Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structures of mannose-6-phosphate isomerase from *Salmonella typhimurium* bound to metal atoms and substrate: implications for catalytic mechanism

S. R. Sagurthi,[a] Giri Gowda,[a] H. S. Savithri[b] and M. R. N. Murthy[a]*

[a]Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India, and [b]Department of Biochemistry, Indian Institute of Science, Bangalore 560 012, India

Correspondence e-mail: mrn@mbu.iisc.ernet.in

Received 11 February 2009
Accepted 8 April 2009

PDB References: *St*-MPI–EDO, 2wfp; *St*-MPI–Zn^{2+}, 3hlm; *St*-MPI–Zn^{2+}–Y^{3+}, 3hlw; *St*-MPI–F6P–Zn^{2+}, 3hly.

Mannose-6-phosphate isomerase (MPI) catalyzes the interconversion of mannose 6-phosphate and fructose 6-phosphate. X-ray crystal structures of MPI from *Salmonella typhimurium* in the apo form (with no metal bound) and in the holo form (with bound Zn^{2+}) and two other structures with yttrium bound at an inhibitory site and complexed with Zn^{2+} and fructose 6-phosphate (F6P) were determined in order to gain insights into the structure and the isomerization mechanism. Isomerization involves acid/base catalysis with proton transfer between the C1 and C2 atoms of the substrate. His99, Lys132, His131 and Asp270 are close to the substrate and are likely to be the residues involved in proton transfer. The interactions observed at the active site suggest that the ring-opening step is probably catalyzed by His99 and Asp270. An active-site loop consisting of residues 130–133 undergoes conformational changes upon substrate binding. Zn^{2+} binding induces structural order in the loop consisting of residues 50–54. The metal atom appears to play a role in substrate binding and is probably also important for maintaining the architecture of the active site. Isomerization probably follows the previously suggested *cis*-enediol mechanism.

1. Introduction

Mannose-6-phosphate isomerase (MPI) catalyzes the interconversion of mannose 6-phosphate (M6P) and fructose 6-phosphate (F6P), which are substrates for glycolysis and gluconeogenesis. MPI is essential for bacterial growth on mannose as the sole source of carbon. The absence of MPI activity in yeast causes cell lysis (Smith *et al.*, 1995). MPI has been found to be essential for the production of the exopolysaccharide alginate, which coats the bacteria and protects them from antibiotics and the host's innate immunity (Shinabarger *et al.*, 1991). Deficiency of MPI activity leads to a human disease known as congenital disorder of glycosylation. The disorder results from the hypoglycosylation of serum and other glycoproteins (Jaeken *et al.*, 1998; De Koning *et al.*, 1998). Ablation of MPI causes an accumulation of mannose 6-phosphate in mouse, leading to toxicity and embryonic lethality (DeRossi *et al.*, 2006). MPI is also important for the virulence of the protozoan parasite *Leishmania mexicana* (Awadalla *et al.*, 1987). The amino-acid sequence of human MPI does not share significant identity with those of pathogenic organisms. Therefore, MPI has been considered to be an important drug target.

MPIs have been classified broadly into two groups, type I and type II, based on domain organization (Jensen & Reeves,

1998). Type I enzymes, which have been shown to be zinc-dependent, have one metal atom per monomer and are monofunctional, catalyzing a single isomerization reaction. Type II MPIs are bifunctional enzymes that possess both MPI and guanosine diphospho-D-mannose pyrophosphorylase activities in separate catalytic domains. Although they have no overall sequence similarity, type I and type II MPIs share a conserved catalytic motif. Recent inhibition studies on type I and type II MPIs from baker's yeast and *Pseudomonas aeruginosa*, respectively, showed that the intermediate analogue D-arabinohydroxamic acid (5PAH) is the most potent inhibitor of MPI (Roux *et al.*, 2004). Mass-spectrometric studies on *Escherichia coli* MPI showed that binding of metal ions increases the binding affinity of the substrate by a factor of five (Gao *et al.*, 2005). Biochemical studies on MPI from *Xanthomonas campestris* showed that it had optimum activity at pH 7.0 and its activation by metals was in the order $Co^{2+} > Zn^{2+} > Mn^{2+} > Ni^{2+} > Ca^{2+}$ (Papoutsopoulou & Kyriakidis, 1997). In addition to MPI, most organisms also possess a distinct enzyme that catalyzes the interconversion of glucose 6-phosphate (G6P), the C2 epimer of M6P, and F6P. A member of the phosphoglucose isomerase (PGI) superfamily from the crenarchaeon *Pyrobaculum aerophilum* has been found to possess an enzyme with a novel PGI/MPI dual function (Swan *et al.*, 2004). The reaction catalyzed by MPI involves the stereospecific transfer of a proton between the C1 and C2 positions of the substrate (Seeholzer, 1993). Two alternative mechanisms, the hydride-shift mechanism and the *cis*-enediol mechanism, have been suggested for the reaction catalyzed by MPI. The finely tuned specificity of these enzymes could be understood if the three-dimensional structures of several of these isomerases were available for comparative analysis.

The crystal structure of MPI from *Candida albicans* (*Ca*-MPI; PDB code 1pmi), a zinc-dependent enzyme, has been determined at 1.7 Å resolution (Cleasby *et al.*, 1996). The enzyme has three distinct domains: two are β-domains while the third is an almost exclusively α-helical domain. The active site lies in the central β-domain. *Ca*-MPI has been shown to be inhibited by sugar phosphates such as erythrose 4-phosphate at millimolar concentrations (Proudfoot *et al.*, 1994). In addition to the zinc at the catalytic site, it has been shown that MPI binds Zn^{2+} at two other sites with lower affinity. Binding at these sites leads to inhibition of the enzyme (Wells *et al.*, 1993).

The structures of type 1 MPI from *Bacillus subtilis* (*Bs*-MPI; PDB code 1qwr; Y. Kim, L. Lezondra & A. Joachimiak, unpublished work) and *Archeoglobus fulgidus* (*Af*-MPI; PDB code 1zx5; M. E. Cuff, T. Skarina, A. Edwards, A. Savchenko & A. Joachimiak, unpublished work) have been determined and deposited in the PDB as part of structural genomics initiatives. However, these structures have not been described in the literature. Also, structures of substrate or inhibitor complexes of *Ca*-MPI were not determined. There is thus a paucity of structures of MPI and its complexes with metal ions and inhibitors for comparative analysis and for understanding the functional mechanism of MPI.

In the present communication, we describe the X-ray structures of MPI from *Salmonella typhimurium* (*St*-MPI) with different bound metal ions (Zn^{2+} and Y^{3+}), with ethylene glycol in the place of metal ion and of the complex with F6P and Zn^{2+}. The zinc-bound structure represents the active form of the enzyme, while the yttrium-bound structure represents the enzyme form with occupied inhibitory site. A comparative analysis of the structures of MPIs available in the PDB is presented. Unlike *Ca*-MPI, *St*-MPI is activated more by Mg^{2+} than by Zn^{2+}. The metal atom appears to be important for maintenance of the active-site geometry and plays a role in substrate binding.

