Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of mannosyl-3-phosphoglycerate phosphatase from *Pyrococcus horikoshii*

Takashi Kawamura,[a] Nobuhisa Watanabe[b] and Isao Tanaka[c]*

[a]Division of Biological Sciences, Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan, [b]Synchrotron Radiation Research Center, Nagoya University, Nagoya 464-8603, Japan, and [c]Division of Molecular Life Science, Faculty of Advanced Life Science, Hokkaido University, Sapporo 060-0810, Japan

Correspondence e-mail: tanaka@castor.sci.hokudai.ac.jp

Received 12 June 2008
Accepted 16 October 2008

PDB References: mannosyl-3-phosphoglycerate phosphatase, apo form, 2zos, r2zossf; holo form, 1wzc, r1wzcsf.

Mannosyl-3-phosphoglycerate phosphatase (MPGP) catalyzes the dephosphorylation of α-mannosyl-3-phosphoglycerate (MPG) to produce α-mannosylglycerate (MG). In the hyperthermophile *Pyrococcus horikoshii*, MPGP plays a role in a series of enzyme reactions that are involved in the MG-biosynthesis pathway, which is important for maintaining life under conditions of high salt concentration. Crystal structures of *P. horikoshii* MPGP (*Pho*MPGP) in the holo form and in the apo form lacking the magnesium ion were determined by the multiple-wavelength anomalous diffraction method using SeMet-substituted *Pho*MPGP. *Pho*MPGP consists of two domains: a core domain that is conserved in the haloacid dehalogenase superfamily and a cap domain that is specific to the C2B cap subclass of the superfamily. Apo-form crystals contain two *Pho*MPGP molecules: one in the open conformation and the other in the closed conformation. In holo-form crystals both of the two molecules are in the closed conformation with phosphate and magnesium ions. *Pho*MPGP has a specific hairpin loop that is bent towards the active site in the closed conformation of both the apo and holo forms. *Pho*MPGP has a cavity between the two domains which is considered to be the substrate-binding site as a phosphate ion is located in the cavity, mimicking the binding manner of the phosphate group of MPG. The cavity is sequestered in the closed conformation such that a conformational change is indispensable for the release of products. A salt bridge from the general acid/base Asp10 to Arg170 is observed in the holo-form *Pho*MPGP which is not present in the open form. The importance of the conformational change in the activity of *Pho*MPGP is discussed.

1. Introduction

Mannosyl-3-phosphoglycerate phosphatase catalyzes the dephosphorylation of mannosyl-3-phosphoglycerate (MPG) to produce α-mannosylglycerate (MG), which is one of the major compatible solutes in thermophiles and hyperthermophiles (Empadinhas & Costa, 2006). Compatible solutes are defined as low-molecular-weight hydrophilic organic compounds that are accumulated in cells with no harm. MG is thought to act as a thermoprotectant and osmoprotectant because it is accumulated in the thermophile *Rhodothermus marinus* to confer resistance to high-temperature shock and high-salt concentration shock using different enzymes (Borges *et al.*, 2004). Non-thermophilic bovine RNase A was stabilized by the addition of MG at relatively high temperatures (Faria *et al.*, 2003).

MG is synthesized by two specific enzymes: MPGP and MPG synthase. The genes encoding these enzymes are con-

served in thermophiles and hyperthermophiles such as *R. marinus*, *Thermus thermophilus* HB27 and the mesophile *Dehalococcoides ethenogenes* (Borges *et al.*, 2004; Empadinhas *et al.*, 2003, 2004). In addition to these genes, the hyperthermophile *Pyrococcus horikoshii* possesses two further genes encoding α-phosphomannomutase and a bifunctional phosphomannose isomerase/mannose-1-phosphate guanylyltransferase located in the genome in a tandem manner. This suggests the existence of a pathway for MG synthesis (Fig. 1) from fructose-6-phosphate through five reactions performed by the four enzymes (Empadinhas *et al.*, 2001). The final reaction in this pathway is performed by *P. horikoshii* MPGP (*Pho*MPGP).

