r.m.s.d.s) (Å) among the four structures.

The number of C$^\alpha$ pairs is given in parentheses.

	A. thaliana	*F. neoformans*	*S. aureus*
M. tuberculosis	0.80 (169)	0.62 (143)	1.04 (145)
A. thaliana		0.89 (164)	0.97 (154)
F. neoformans			1.03 (152)

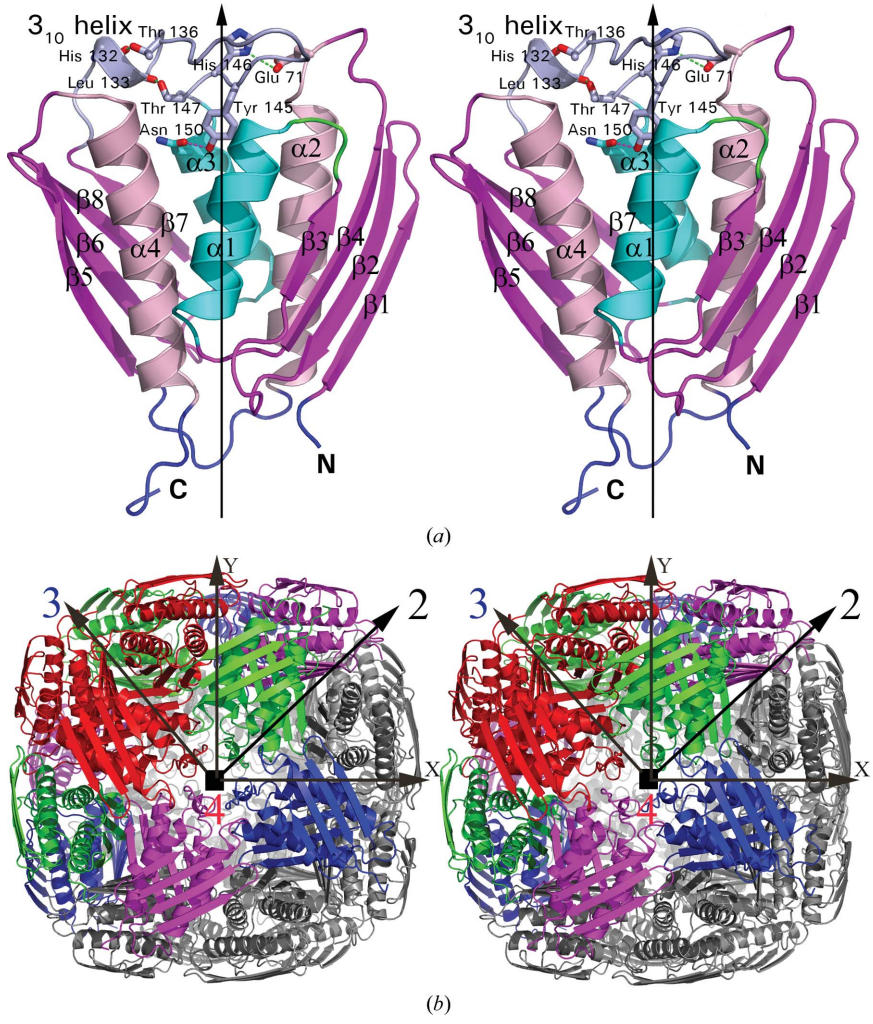

Figure 2
Tertiary and quaternary structures of *Mtb* IGPD. (*a*) Stereoview of a cartoon diagram of IGPD showing secondary-structural elements in ribbon representation. Secondary structures are labelled with Greek letters. The pseudo-twofold symmetry axis is shown by a vertical arrow. The residues protruding from the loop that interact and make hydrogen bonds with the other residues are shown in stick representation. Hydrogen-bonding interactions are shown as dotted lines. (*b*) Stereoview of the 24-meric self-assembled IGPD. The fourfold, threefold and twofold symmetry axes of the 432 point symmetry with respect to an orthogonal coordinate system are shown. Application of the fourfold symmetry about the axis with direction cosines 0.000, 0.000 and 1.000 generates the front four molecules shown in red, green, blue and magenta. The threefold symmetry about the axis with direction cosines 0.577, 0.577 and 0.577 then generates 12 molecules (all of the coloured molecules). The 24-mer assembly is generated by the application of twofold symmetry along the axis with direction cosines 0.707, 0.707 and 0.000. The 12 molecules coloured red, green, blue and magenta and the 12 molecules coloured grey are related by the twofold symmetry. The assembly contains 24 identical active sites.