2. Materials and methods

2.1. Cloning and purification of *St*-MPI

The cloning, expression, purification, crystallization and preliminary X-ray crystallographic analysis of *St*-MPI have been reported previously (Gowda *et al.*, 2008). Briefly, *St*-MPI was cloned in pRSET-C vector and overexpressed in BL21 (DE3) pLysS with an N-terminal hexahistidine tag. *St*-MPI was purified to homogeneity using nickel–nitrilotriacetic acid affinity column chromatography. The purity of the enzyme was checked by mass spectrometry and on SDS–PAGE.

2.2. Crystallization, structure solution and refinement

Crystals suitable for X-ray diffraction studies were obtained at 293 K using the microbatch method. Crystallization drops were prepared by mixing 3 µl protein solution with 3 µl of a crystallization cocktail containing protein at a concentration of 4 mg ml^{-1}, 0.1 *M* magnesium acetate, 0.2 *M* sodium cacodylate pH 6.5, 20% PEG 8000 and 5% dioxane. The crystal parameters and data-collection statistics are summarized in Table 1. Four independent data sets extending to 2.5–1.7 Å resolution were collected after soaking the crystals in cryoprotectant [crystallization condition containing 20%(*v*/*v*) ethylene glycol] for between 30 s and 18 h. It was found that the crystals diffracted to a much higher resolution (1.7 Å) when soaked in the cryoprotectant for 18 h. This data set is referred to as *St*-MPI–EDO. The apo and holo enzymes were also cocrystallized with 10 m*M* magnesium, manganese, cobalt, nickel, yttrium and other transition-metal ions. To obtain protein complexed with Zn^{2+} and F6P (*St*-MPI–F6P–Zn^{2+}), *St*-MPI crystals were transferred to wells containing 10 m*M* Zn^{2+} and progressively increasing concentrations of F6P (10, 20, 40, 100 and 250 m*M*). Zn^{2+} and F6P were dissolved in distilled water to the required concentrations. Soaking was for 1 h at each concentration.

X-ray diffraction data sets for all crystals were collected using an image-plate detector mounted on a Rigaku UltraX rotating-anode X-ray generator. The data were processed and scaled using the programs *MOSFLM* and *SCALA* from the *CCP*4 package (Collaborative Computational Project, Number 4, 1994). All crystals belonged to space group $P2_12_12_1$ with identical unit-cell parameters. The volume of the asymmetric unit in these crystals was compatible with the presence

Table 1
Data-collection statistics.

Values in parentheses are for the last resolution shell.

	St-MPI–EDO†	*St*-MPI–F6P–Zn^{2+}	*St*-MPI–Zn^{2+}	*St*-MPI–Zn^{2+}–Y^{3+}
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters				
a (Å)	35.90	35.89	36.03	36.01
b (Å)	92.47	92.25	92.4	92.08
c (Å)	111.87	111.53	111.66	111.5
Resolution range (Å)	30–1.66 (1.7–1.66)	30–2.0 (2.1–2.0)	30–2.5 (2.6–2.5)	30–1.9 (2.0–1.9)
Total no. of reflections	863957	155281	47822	126750
No. of unique reflections	44547	24243	12581	27871
Overall completeness (%)	98.9 (89.6)	98.0 (87.3)	93.6 (90.0)	89.8 (81.6)
R_{merge}‡ (%)	6.2 (55.6)	8.9 (34.6)	13.5 (43.3)	6.1 (29.5)
$R_{p.i.m.}$§ (%)	1.5 (12.1)	3.2 (14.7)	7.2 (23.8)	3.7 (15.7)
$\langle I/\sigma(I)\rangle$	48.5 (5.2)	14.9 (4.7)	7.6 (2.6)	18.9 (4.2)
Multiplicity	19.4 (16)	6.4 (6.2)	3.8 (3.8)	4.5 (4.2)

† The cloning and data-collection statistics of *St*-MPI–EDO have been reported previously (Gowda *et al.*, 2008). ‡ $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the *i*th observation of reflection *hkl* and $\langle I(hkl)\rangle$ is its mean intensity. § $R_{p.i.m.} = \sum_{hkl}[1/(N-1)]^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the *i*th observation of $I(hkl)$, $\langle I(hkl)\rangle$ is its mean intensity and *N* is the number of times a given reflection is measured.

Table 2
Refinement statistics.

	St-MPI–EDO†	*St*-MPI–F6P–Zn^{2+}	*St*-MPI–Zn^{2+}	*St*-MPI–Zn^{2+}–Y^{3+}
No. of atoms in structure	3473	3254	3098	3290
No. of water molecules	371	218	103	239
No. of ethylene glycols	16	6	3	6
No. of metal ions	—	1	1	2
No. of ligands (F6P)	—	1	—	—
Model quality				
Final *R* factor (%)	18.6	21.4	21.9	18.6
Final R_{free} (%)	21.4	24.5	26.8	24.2
R.m.s.d. from ideal values				
Bond lengths (Å)	0.01	0.003	0.004	0.01
Bond angles (°)	1.47	0.83	0.97	1.33
Chirality	0.09	0.05	0.06	0.08
Average *B* factor (Å^2)				
Protein atoms	16.01	18.7	19.7	21.4
Waters	33.7	30.5	18.9	26.50
Ethylene glycols	30.1	22.9	24.3	20.7
Metal ions	—	19.7†	27.1†	17 (Zn)†, 17.4 (Y)†
Ligand (F6P)	—	20.0†	—	—
Ramachandran statistics (%)				
Most favoured region	91.1	91.4	90.8	90.8
Generously allowed region	8.3	8.0	8.6	8.6
Additionally allowed region	—	—	—	—
Disallowed region	0.6	0.6	0.6	0.6

† F6P, Zn^{2+} and Y^{3+} were refined with 0.5, 0.6 and 0.5 occupancy, respectively.

of one monomer. Gel-filtration and DLS studies on *St*-MPI also suggested that it was a monomer (radius of hydration 31.4 Å; data not shown) in solution.

The Matthews coefficient (V_M) and solvent content were 2.08 Å^3 Da^{-1} and 41.2%(*v*/*v*), respectively.

St-MPI has a sequence identity of 29.7% with MPI from *C. albicans* (PDB code 1pmi). Molecular-replacement trials carried out with *Phaser* (McCoy *et al.*, 2005) using 1pmi as the phasing model failed to give a satisfactory solution, probably owing to the low sequence identity and the longer length of the *Ca*-MPI polypeptide. The phasing model was edited so as to keep conserved residues and mutate other residues to alanine. Further attempts were made with the modified model, allowing a small number of short contacts. A reasonable solution was obtained with a log-likelihood gain (LLG) of 78.2 and a *Z* score of 7.7. Refinement was initiated with the program *REFMAC*5 (Murshudov *et al.*, 1997) followed by manual model building in *Coot*. The progress of refinement was monitored by using 5% of independent reflection measurements for calculation of the free *R* factor. The $2F_o - F_c$ map calculated after initial positional refinement showed a good fit for the majority of the main chain in the two β-sheet domains. These domains were built in the first round of manual model building. The electron density corresponding to the helical segments was less well defined, probably owing to differences in the organization of the helical domain from that of the model. The model corresponding to the β-sheet domains was used as input to *ARP*/*wARP* v.6.1 for automated model building and 85% of the residues were built by the program. The rest of the residues were built manually in subsequent cycles of refinement. In the final stages of refinement, inspection of the difference density revealed a few well defined regions of electron density that were not accounted for by the polypeptide. These were interpreted as bound ethylene glycol (EDO) molecules. One of the EDO-binding sites overlapped with the metal-binding site of other MPIs.