*Pho*MPGP belongs to the C2 cap group of the haloacid dehalogenase superfamily (HADSF), which is a large family containing haloacid dehalogenases, P-type ATPases, β-phosphoglucomutases and various phosphatases. The sequences and structures of the HADSF have been investigated (Burroughs *et al.*, 2006) and the members of this family have been classified into three groups: the C0 cap group consisting of single core-domain enzymes, the C1 cap group containing a cap domain consisting of a four-helix bundle and the C2 cap group containing a cap domain consisting of an $\alpha+\beta$ fold. HADSF has four characteristic sequence motifs, among which –D*X*D– in sequence motif 1 is highly conserved owing to catalytic necessity.

The catalytic mechanism of the phosphatase reaction in the HADSF has been investigated in the case of *Methanococcus* (*Methanocaldococcus*) *jannaschii* phosphoserine phosphatase (*MJ*PSP; Wang *et al.*, 2002). The suggested mechanism is as follows. (i) Phospho-L-serine is bound to the active site. (ii) The side chain of Asp11, the first Asp of sequence motif 1, attacks the P atom, generating the pentacovalent phosphorus intermediate. (iii) The covalent bond between the P atom and the O atom of the leaving serine molecule is broken. (iv) The phosphate moiety transferred to the side chain of Asp11 is attacked by a water molecule, again generating the pentacovalent phosphorus intermediate. (v) Finally, the covalent bond between the P atom and the O atom of the side chain of Asp11 is cleaved and the phosphate moiety is released as a phosphate ion. In this series of reactions, a magnesium ion is required for activation of the phosphate group. Asp13, the second Asp in sequence motif 1, is considered to be the general acid/base.

We determined the crystal structure of the *Pho*MPGP apo form lacking magnesium ions using the multiple-wavelength anomalous diffraction (MAD) method and the structure of the magnesium-bound holo form using the molecular-replacement (MR) method. In these structures, open and closed conformations of *Pho*MPGP were observed. The conformational differences between the two forms and their importance in the catalytic mechanism are discussed.

2. Materials and methods

2.1. *Pho*MPGP expression

All reagents used in this study were of analytical grade. The columns used for chromatography were purchased from Amersham Biosciences (now GE Healthcare Biosciences).

The *Pho*MPGP gene was amplified by PCR using *Pfu* Turbo polymerase (Stratagene) and two primers (5′-GGAATTC-**CATATG**ATTAGGTTAATATTCCTGG-3′ and 5′-CCG**CTC-GAG**TTTGATCACCTCCAAAACATC-3′; *Nde*I and *Xho*I recognition sites are in bold) with the *P. horikoshii* genome as the template. The amplified gene was inserted into the *Nde*I/*Xho*I restriction sites of pET-22b vector (Novagen) for expression of the C-terminally His_6-tagged fusion protein.