and Mn2) located about 6 Å apart; each N atom of the imidazole ring interacts with a manganese ion (Fig. 5*b*). Both manganese ions are octahedrally coordinated. Mn1 is coordinated to the $N^{\varepsilon 2}$ atoms of His47*A*, His176*A* and His74*B*, the $O^{\varepsilon 1}$ atom of Glu180*A* and the N2 and O3 atoms of the substrate IGP, whereas Mn2 is coordinated to the $N^{\varepsilon 2}$ atoms of His73*B*, His152*B* and His177*A*, the $O^{\varepsilon 1}$ atom of Glu77*B*, the N1 atom of IGP and W460. The N and C atoms of the imidazole ring of the IGP make van der Waals interactions with the C^{ε} atom of Met107*A*. The $N^{\eta 1}$ and $N^{\eta 2}$ atoms of Arg99*C* are involved in hydrogen-bonding interactions with the OP5 and OP6 atoms of IGP, respectively (Fig. 5*b*). Lys184*A* is involved in substrate recognition through a hydrogen-bonding interaction between its N^{ζ} atom and OP5 of the substrate. The $O^{\varepsilon 1}$ atom of Glu21*B* interacts with O2 of IGP through a hydrogen bond. Comparison of the substrate-bound structure with that of the native enzyme by least-squares superposition shows that while the conformations of most of the active-site residues remain essentially the same,

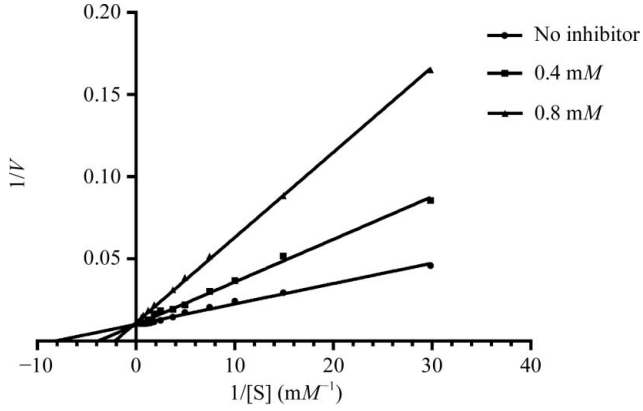

Figure 3
Enzyme kinetics. Lineweaver–Burk plot for the substrate and inhibitor kinetics.

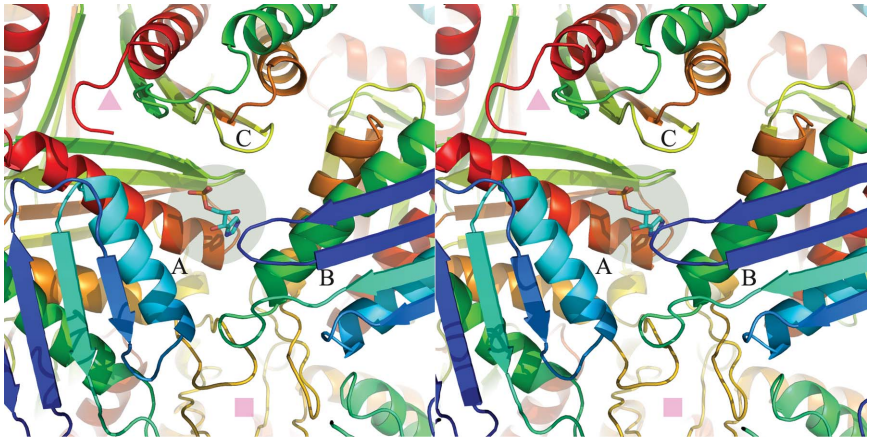

Figure 4
Mapping of the IGPD active site. Stereoview of the one of the 24 catalytic centres located at the interface of three monomers; the filled triangle and square symbolize the locations of the threefold and fourfold symmetry axes of the 432 point symmetry, respectively. Three molecules that form an active site are denoted *A*, *B* and *C*. The substrate-binding region is highlighted by a shaded circle.

the side chain of Arg121 undergoes a significant conformational change upon substrate binding and its $N^{\eta 1}$ atom is involved in a hydrogen-bonding interaction with the OP6 O atom of the phosphate (Fig. 6). Understandably, these active-site residues, with the exception of His152, are conserved among the IGPDs (which share as little as 21% overall amino-acid sequence identity) from *Mtb*, *A. thaliana*, *F. neoformans* and *S. aureus* for which crystal structures have been determined. His152 is conserved in all of these IGPDs except that from *S. aureus* (Supplementary Fig. S4).