The other structures were determined using difference Fourier maps and refinement. The structures obtained by brief soaking (30 s) in the cryoprotectant did not contain bound EDO. Instead, a probable metal ion was found at the same position. This was followed by identification of bound waters using the automated water-picking algorithm of *Coot* (Emsley & Cowtan, 2004). The positions of these waters were manually checked. Some further water molecules were manually identified on the basis of electron density contoured at 1.0σ in the $2F_o - F_c$ map and 3.0σ in the $F_o - F_c$ map.

2.3. Quality of the structures

The refinement statistics are summarized in Table 2. The final refined structure consists of the polypeptide chain extending from residues −2 to 391 in *St*-MPI–EDO (three residues originating from the vector are also ordered) and from −1 to 391 in the other structures. The final models were examined with *PROCHECK* (Laskowski *et al.*, 1993). All

residues were within the allowed or generously allowed regions of the Ramachandran plot except for Ala147 and Asp355 in all the structures determined. Although the Ramachandran angles corresponding to Ala147 ($\varphi = -73.6^\circ$, $\psi = -162.0^\circ$) of *St*-MPI are retained in *Ca*-MPI ($\varphi = -79.0^\circ$, $\psi = -162.0^\circ$), they do not fall in the disallowed region in the latter as Ala has been replaced by Gly. The strained conformation of Ala147 in *St*-MPI-EDO is stabilized by strong hydrogen bonding between Ala147 N and Pro254 O, between Asn146 O and Asn237 N and between Ala147 O and Asn237 N. Asp355 is present on a loop that is exposed to the solvent and has a relatively high *B* factor (25 Å^2 at C$^\alpha$ compared with the overall value of 16 Å^2 for all protein atoms); its φ, ψ angles fall on the border of the generously allowed region. Electron density was not well defined for the loops consisting of residues 50–54 in *St*-MPI–EDO, 50–52 in *St*-MPI–F6P–Zn^{2+} and residue 52 in *St*-MPI–Zn^{2+}. These residues were not modelled in their respective structures. In contrast, clear density was observed for these residues in the *St*-MPI–Mg^{2+} structure (not shown).

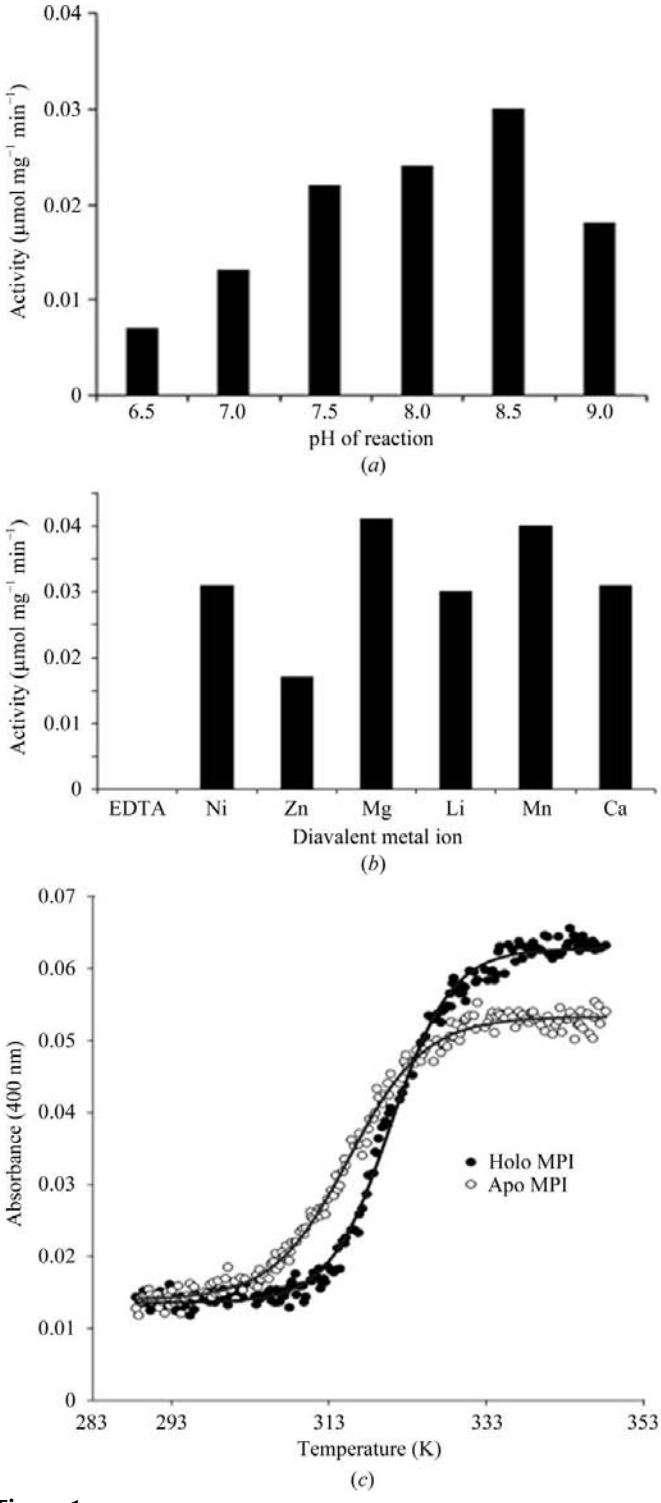

Figure 1
Biochemical characterization of *St*-MPI. (*a*) Effect of pH on the catalytic activity of *St*-MPI when M6P is used as the substrate. (*b*) Effect of divalent cations on the catalytic activity of *St*-MPI. (*c*) Thermal denaturation profile of *St*-MPI. The T_m values for the zinc-bound and apo forms of the enzyme are 321 and 316 K, respectively.

2.4. Structure analysis

The crystal structures were visually examined using *Coot* (Emsley & Cowtan, 2004) and *PyMOL* (DeLano, 2002). Superpositions of the MPI structures were carried out using *ALIGN* (Cohen, 1997) and *MUSTANG* (Konagurthu *et al.*, 2006). Structure-based sequence alignment of the amino-acid sequences of *St*-MPI, *Ca*-MPI, *Af*-MPI and *Bs*-MPI was achieved using *MUSTANG* and represented using *ESPript* (Gouet *et al.*, 2003). All interactions were identified using the *CONTACT* program from the *CCP*4 package (Collaborative Computational Project, Number 4, 1994), considering distance cutoffs of 3.5 Å for hydrogen bonds and 4.0 Å for van der Waals interactions. Average *B* factors for protein atoms and water molecules were calculated using the *BAVERAGE* program from the *CCP*4 package (Collaborative Computational Project, Number 4, 1994). The program *GRASP* was used for analysis of the packing of residues and identification of cavities (Nicholls *et al.*, 1991).