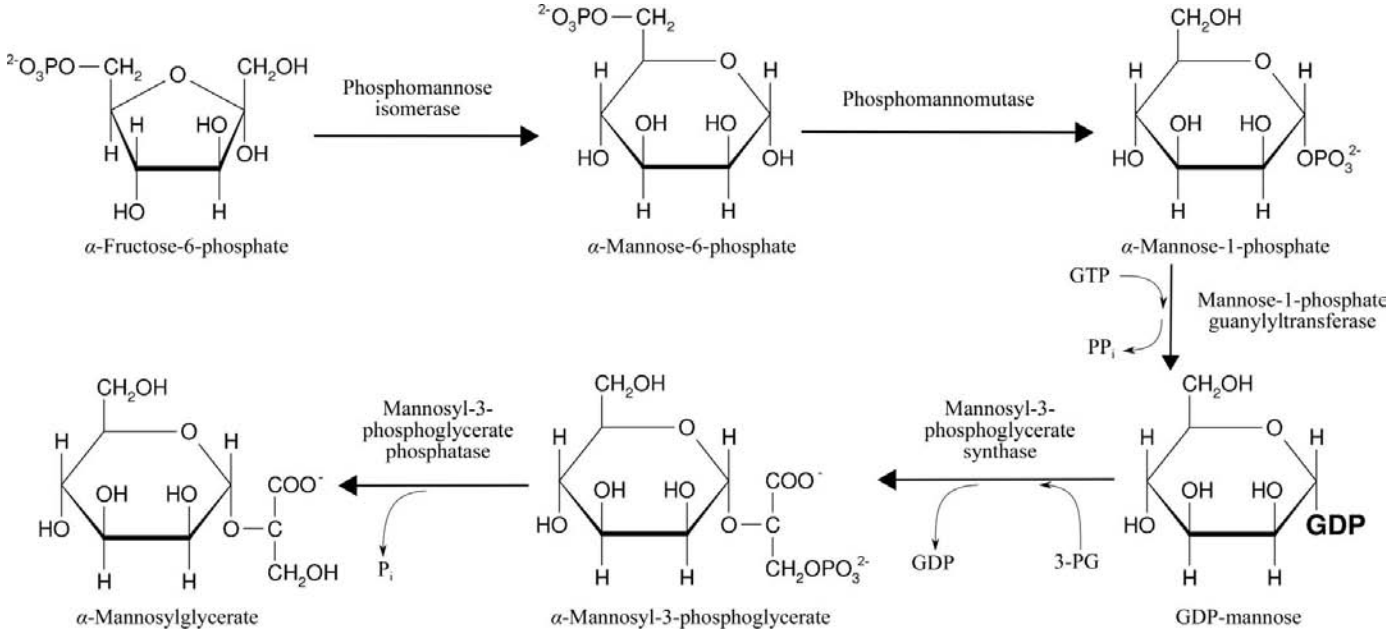

Figure 1
The mannosylglycerate-biosynthesis pathway in the hyperthermophile *P. horikoshii* (Empadinhas *et al.*, 2001).

Table 1
Data-collection statistics.

Values in parentheses are for the outer shell.

	SeMet			
	Peak	Edge	Remote	Native
Space group	$P2_1$			$P2_1$
Unit-cell parameters				
a (Å)	58.4			59.8
b (Å)	68.6			70.7
c (Å)	67.5			67.9
β (°)	96.5			98.1
Wavelength (Å)	0.9791	0.9793	0.9641	1.0000
Resolution range (Å)	50.0–1.69 (1.75–1.69)	50.0–1.69 (1.75–1.69)	50.0–1.65 (1.73–1.67)	50.0–1.90 (1.97–1.90)
No. of reflections	213941	213644	222313	167220
No. of unique reflections	59147	59034	61540	44398
Completeness (%)	99.5 (98.6)	99.5 (95.0)	99.8 (98.2)	99.0 (100)
$R_{\rm merge}$† (%)	7.1 (30.6)	6.7 (21.8)	6.9 (30.6)	4.7 (40.0)
$I/\sigma(I)$	16.0 (3.0)	16.8 (3.4)	14.6 (2.3)	17.8 (4.1)
Redundancy	3.6 (2.9)	3.6 (2.9)	3.6 (2.8)	3.8 (3.7)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean intensity of a set of equivalent reflections.

Table 2
Refinement statistics.

	Apo form (2zos)	Holo form (1wzc)
Resolution (Å)	50.0–1.70	50.0–1.90
$R_{\rm work}$†	0.211	0.219
$R_{\rm free}$‡	0.234	0.253
No. of atoms in refined model		
Protein	3912	3896
Water	291	102
Phosphate ion		10
Magnesium ion		2
R.m.s. deviations from ideal		
Bond lengths (Å)	0.005	0.009
Bond angles (°)	1.2	1.5
Dihedral angles (°)	23.0	23.3
Improper angles (°)	0.8	0.9
Ramachandran plot		
Most favoured (%)	91.6	89.4
Additional allowed (%)	7.4	10.6
Generously allowed (%)	1.0	0.0
Estimated coordinate error		
From Luzzati plot (Å)	0.21	0.26
From σ (Å)	0.11	0.25

† $R_{\rm work} = \sum |F_{\rm obs} - F_{\rm calc}| / \sum F_{\rm obs}$, where $F_{\rm obs}$ and $F_{\rm calc}$ are the observed and calculated structure-factor amplitudes, respectively. ‡ The $R_{\rm free}$ value was calculated with a random 5% subset of all reflections excluded from the refinement.