In order to better understand the molecular basis of the inhibitory effect of triazole-derivative compounds against IGPD, the structure of IGPD bound to the inhibitor ATZ was elucidated. An examination of the electron-density maps in the active-site region revealed unambiguous density for the inhibitor (Fig. 7). The 1,2,4-triazole ring of the inhibitor binds to the enzyme in the same place as the imidazole moiety of the IGP. The N1 and N4 atoms of ATZ are coordinated to the Mn1 and Mn2 ions, respectively. The ring N and C atoms of the inhibitor make van der Waals interactions with the C^{ε} atom of Met107 (Fig. 7). Notably, the side chain of Arg121 of the inhibitor-bound structure undergoes a conformational change similar to that occurring upon IGP binding. Since the triazole ring of the ATZ occupies the same position in the active site as the imidazole moiety of the substrate, it is reasonable to hypothesize that inhibitor-bound IGPD may preclude substrate binding.

4. Discussion

One third of the world's population are infected with TB. From this pool, about ten million new cases of active TB arise and between two and three million deaths occur every year globally despite the worldwide use of the Bacille Calmette–Guerin (BCG) vaccine and several antibiotics. The problem is further compounded by the emergence of extensively drug-resistant (XDR) and multi-drug resistant (MDR) strains of TB. Unfortunately, no new antituberculosis compound has been introduced in the past 45 years, although a few are in the pipeline (Laughon, 2007). Therefore, it is of paramount importance to carry out studies of the molecules pertaining to *Mtb* infections, latency and growth through a multidisciplinary approach. Importantly, it has been demonstrated that *Mtb* tryptophan and histidine auxotrophs failed to survive single-amino-acid starvation (Parish, 2003). In another independent study, it has also been shown that the histidine biosynthesis pathway is essential for the optimal growth of *Mtb* (Sassetti *et al.*, 2003). Moreover, the Tropical Disease Research (TDR) Targets database has ranked *Mtb* IGPD (Rv1601) among the top 50 targets for anti-tuberculosis agent development (Agüero *et al.*, 2008). The enzymes of the histidine