2.5. Activity assay and kinetic studies

MPI was assayed spectrophotometrically following published procedures (Roux *et al.*, 2007). The standard assay mixture (1 ml) contained 50 m*M* HEPES pH 8.5, 1 m*M* mannose 6-phosphate and 3 m*M* magnesium chloride. The assay mixture was incubated at 310 K for 5 min. The reaction was initiated by the addition of 40 µg *St*-MPI. After 30 min, the reaction was stopped by the addition of 2 ml 0.01 *M* thiobarbituric acid (TBA) in concentrated HCl. After heating to 353 K for 6 min, the reaction mixture was cooled to room temperature in an ice-cold water bath. The amount of TBA–F6P complex (hydroxymethylfurfural) formed was measured spectrometrically using a Jasco UV–visible spectrophotometer model V-530 (Japan Spectroscopic Co.) at 434 nm. Experiments were repeated at least three times from three independent purifications. Kinetic parameters were calculated by fitting the initial velocity *versus* substrate concentration to the Michaelis–Menten equation, $V = (V_{max}[S])/(K_m + [S])$, using the nonlinear regression analysis option of the *GraphPad Prism* software.

2.6. Thermal denaturation experiments

Thermal denaturation experiments were carried out using a Jasco UV–visible spectrophotometer model V-530 (Japan Spectroscopic Co.) in a quartz cell of 10 mm path length. The temperature was varied from 288 to 353 K at a rate of 1 K min^{-1}. The absorbance was measured at 400 nm.

3. Results and discussion

3.1. Biochemical studies on *St*-MPI

The enzymatic activity was measured by monitoring the increase in absorbance at 435 nm arising from the product F6P formed from the isomerization of M6P by purified *St*-MPI. pH-profile experiments showed that *St*-MPI has significant activity over a wide pH range (6.5–9.0), with optimum activity at pH 8.5 (Fig. 1*a*). The optimum temperature for catalytic activity was found to be 310 K. The kinetic constants K_m and V_{max} of *St*-MPI were 1.34 m*M* and 833.3 µmol min^{-1} mg^{-1}, respectively. These constants are comparable to those of fungal and mammalian MPIs. The K_m values of *C. albicans*, *Saccharomyces cerevisiae* and human MPI are 1.24, 0.65 and 0.25 m*M*, respectively, and their V_{max} values are 1200, 980 and 110 µmol min^{-1} mg^{-1}, respectively (Proudfoot *et al.*, 1994). The requirement of divalent cations for enzyme activity was examined by estimating the activity in the presence of 1 m*M* ethylenediaminetetraacetic acid (EDTA); it was found that the enzyme was inactive. It was reactivated by the addition of divalent metal cations such as Zn^{2+}, Mn^{2+}, Co^{2+}, Mg^{2+} and Ni^{2+} (Fig. 1*b*). The activation of the enzyme was greater in the presence of Mg^{2+} than of Zn^{2+}. In contrast, yeast MPI showed maximum activity with Zn^{2+} (Gracy & Noltmann, 1968). Thermal denaturation studies have shown that *St*-MPI is stabilized by metal binding (T_m of 321 K compared with 316 K without bound zinc; Fig. 1*c*).

3.2. Structure of *St*-MPI

The ellipsoidally shaped *St*-MPI has approximate dimensions of 47.1 × 67.9 × 59.3 Å. As determined by the *DSSP* program (Kabsch & Sander, 1983), the polypeptide contains 21 strands and 14 helices constituting 37.8 and 28.7% of the total residues, respectively. Although *St*-MPI has low sequence identity (29%) to *Ca*-MPI, their overall folds are similar. Like *Ca*-MPI, it has three distinct spatial domains. Two of the domains have a core of antiparallel β-strands resembling the cupin fold (Dunwell, 1998). The third domain consists of only helical segments. A cartoon representation of the *St*-MPI polypeptide is shown in Fig. 2. The central β-domain is flanked by the helical domain on one side and the carboxy domain on the other. The N- and C-termini of the polypeptide are in the outer β-sheet domain. Most of the residues in this domain (1–6 and 305–391) are from the carboxyl end of the molecule. Residues 1–5 form a strand (β1). The carboxy-terminal residues fold into eight antiparallel strands (β14, β15, β16, β17, β18, β19, β20 and β21). The loops connecting the strands are relatively short, whereas the loops in the central catalytic domain are longer. Only *St*-MPI possesses a short 3_{10}-helix (residues 372–375). The overall folding pattern of the central catalytic domain shows similarities to that of the C-terminal domain. This is the major domain and consists of 210 residues (7–149 and 238–304). It has a β-sheet core consisting of ten strands (β2, β3, β6, β7, β8, β9, β10, β11, β12 and β13) and two additional strands β4 and β5. The core is flanked by seven helices (α1, α2, α3, α4, α5, α12 and α13). There is a large cleft in the catalytic domain between β7 and β8. This cleft contains the metal-ion and substrate-binding sites. The interface between the C-terminal domain and the catalytic domain is hydrophobic. No solvent molecules were located between the adjacent β-sheets that form the interface between the C-terminal and catalytic domains. The helical domain consisting of residues 150–237 is an insertion between β9 and β10 and contains six helical segments (α6, α7, α8, α9, α10 and α11) connected by short turns and loops. The corresponding helical domains in *Bs*-MPI and *Af*-MPI are very small and consist of only 23 residues (Fig. 3*a*). The r.m.s.d. for the superposition of the C^{α} atoms of *St*-MPI and *Ca*-MPI is 1.58 Å. In contrast, the r.m.s.d. for the superposition of the helical domains of the two structures is 2.16 Å. As the catalytic domain is the most conserved of the three domains and is placed in the middle of the structure, the matrix obtained by superposing the catalytic domains of *Ca*-MPI and *St*-MPI was used to transform the entire molecule. Examination of the resulting superposition clearly shows that there is a reorganization of the helical domain between the two structures (Fig. 3*b*). The *Ca*-MPI helical domain and the transformed *St*-MPI helical domain are related by a rotation of 12°. Among the MPI structures, a fructose-bound structure is only available for *Af*-MPI. However, fructose is not bound at the active site, probably because the metal-binding site in this structure is occupied by a glycerol molecule.