Recombinant *Pho*MPGP was expressed in *Escherichia coli* B834 (DE3) cells. LB broth or S9 minimum broth with addition of seleno-L-methionine (SeMet) for SeMet-substituted *Pho*MPGP expression were used as media. Bacterial cultures were grown at 310 K with each medium containing 50 µg ml^{-1} ampicillin to an OD_{600} of 0.6–1.0. *Pho*MPGP expression was induced by addition of isopropyl β-D-1-thiogalactopyranoside to a final concentration of 0.4 m*M*. Cells were cultivated for 16 h after induction and then harvested by centrifugation at 4000*g* for 15 min. To remove residual broth, the cell pellet was suspended in cell-lysis buffer (50 m*M* Tris–HCl pH 7.5, 150 m*M* NaCl) and centrifuged again under the same conditions. Cell pellets were frozen until further analysis.

2.2. *Pho*MPGP purification

Cells were thawed on ice and suspended in cell-lysis buffer. Lysozyme (1 mg ml^{-1} final concentration) and DNase I were added to the suspended cells, which were incubated for 1 h on ice. After the cells had been lysed by ultrasonication on ice, cell debris was removed by centrifugation (40 000*g*, 277 K, 30 min). The supernatant was incubated in a water bath at 343 K for 30 min and then centrifuged (40 000*g*, 277 K, 30 min) to remove aggregated host-cell proteins. Ammonium sulfate was dissolved in the supernatant to 60% saturation and the precipitate was retrieved by centrifugation (20 000*g*, 277 K, 30 min). The precipitate was dissolved and dialyzed against dialysis buffer (20 m*M* Tris–HCl pH 7.5, 150 m*M* NaCl).

Dialyzed samples were loaded onto a HiTrap Chelating HP column containing immobilized Ni^{2+} ions and equilibrated with dialysis buffer. To remove nonspecifically bound protein, the column was washed with 15 column volumes of dialysis buffer and ten column volumes of dialysis buffer containing 10 m*M* imidazole. *Pho*MPGP was eluted with dialysis buffer containing 200 m*M* imidazole. Fractions containing *Pho*MPGP were pooled and loaded onto a HiLoad Superdex 75 26/60 prep-grade column equilibrated with gel-filtration buffer (20 m*M* Tris–HCl pH 7.5, 150 m*M* NaCl), which was used for isocratic elution of *Pho*MPGP. Fractions containing *Pho*MPGP were pooled and dialyzed against binding buffer (20 m*M* Tris–HCl pH 7.5) and loaded onto a Source 15Q 4.6/100 PE column equilibrated with binding buffer. *Pho*MPGP was eluted with a linear NaCl gradient (0–500 m*M*) in 30 column volumes. Fractions containing *Pho*MPGP were dialyzed against crystallization buffer [10 m*M* MOPS–NaOH pH 7.0, 5%(*v*/*v*) glycerol].

2.3. Crystallization of *Pho*MPGP

Purified *Pho*MPGP solution was concentrated by ultrafiltration using Amicon Ultra 4 30 kDa cutoff filters (Millipore) and mixed with an equal volume of crystallization solution containing 16–20%(*w*/*v*) PEG 3350, 0.2 *M* ammonium chloride and 0.1 *M* MES–NaOH pH 6.0. *Pho*MPGP crystals were grown in 4 µl of the mixed solutions by the hanging-drop vapour-diffusion method using the crystallization solution as the mother liquor. Native *Pho*MPGP and SeMet-substituted *Pho*MPGP were purified and crystallized according to the same procedure.