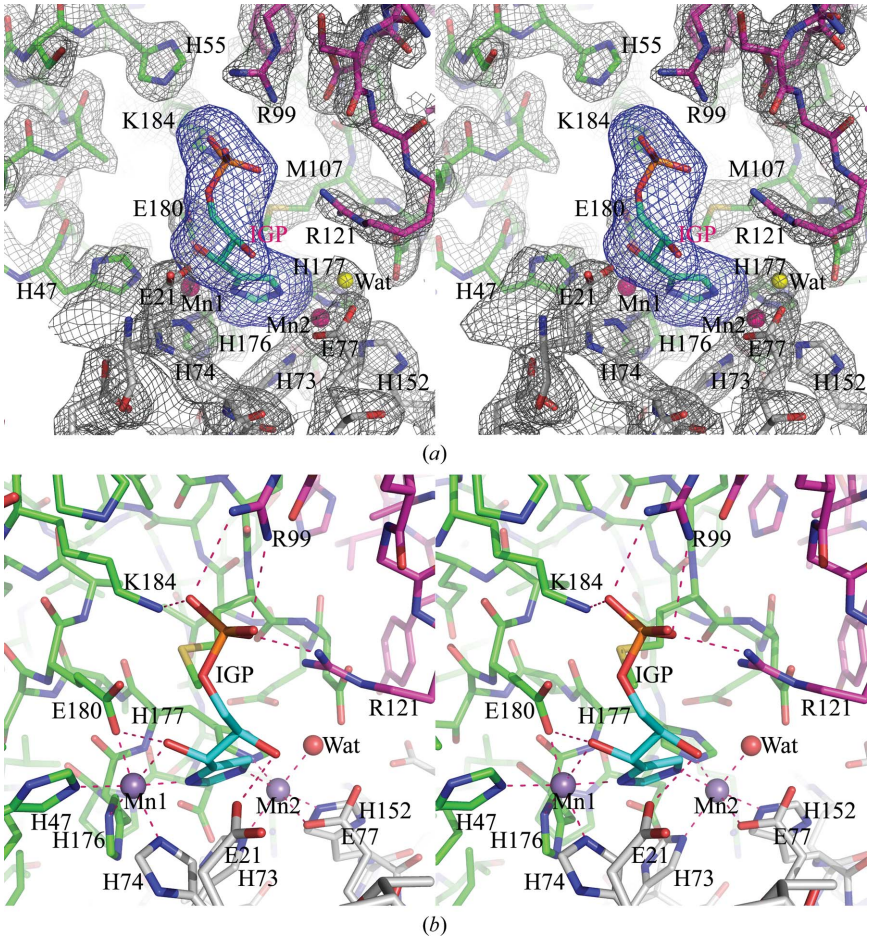

Figure 5
(*a*) The $|F_o| - |F_c|$ OMIT electron-density map contoured at the 3σ level for the substrate IGP and the $2|F_o| - |F_c|$ map contoured at the 1σ level for the residues lining the active site. (*b*) The interactions between the enzyme and the substrate IGP. The refined IGP was superimposed onto the electron-density map. The C atoms of residues of molecules *A*, *B* and *C* are coloured green, grey and pink, respectively. The dotted lines show coordination bonds or hydrogen bonds.

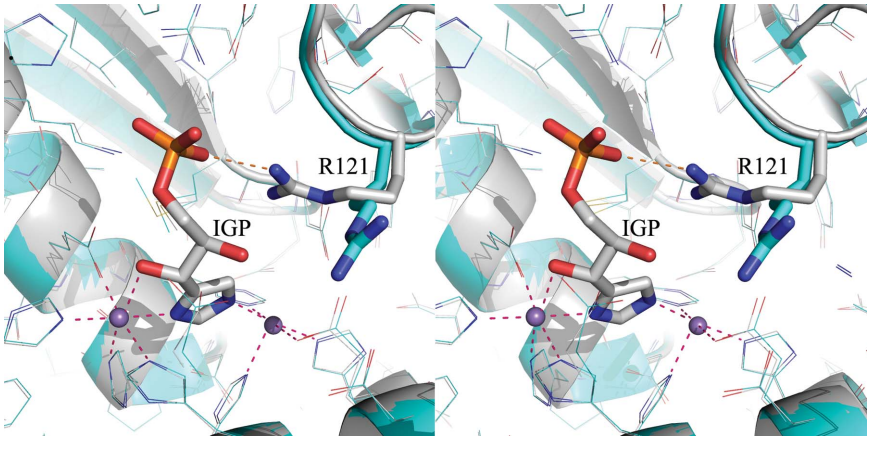

Figure 6
Substrate-induced conformational change. Stereoview of the superimposed structures of the native and substrate-bound forms. The arginine (121) residues of the native and substrate-bound structures are shown in a stick representation with C atoms in cyan and grey, respectively. A hydrogen bond between a phosphate O atom of IGP and the $N^{\eta 1}$ atom of Arg121 is shown as an orange dotted line.

biosynthesis pathway are therefore among the attractive targets for the development of new anti-TB agents. In this context, the present study involving the structural and biochemical characterization of *Mtb* IGPD is relevant in many aspects.

Firstly, both structural and biochemical data show that *Mtb* IGPD is a 24-meric monofunctional enzyme. The 24-meric IGPD possesses 432 molecular symmetry, with symmetry axes coinciding with the crystallographic symmetry axes of space group *P*432. The counterparts of *Mtb* IGPD in *A. thaliana*, *F. neoformans* and *S. aureus* also exhibit the same oligomeric assembly in the crystal, suggesting that the biological functional unit of IGPD is perhaps a 24-mer in prokaryotes and plants. However, in contrast to IGPD from other organisms, the monomeric *Mtb* IGPD possesses a unique structural feature: the insertion of a 3_{10}-helix followed by a loop of ten residues in length. In some prokaryotes such as *E. coli*, *S. typhimurium* and *A. brasilense* the sixth step is catalysed by the dehydratase domain of a bifunctional enzyme (encoded by the *hisB* gene) which has both dehydratase and phosphatase activities. The phosphatase activity of this bifunctional enzyme is required to catalyse the eighth step in the pathway. However, the enzyme that is involved in the catalysis of the eighth step in *Mtb* remains unclear. However, the literature suggests that it may be Rv0114 (similar to HisB) which catalyses this step (Cole *et al.*, 1998).