Structural comparisons using the *DALI* server (Holm & Sander, 1995) showed that *St*-MPI is most similar to *Ca*-MPI (*Z* score 42.3), followed by *Af*-MPI and *Bs*-MPI. *St*-MPI also shows similarity to some functionally distinct proteins such as

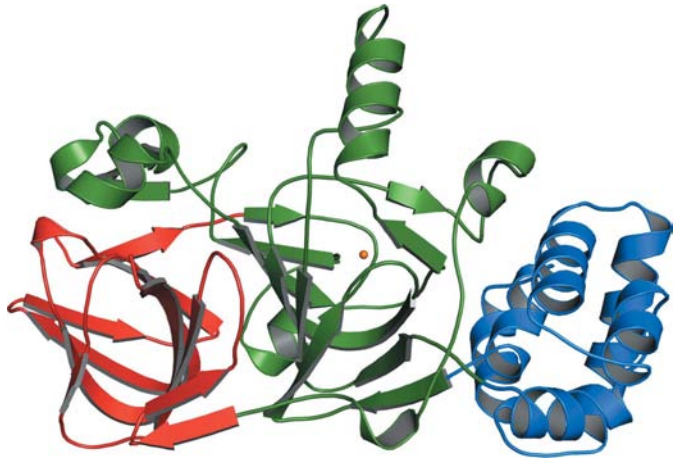

Figure 2
Cartoon representation of the polypeptide fold of *St*-MPI. The carboxy, catalytic and helical domains are shown in red, green and blue, respectively. This figure was prepared using *PyMOL* (DeLano, 2002).

pirin-like proteins (PDB code 2p17; J. Zhu, J. T. Swindell II, L. Chen, A. Ebihara, A. Shinkai, S. Kuramitsu, S. Yokoyama, Z.-Q. Fu, J. P. Rose & B.-C. Wang, unpublished work), quercetin 2,3-dioxygenase (PDB code 1tq5; Adams & Jia, 2005) and canavalin (PDB code 2cav; De Koning *et al.*, 2000).

3.3. *St*-MPI–Zn^{2+} structure (holo structure)

The *St*-MPI–Zn^{2+} structure was determined at 2.5 Å resolution. Clear electron density was located in the final difference map at a site equivalent to that of the Zn^{2+}-binding site of *Ca*-MPI. Glu134 (OE2, 2.1 Å), His99 (NE2, 2.1 Å), His255 (NE2, 2.1 Å) and HOH (1.8 Å) are liganded to Zn^{2+} in *St*-MPI. The coordination of the Zn^{2+} atom can be described as a distorted tetrahedron. The coordination of the Zn atom in *Bs*-MPI is also a distorted tetrahedron. In contrast, the coordination can be classified either as trigonal bipyramidal or octahedral (Cleasby *et al.*, 1996) in *Ca*-MPI. In *Af*-MPI, the probable metal-binding position is occupied by glycerol. Gln97 is conserved in MPIs. It is liganded to the metal in *Ca*-MPI but not in *St*-MPI owing to a change in the conformation of its side chain. The distance between Zn^{2+} and Gln OE2 in *St*-MPI-Zn^{2+} is 3.59 Å, while it is less than 2.2 Å in *Ca*-MPI.

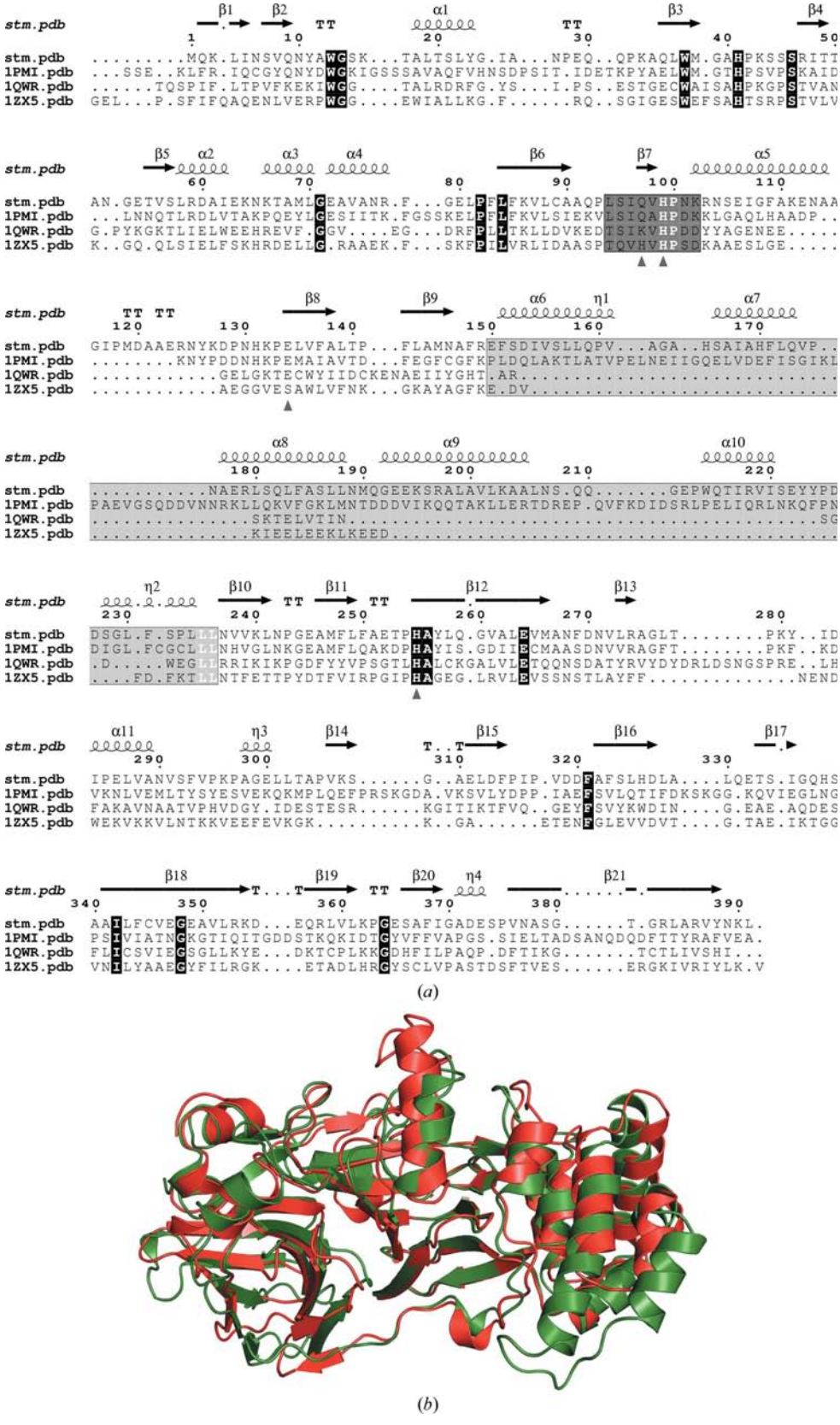

Figure 3
(*a*) Structure-based sequence alignment of *St*-MPI with three other MPIs of known structure: *Ca*-MPI (PDB code 1pmi), *Bs*-MPI (PDB code 1qwr) and *Af*-MPI (PDB code 1zx5). The sequence numbering and secondary-structure assignment shown at the top correspond to *St*-MPI. Residues that are conserved in all the sequences are shown as blocks with a black background. The dark grey box represents the conserved sequence of the type 1 MPI motif (residues 94–102 in *St*-MPI). The block in light grey represents the helical domain of variable length (23–113). The residues involved in metal binding are represented by a triangle at the bottom. Structure-based sequence alignment was performed using *MUSTANG* (Konagurthu *et al.*, 2006) and was illustrated using *ESPript* (Gouet *et al.*, 2003). (*b*) Structural superposition of *St*-MPI and *Ca*-MPI. The two structures are represented in red and green, respectively.