2.4. X-ray data collection and data reduction

Diffraction data from SeMet-substituted *Pho*MPGP for use in the MAD method and native data were collected using

X-ray beamlines BL-5A and BL-6A at the Photon Factory, Japan. For low-temperature measurement, a *Pho*MPGP crystal was dipped into a drop of crystallization solution containing cryoprotectant [20%(*v*/*v*) glycerol, 20%(*w*/*v*) PEG 3350, 0.2 *M* ammonium chloride, 0.1 *M* MES–NaOH] and then mounted on a nylon loop. The crystal was cooled by sudden exposure to a stream of cold nitrogen gas. Data collection was performed at 95 K. The collected data were processed using the *HKL*-2000 program package (Otwinowski & Minor, 1997). Scaled intensity data were converted into amplitude data using the program *TRUNCATE* (French & Wilson, 1978). Statistics of the diffraction data are listed in Table 1.

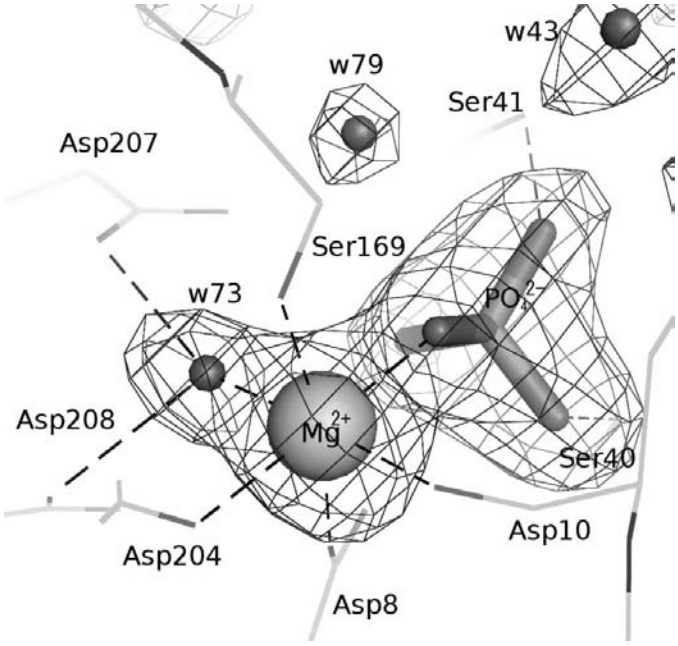

Figure 2
The phosphate ion (stick model) and the magnesium ion (large sphere) in the active site of the holo-form *Pho*MPGP crystal structure. The $F_o - F_c$ OMIT map (contoured 3σ) was calculated with the program *CNS* using observed intensities and two *Pho*MPGP molecules as the model for phase calculation. The dashed lines indicate electrostatic interactions between *Pho*MPGP (shown as solid lines) and the ligands. Small spheres indicate water molecules (numbered w43, w73 and w79 in the *Pho*MPGP structure deposited in the PDB with code 1wzc).

2.5. Structure determination

Initially, the SeMet-substituted *Pho*MPGP crystal structure was determined using the MAD method. Nine Se atoms were detected (of the total of 14 Se atoms, including the N-terminal SeMet) by the program *SHELXD* (Sheldrick *et al.*, 2001). The coordinates of the Se atoms were refined and phase calculation was performed with the program *SHARP* (de La Fortelle & Bricogne, 1997). After density modification with the program *DM* (Cowtan, 1994), the phases were refined using the program *SOLVE* (Terwilliger, 2000) and an initial model was built with the program *ARP*/*wARP* (Perrakis *et al.*, 1997). The $2mF_o - DF_c$ and $mF_o - DF_c$ maps generated by the program *CNS* (Brünger *et al.*, 1998) were used for additional model building and manual fitting with the program *O* (Jones *et al.*, 1991). Positional and temperature-factor refinement and simulated annealing were performed with the program *CNS*, using data in the resolution range 50–1.7 Å. A randomly chosen 5% of all observed reflections were set aside for cross-validation analysis.