Secondly, the substrate-bound crystal structure revealed that the active site of the enzyme is made up of residues protruding from three subunits. Moreover, the complex structure delineates the residues involved in substrate binding. Analyses of the enzyme–substrate interactions suggest that His47, His73, His74, His152, His176 and His177 are involved in metal binding, that Glu21 and Glu180 may be responsible for catalysis and that Arg99, Arg121 and Lys184 are involved in the stabilization of the phosphate group of the substrate. However, further experimental studies are needed to address this. Since the crystals of the enzyme–substrate complex were obtained by cocrystallization, one would anticipate the presence of the product IAP instead of the observed substrate IGP in the crystal structure. It may be that the solution environment (particularly the presence of PEG 1500) in the

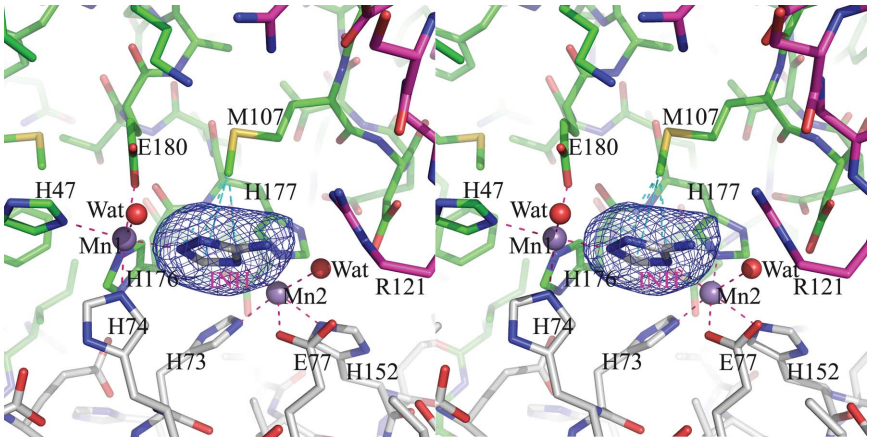

Figure 7
The $|F_o| - |F_c|$ OMIT electron-density map (shown in blue) contoured at the 3σ level for the inhibitor ATZ. Also shown in the figure are the interactions between the enzyme and the ATZ. The refined ATZ was superimposed onto the electron-density map. The dotted lines show the coordination bonds or van der Waals interactions. The C atoms of residues of molecules *A*, *B* and *C* are coloured green, grey and pink, respectively.

crystallization drop was unfavourable for the enzymatic reaction to occur.

Thirdly, in conjunction with the kinetic data, the structure of the *Mtb* HisB–ATZ complex demonstrates that ATZ inhibits HisB competitively. In light of the fact that IGPD is absent in humans and that histidine is required for the survival of *Mtb*, the structures of the IGPD–IGP and IGPD–ATZ complexes provide a platform which could help in the design of potent inhibitors through a structure-guided approach.

Fourthly, it has been established that IGPD is a major target for herbicide discovery. Therefore, in the context of developing efficient new herbicides to control weeds, these structural studies will undoubtedly be helpful. It is worth mentioning here that among all of the enzymes of the histidine biosynthesis pathway, only HisB functions as a 24-mer with protomers arranged with 432 point-group symmetry. This might allow the formation of active sites constituted by three protomers.

In conclusion, we have reported the first crystal structures of substrate-bound and inhibited forms of IGPD, which provide valuable insights pertaining to IGPD structure, function and inhibition.