3.4. Structure of *St*-MPI–EDO (apo structure)

The *St*-MPI–EDO structure was solved at a relatively high resolution (1.66 Å). Owing to the long soaking (18 h at 277 K), the metal ion was replaced by EDO. Clear density was observed for EDO (Fig. 4*a*) bound at the active site. An additional 16 EDO molecules were also bound at various sites. In MPIs, although a metal ion is necessary for activity, it is only weakly bound and can easily be removed or replaced. The O1 and O2 atoms of EDO are hydrogen bonded to Glu134 OE2; O1 of EDO is also hydrogen bonded to His99 NE2 and His255 NE2. The loop residues 50–54 (the region between strands $\beta3$ and $\beta4$) are disordered in this structure, while they are ordered in the metal-bound structure (see below). Replacement of the metal also affects the conformation of the active-site loop residues 130–133 (between $\alpha5$ and $\beta8$). The loop between $\beta4$ and $\beta5$ is disordered in this structure, although it is distant from the active site and the metal-binding position.

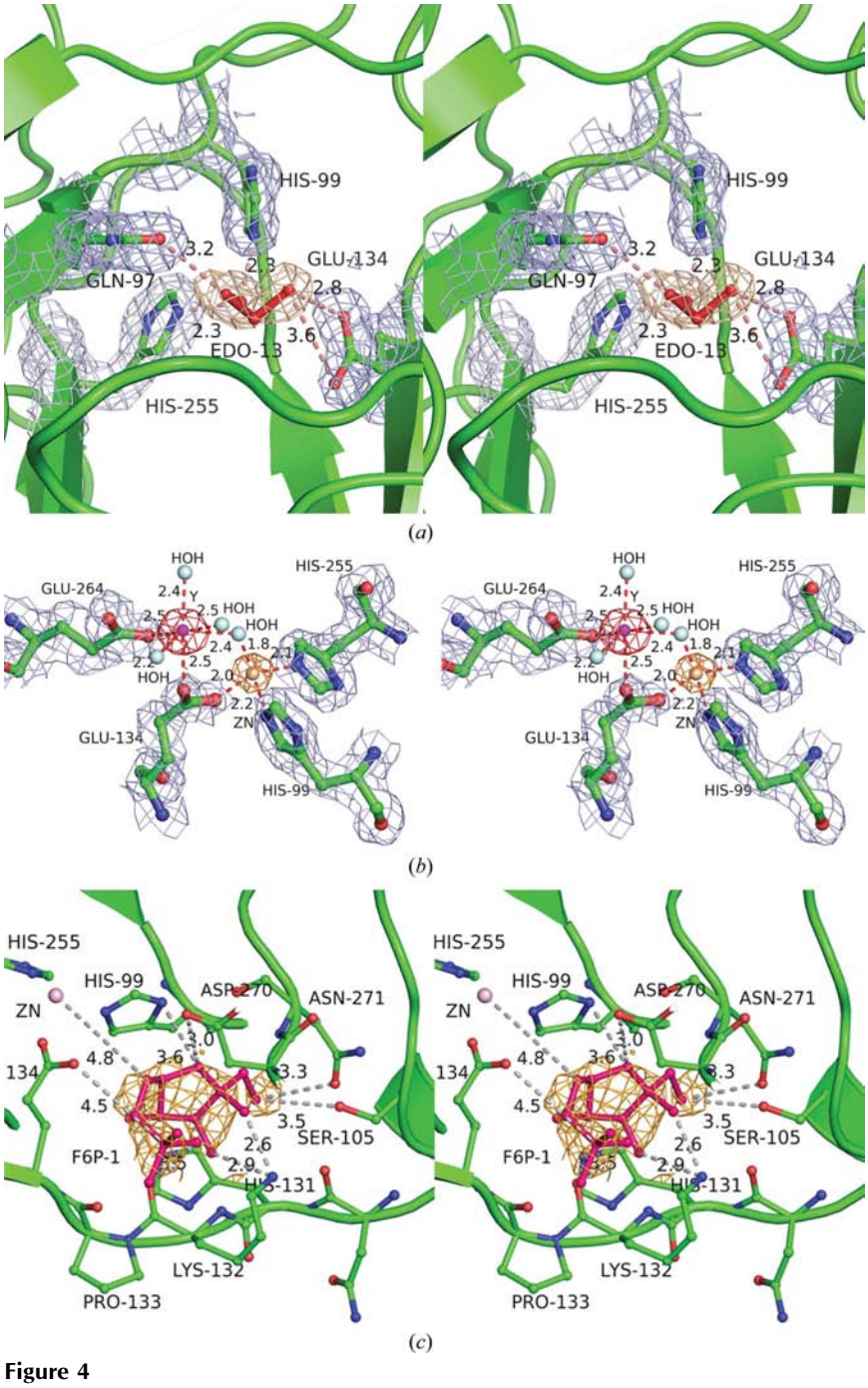

Figure 4
Stereoviews of electron-density maps. (*a*) *St*-MPI–EDO (ethylene glycol bound near the proposed metal-binding site). EDO is represented as a red stick and the corresponding electron density is shown in wheat ($2F_o - F_c$ contoured at 1.0σ). The O1 and O2 atoms of EDO are hydrogen bonded to Glu134 OE2; O1 of EDO is also hydrogen bonded to His99 NE2 and His255 NE2. (*b*) Metal atoms and the associated difference densities (8σ) at the active site of *St*-MPI–Zn^{2+}–Y^{3+}. Some of the key residues in the vicinity of the metal atoms are shown. (*c*) The geometry of the substrate-bound active site of *St*-MPI–F6P–Zn^{2+}. Zn^{2+} is represented by an orange–brown spherical ball. F6P is represented by pink sticks. The difference density map contoured at 2.4σ is shown. This figure was prepared using the program *PyMOL* (DeLano, 2002).

3.5. *St*-MPI–Zn^{2+}–Y^{3+} structure (inhibitory metal-bound structure)

It was possible to replace the bound metal ion by other divalent metal ions such as Zn^{2+}, Ni^{2+} and Mn^{2+} by successive dialysis against buffers containing 10 m*M* EDTA followed by dialysis against buffers containing the respective divalent ions. Difference density maps were calculated using the data collected from crystals obtained in the presence of different metal atoms, substrates and inhibitors. For crystals obtained in the presence of 10 m*M* Zn^{2+} and 10 m*M* yttrium (Y^{3+}), a strong difference peak ($>12\sigma$) in the electron-density map was observed at a site distinct from the Zn^{2+} position. A peak corresponding to bound zinc was observed at the position corresponding to Zn in *St*-MPI–Zn^{2+}. The Zn^{2+} is coordinated by the same residues as in *St*-MPI–Zn^{2+}. An yttrium placed at the 12σ peak could be refined to a reasonable *B* factor of 17.4 Å^2, which is comparable to those of the surrounding atoms (17 Å^2), with an occupancy of 0.5. Yttrium is liganded by the side chains of Glu134 (OE1, 2.4 Å), Glu264 (OE2, 2.5 Å) and four ordered water molecules (2.2, 2.4, 2.5 and 1.8 Å) (Fig. 4*b*). The coordination can be described as a distorted octahedron. The PDB does not contain any entries with a bound yttrium. The side chain of Glu134 contributes a ligand to both Y^{3+} and Zn^{2+}. Structural comparisons suggests that the position of yttrium in *St*-MPI corresponds to that of the bound substrates and inhibitors in several other cupin-domain proteins (PDB codes 1gqh, 1j58 and 1lrh) and hence most likely

corresponds to the inhibitory metal-binding site proposed previously on the basis of biochemical studies. Structural superposition with other MPIs shows that most of the residues in the vicinity of yttrium are conserved. These observations suggest that the inhibitory metal-binding site is probably important for the binding of substrate or intermediates during catalysis.