For native structure determination, the structure of SeMet *Pho*MPGP was used as the search model in the MR method

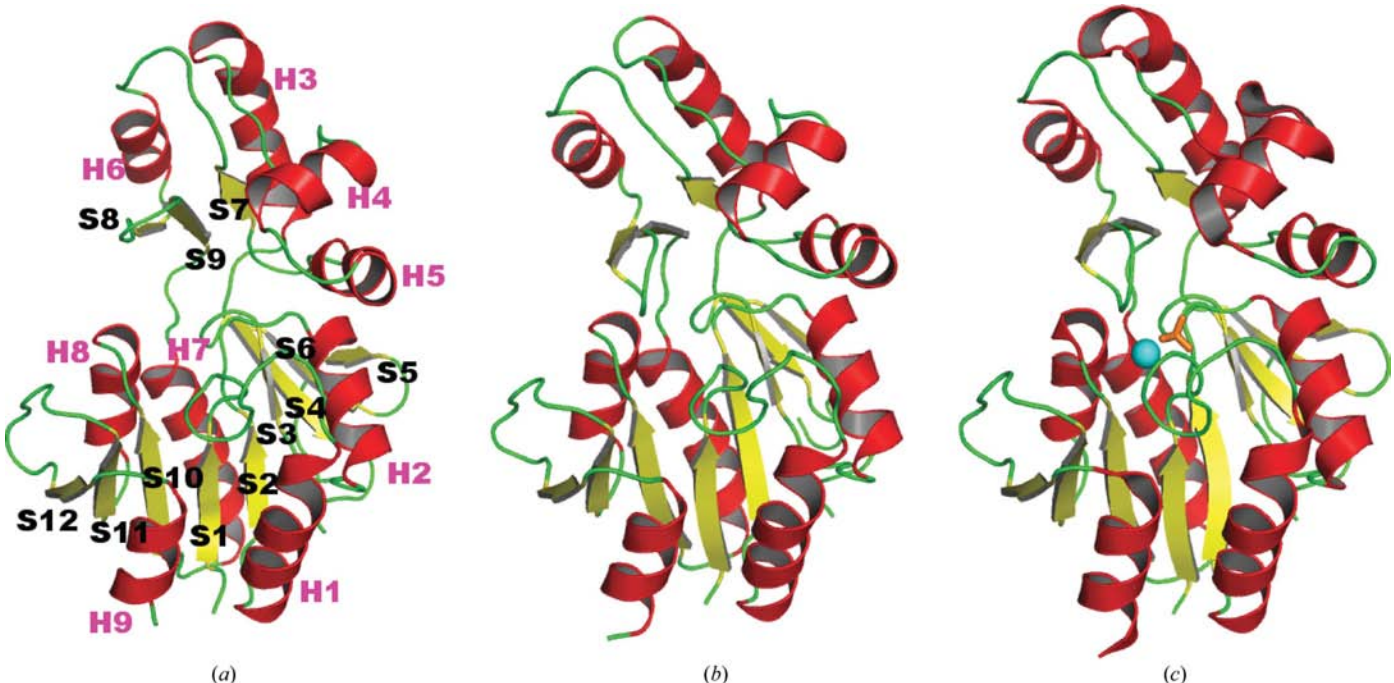

Figure 3
Monomer structures observed in *Pho*MPGP crystals. (*a*) Apo-form open conformation. (*b*) Apo-form closed conformation. (*c*) Holo-form closed conformation, with the phosphate and magnesium ions shown as a stick model and as a sphere, respectively.

using the program *MOLREP* (Vagin & Teplyakov, 1997). Model refinement of the native *Pho*MPGP structure was performed in the same way as in the SeMet case. Refined structures were validated with the program *PROCHECK* (Laskowski *et al.*, 1993). Statistics of the final structures are listed in Table 2.

For structure analysis, the following programs were used: *LSQKAB* (Kabsch, 1976) for protein-structure superposition, *DynDom* (Hayward & Berendsen, 1998) for detecting hinge motion and *VOIDOO* (Kleywegt & Jones, 1994) for cavity detection. Figures of protein structures were prepared with *PyMOL* (DeLano, 2002). The topology diagram of the secondary structure was prepared with the program *TopDraw* (Bond, 2003). *TRUNCATE*, *DM*, *MOLREP*, *PROCHECK*, *LSQKAB* and *TopDraw* are from the *CCP*4 program suite (Collaborative Computation Project, Number 4, 1994).