The authors thank Professor A. Surolia and Dr C. Shaha, the former and current directors of NII, respectively, for their encouragement and support. X-ray data were collected using the in-house X-ray facility which was established with financial support from the Department of Biotechnology (DBT), Government of India and on synchrotron beamline BM14 at the ESRF, Grenoble, France under a program supported by the DBT, Government of India. We thank Ravikant Pal for his help during data collection. BKB received funding from the Indian Council of Medical Research (Reference No. 5/8/5/4/2010-ECD-I) and the National Institute of Immunology, New Delhi, India.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Agüero, F. *et al.* (2008). *Nature Rev. Drug Discov.* **7**, 900–907.
Ahangar, M. S., Khandokar, Y., Nasir, N., Vyas, R. & Biswal, B. K. (2011). *Acta Cryst.* F**67**, 1451–1456.
Alifano, P., Fani, R., Liò, P., Lazcano, A., Bazzicalupo, M., Carlomagno, M. S. & Bruni, C. B. (1996). *Microbiol. Rev.* **60**, 44–69.
Ames, B. N., Garry, B. & Herzenberg, L. A. (1960). *J. Gen. Microbiol.* **22**, 369–378.
Bazzicalupo, M., Fani, R., Gallori, E., Turbanti, L. & Polsinelli, M. (1987). *Mol. Gen. Genet.* **206**, 76–80.
Brilli, M. & Fani, R. (2004). *Gene*, **339**, 149–160.
Chiariotti, L., Nappo, A. G., Carlomagno, M. S. & Bruni, C. B. (1986). *Mol. Gen. Genet.* **202**, 42–47.
Cole, S. T. *et al.* (1998). *Nature (London)*, **393**, 537–544.
DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Fujimori, K. & Ohta, D. (1998). *FEBS Lett.* **428**, 229–234.
Glynn, S. E., Baker, P. J., Sedelnikova, S. E., Davies, C. L., Eadsforth, T. C., Levy, C. W., Rodgers, H. F., Blackburn, G. M., Hawkes, T. R., Viner, R. & Rice, D. W. (2005). *Structure*, **13**, 1809–1817.
Gohda, K., Kimura, Y., Mori, I., Ohta, D. & Kikuchi, T. (1998). *Biochim. Biophys. Acta*, **1385**, 107–114.
Karplus, P. A. & Diederichs, K. (2012). *Science*, **336**, 1030–1033.
Klem, T. J. & Davisson, V. J. (1993). *Biochemistry*, **32**, 5177–5186.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Laughon, B. E. (2007). *Curr. Top. Med. Chem.* **7**, 463–473.
Mano, J., Hatano, M., Koizumi, S., Tada, S., Hashimoto, M. & Scheidegger, A. (1993). *Plant Physiol.* **103**, 733–739.
Martin, R. G. & Goldberger, R. F. (1967). *J. Biol. Chem.* **242**, 1168–1174.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Mori, I., Fonne-Pfister, R., Matsunaga, S., Tada, S., Kimura, Y., Iwasaki, G., Mano, J., Hatano, M., Nakano, T., Koizumi, S., Scheidegger, A., Hayakawa, K. & Ohta, D. (1995). *Plant Physiol.* **107**, 719–723.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Parish, T. (2003). *J. Bacteriol.* **185**, 6702–6706.
Parker, A. R., Moore, T. D., Edman, J. C., Schwab, J. M. & Davisson, V. J. (1994). *Gene*, **145**, 135–138.
Sassetti, C. M., Boyd, D. H. & Rubin, E. J. (2003). *Mol. Microbiol.* **48**, 77–84.
Schweitzer, B. A., Loida, P. J., CaJacob, C. A., Chott, R. C., Collantes, E. M., Hegde, S. G., Mosier, P. D. & Profeta, S. (2002). *Bioorg. Med. Chem. Lett.* **12**, 1743–1746.
Schweitzer, B. A., Loida, P. J., Thompson-Mize, R. L., CaJacob, C. A. & Hegde, S. G. (1999). *Bioorg. Med. Chem. Lett.* **9**, 2053–2058.
Sinha, S. C., Chaudhuri, B. N., Burgner, J. W., Yakovleva, G., Davisson, V. J. & Smith, J. L. (2004). *J. Biol. Chem.* **279**, 15491–15498.
Staples, M. A. & Houston, L. L. (1979). *J. Biol. Chem.* **254**, 1395–1401.
Stepansky, A. & Leustek, T. (2006). *Amino Acids*, **30**, 127–142.
Tada, S., Hatano, M., Nakayama, Y., Volrath, S., Guyer, D., Ward, E. & Ohta, D. (1995). *Plant Physiol.* **109**, 153–159.
Tada, S., Volrath, S., Guyer, D., Scheidegger, A., Ryals, J., Ohta, D. & Ward, E. (1994). *Plant Physiol.* **105**, 579–583.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.