3.6. Structure of *St*-MPI–F6P–Zn^{2+} (substrate-bound structure)

The *St*-MPI–F6P–Zn^{2+} structure was determined at 2.0 Å resolution. Density corresponding to F6P and Zn^{2+} could be observed at the active site. The zinc is liganded to Glu134 (OE2, 2.06 Å), His99 (NE2, 2.05 Å) and His255 (NE2, 2.09 Å) (Fig. 4*c*). The coordination geometry of the metal ligand is a distorted tetrahedron. The metal–ligand distances are less than 2.2 Å. Conversion of F6P to M6P by MPI involves intermediate open forms of the substrate. Therefore, although F6P was used in soaking, the compound bound to the enzyme could be F6P or M6P in the open or closed form. The density was carefully analyzed in order to distinguish between these possibilities. The closed form of F6P filled the density better than any other form. As the electron density for F6P was weak, it was assigned an occupancy of 0.5. Subsequent refinement resulted in reasonable temperature factors that were similar to those of the surrounding atoms. Fig. 4(*c*) shows the fit of the closed form of F6P to the electron-density map. The lack of density for a few atoms could be the consequence of partial ring opening. In the closed form, F6P forms hydrogen bonds to backbone amide and carboxylate groups from several active-site residues. The phosphate O atoms O1P, O2P and O3P form hydrogen bonds to Lys132 NZ, Pro130 O, Ala267 N and Ala267 O. The bridging phosphate O6P forms a hydrogen bond to Glu134 OE2. Hydroxyl O atoms O3, O4 and O5 form hydrogen bonds to Asp270 OD2, His99 ND1, Asn271 O, Lys132 O, water molecule 123, Lys132 N and His131 ND1. F6P O1 forms hydrogen bonds to Asn271 OD1, Ser105 OG and Pro100 N. The conformation of loop 130–133 in *St*-MPI–F6P–Zn^{2+} differs from those in *St*-MPI–Zn^{2+} and *St*-MPI–Zn^{2+}–Y^{3+}. Only the conformation observed in *St*-MPI–F6P–Zn^{2+} is compatible with substrate binding. In the other conformations the loop would clash with the substrate. As revealed by the bound substrate, the catalytic site of MPI is located in a deep cavity in the central domain of the enzyme. The surface charge distribution in *St*-MPI is illustrated in Fig. 5(*a*). The active site contains a positively charged patch. There is an extensive hydrogen-bonding network in the metal-binding pocket involving water molecules and polar residues lining the walls of the cleft. Asp320 forms a hydrogen bond to Lys132, which in turn is within hydrogen-bonding distance of several well defined water molecules. The network extends to Glu134 and Ser105. Flexibility in the active site is a characteristic feature of cupin proteins (Dunwell *et al.*, 2001). The active-site residues 130–133 of *St*-MPI (apo and holo) occur in two different conformations (Fig. 5*b*) in these structures.

3.7. Implications for the reaction catalyzed by *St*-MPI

In the *St*-MPI–F6P–Zn^{2+} structure, the distances between the ligand F6P atoms and the bound Zn^{2+} are large for Zn to play a direct role in ring-opening step of catalysis. However, the metal is likely to contribute to the correct geometry of the active site as well as to stabilize the intermediates formed during isomerization. The mode of binding of F6P in the active site and its interactions with the polypeptide suggest a mechanism for the ring-opening step in the F6P→M6P

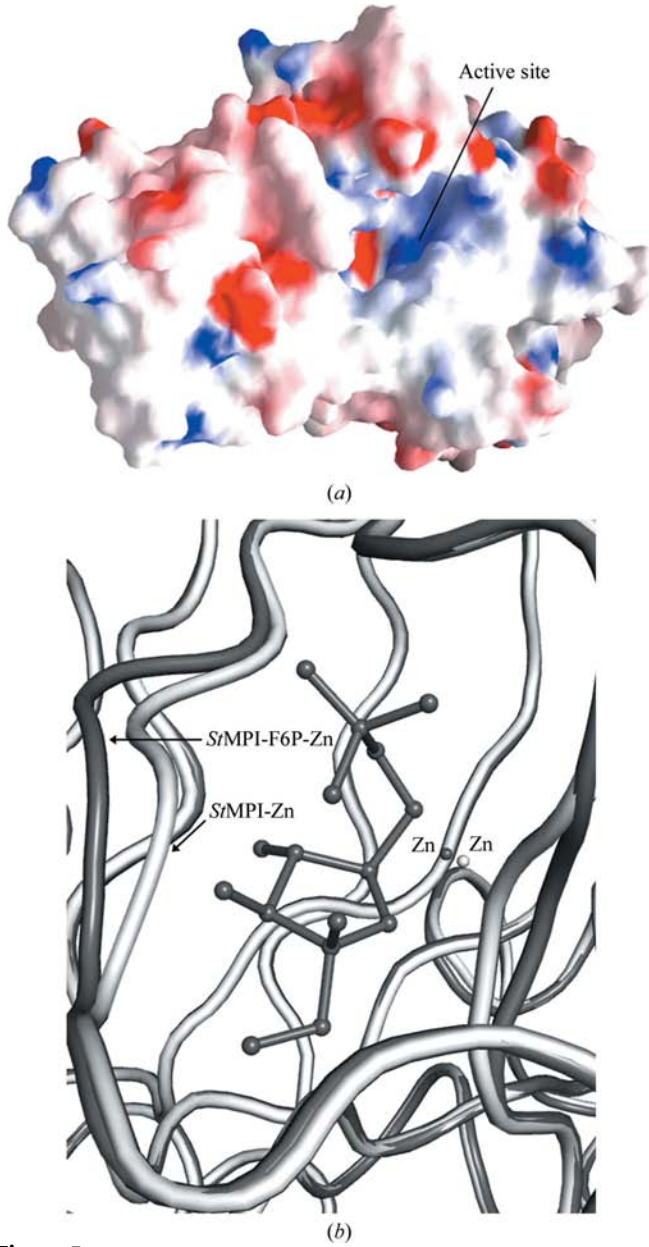

Figure 5
Active site of *St*-MPI. (*a*) Surface electrostatic potential illustrating the charge distribution in *St*-MPI. The figure was generated using *GRASP* (Nicholls *et al.*, 1991). Positively charged regions are shown in blue, negatively charged regions in red and neutral regions in white. The positively charged *St*-MPI active site is marked. (*b*) Alternative conformations of the active-site loop 130–133 of *St*-MPI. These residues change conformation upon substrate binding (dark grey, *St*MPI–F6P–Zn^{2+}; light grey, *St*-MPI–Zn^{2+}).