3. Results

3.1. General feature of the crystal structure

The asymmetric unit of the SeMet *Pho*MPGP crystal contains two protein molecules. A total of 482 residues (478 residues including side chains and four residues with only main chain) and 290 water molecules were constructed. The asymmetric unit of the native *Pho*MPGP crystal contains 478 residues from two protein molecules and 102 water molecules. A nonprotein peak was observed in the $F_o - F_c$ electron-density map of the native *Pho*MPGP crystal using *Pho*MPGP molecules for phase calculation. We assigned these peaks as phosphate and magnesium ions (Fig. 2), considering the phosphatase activity and the magnesium-ion dependency of this enzyme. These ions were not introduced in the purification and crystallization procedures and were therefore considered to have been salvaged from the host cell. All *Pho*MPGP structures underwent one spontaneous mutation (A218V) during PCR processing.

As the magnesium ion is required for full activity of *Pho*MPGP (Empadinhas *et al.*, 2001), we called the crystal structure of native *Pho*MPGP the holo form, while the SeMet structure lacking the magnesium ion was designated as the apo form in the following discussion. The apo-form crystal has two

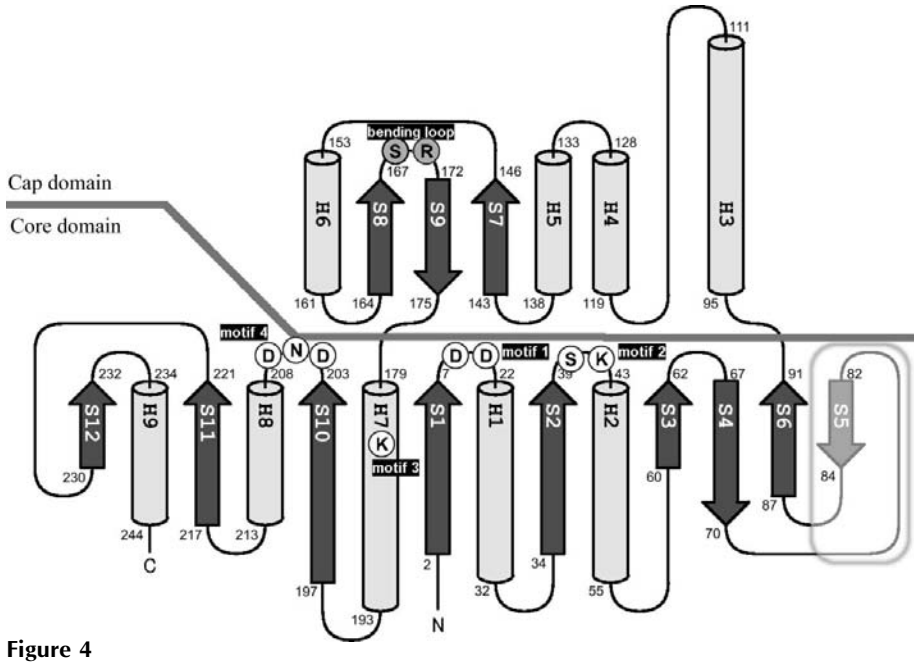

Figure 4
Topology diagram of the secondary structure in *Pho*MPGP. *Pho*MPGP consists of two domains: the cap domain and the core domain. Light grey sticks indicate α-helices. Dark grey arrows indicate β-strands. The conserved residues in the four sequence motifs of the HADSF are indicated by open circles: Asp8 and Asp10 of motif 1 between S1 and H1, Ser40 and Lys42 of motif 2 between S2 and H2, Lys180 of motif 3 in H7 and Asp204, Asn207 and Asp208 of motif 4 between S10 and H8. The grey circles between S8 and S9 indicate Ser169 and Arg170, which participate in the active-site architecture (see Fig. 6*a*). Strand S5, enclosed by the rounded rectangle, is disordered in the closed conformations of apo-form and holo-form *Pho*MPGP.