direction of isomerization. The cyclic form of F6P binds to the active-site cleft formed by His99, Asn271, Asp270, Lys132, His131, Pro130 and Pro100. His99 ND1 and Asp270 OD2 are at distances of 3.35 and 3.06 Å, respectively, from the ring O atom. Hence, these residues are suitable as acid groups to protonate the ring O atom. Lys132 NZ is at a distance of 2.74 Å from the C2 hydroxyl group. Therefore, it is a suitable base for deprotonating the hydroxyl. This acid/base catalysis will open the ring. His99 is conserved in all MPIs (Fig. 3*a*) as well as in other cupin-fold proteins. Asp270 is also conserved in MPIs, except for the archaeal MPI. However, catalysis of the furanose ring-opening step might not be required in the hyperthermophilic environment of *Af*-MPI. There are no suitable groups for the transfer of a proton from C1 to C2. However, as the ring opens the spatial disposition of C1 and C2 will also change and they probably approach other acid/base groups that are suitably placed for proton transfer. The observation of yttrium bound at an inhibitory site distinct from the Zn^{2+}-binding site is suggestive of this rearrangement of substrate atoms upon ring opening. Further steps of isomerization will probably follow the enediol mechanism proposed by Gracy & Noltman (1968).

In conclusion, the present studies suggest that the metal atom is important for maintenance of the active-site geometry and substrate binding. It might also be essential to stabilize the intermediates of isomerization. The conformation of the active site of MPI changes upon substrate binding as in a classical induced-fit mechanism. The first step of the reaction, the opening of the ring, is probably catalyzed by His99 and Asp270. The reaction is likely to follow the *cis*-enediol mechanism. Further insights into the mechanism of MPI catalysis will emerge from the determination of the crystal structures of MPI complexed with the open and cyclic forms of M6P and of F6P or their analogues.

The intensity data were collected using the X-ray Facility for Structural Biology at the Molecular Biophysics Unit, Indian Institute of Science supported by the Department of Science and Technology (DST) and the Department of Biotechnology (DBT) of the Government of India. We thank Professor N. Appaji Rao for useful discussions regarding enzyme assays and chemical kinetics. We also thank the Midwest Centre for Structural Genomics (MCSG) for their work on this enzyme. SSR and GG acknowledge the Council for Scientific and Industrial Research (CSIR) and DBT, Government of India for the award of fellowships.

References

Adams, M. & Jia, Z. (2005). *J. Biol. Chem.* **280**, 28675–28682.
Awadalla, H. N., Mansour, N. S. & Mohareb, E. W. (1987). *Trans. R. Soc. Trop. Med. Hyg.* **81**, 915–917.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cleasby, A., Wonacott, A., Skarzynski, T., Hubbard, R. E., Davies, G. J., Proudfoot, A. E., Bernard, A. R., Payton, M. A. & Wells, T. N. (1996). *Nature Struct. Biol.* **3**, 470–479.
Cohen, G. E. (1997). *J. Appl. Cryst.* **30**, 1160–1161.
De Koning, T. J., Dorland, L., van Diggelen, O. P., Boonman, A. M., de Jong, G. J., van Noort, W. L., De Schryver, J., Duran, M., van den Berg, I. E., Gerwig, G. J., Berger, R. & Poll-The, B. T. (1998). *Biochem. Biophys. Res. Commun.* **245**, 38–42.
De Koning, T. J., Nikkels, P. G., Dorland, L., Bekhof, J., De Schrijver, J. E., van Hattum, J., van Diggelen, O. P., Duran, M., Berger, R. & Poll-The, B. T. (2000). *Virchows Arch.* **437**, 101–105.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* http://www.pymol.org.
DeRossi, C., Bode, L., Eklund, E. A., Zhang, F., Davis, J. A., Westphal, V., Wang, L., Borowsky, A. D. & Freeze, H. H. (2006). *J. Biol. Chem.* **281**, 5916–5927.
Dunwell, J. M. (1998). *Biotechnol. Genet. Eng. Rev.* **15**, 1–32.
Dunwell, J. M., Culham, A., Carter, C. E., Sosa-Aguirre, C. R. & Goodenough, P. W. (2001). *Trends Biochem. Sci.* **26**, 740–746.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Gao, H., Yu, Y. & Leary, J. A. (2005). *Anal. Chem.* **77**, 5596–5603.
Gouet, P., Robert, X. & Courcelle, E. (2003). *Nucleic Acids Res.* **31**, 3320–3323.
Gowda, G., Sagurthi, S. R., Savithri, H. S. & Murthy, M. R. N. (2008). *Acta Cryst.* F**64**, 81–84.
Gracy, R. W. & Noltmann, E. A. (1968). *J. Biol. Chem.* **243**, 5410–5419.
Holm, L. & Sander, C. (1995). *Trends Biochem. Sci.* **20**, 478–480.
Jaeken, J., Matthijs, G., Saudubray, J. M., Dionisi-Vici, C., Bertini, E., de Lonlay, P., Henri, H., Carchon, H., Schollen, E. & Van Schaftingen, E. (1998). *Am. J. Hum. Genet.* **62**, 1535–1539.
Jensen, S. O. & Reeves, P. R. (1998). *Biochim. Biophys. Acta*, **1382**, 5–7.
Kabsch, W. & Sander, C. (1983). *Biopolymers*, **22**, 2577–2637.
Konagurthu, A. S., Whisstock, J. C., Stuckey, P. J. & Lesk, A. M. (2006). *Proteins*, **64**, 559–574.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* D**61**, 458–464.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Nicholls, A., Sharp, K. A. & Honig, B. (1991). *Proteins*, **11**, 281–296.
Papoutsopoulou, S. V. & Kyriakidis, D. A. (1997). *Mol. Cell. Biochem.* **177**, 183–191.
Proudfoot, A. E., Payton, M. A. & Wells, T. N. (1994). *J. Protein Chem.* **13**, 619–627.
Roux, C., Gresh, N., Perera, L. E., Piquemal, J. P. & Salmon, L. (2007). *J. Comput. Chem.* **28**, 938–957.
Roux, C., Lee, J. H., Jeffery, C. J. & Salmon, L. (2004). *Biochemistry*, **43**, 2926–2934.
Seeholzer, S. H. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 1237–1241.
Shinabarger, D., Berry, A., May, T. B., Rothmel, R., Fialho, A. & Chakrabarty, A. M. (1991). *J. Biol. Chem.* **266**, 2080–2088.
Smith, D. J., Proudfoot, A. E., Detiani, M., Wells, T. N. & Payton, M. A. (1995). *Yeast*, **11**, 301–310.
Swan, M. K., Hansen, T., Schonheit, P. & Davies, C. (2004). *Biochemistry*, **43**, 14088–14095.
Wells, T. N., Coulin, F., Payton, M. A. & Proudfoot, A. E. (1993). *Biochemistry*, **32**, 1294–1301.