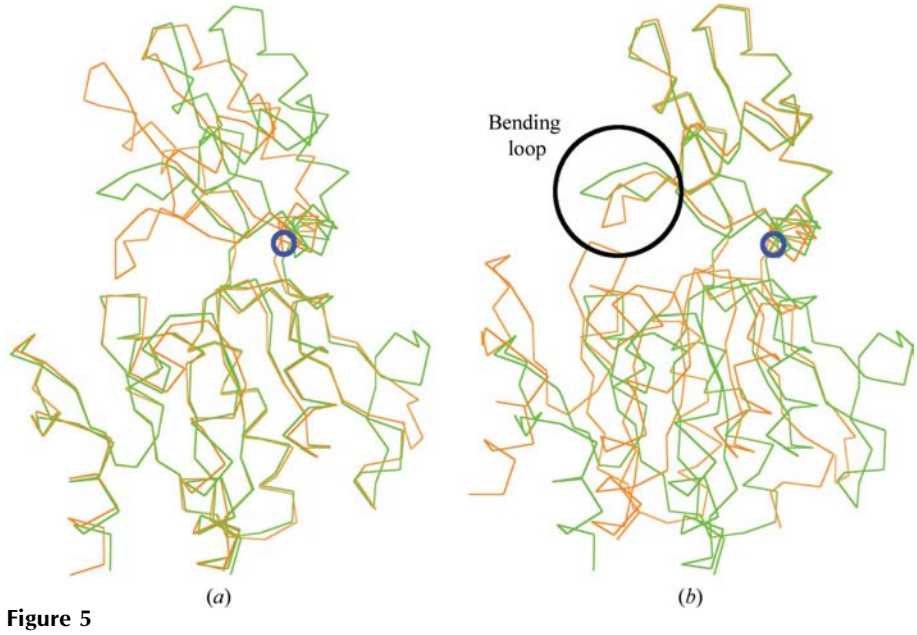

Figure 5
Superposition of the open (green) and closed (orange) conformations of the apo-form *Pho*MPGP crystal. The small blue circles indicate the axes of the hinge motion detected by the program *DynDom* (perpendicular to the paper). (*a*) The main chains of the core domain are superposed. (*b*) The main chains of the cap domains are superposed. The bending loops of the two structures (large circle) are not matched.

*Pho*MPGP molecules in the asymmetric unit that have different domain conformations: one is in the open conformation (Fig. 3*a*) and the other in the closed conformation (Fig. 3*b*). In the holo-form crystal both molecules are in the closed conformation (Fig. 3*c*). The polypeptide chains of residues 77–86 of the closed conformation in the apo form and of residues 84–87 in the holo form were not constructed because the electron density was poor.

3.2. Overall structure

The *Pho*MPGP molecule consists of two domains, the cap domain and the core domain, which are typical of the HADSF. The core domain includes the N- and C-termini of the polypeptide chain; the cap domain (residues 95–175) is inserted in the centre of the polypeptide chain. The core domain has a Rossmann-like fold (Fig. 4), in which one β-sheet (the order of the β-strands is S12, S11, S10, S1, S2, S3, $-$S4, S6 and $-$S5, where '$-$' indicates an inverted direction of the strand) is sandwiched by two sets of α-helices (H9, H1, H2 and H8, H7). The core domain of *Pho*MPGP contains four sequence motifs that are conserved in the HADSF. These motifs make up an active site as observed in other HAD hydrolase structures. The relatively high temperature factors of the active-site residues in the apo form suggest flexibility of the side chains in the active site (data not shown). The cap domain has one $\alpha+\beta$ structure (a sheet consisting of S8, $-$S9 and S7 with H3, H4, H5 and H6). Three β-strands (S7, S8 and S9) in the C-terminus of the cap domain are folded into a three-stranded antiparallel β-sheet that is the basic unit of the C2B cap subclass (Burroughs *et al.*, 2006). Two polypeptide loops (92–94 and 176–178) connect the core domain and the cap domain.

The open and closed conformations were compared by superposing their main chains using the program *LSQKAB*. The core domains of each conformation superposed well, with an r.m.s. distance of 0.54 Å for the main-chain C^{α} atoms (Fig. 5*a*). The cap domains also superposed well, with an r.m.s. distance of 0.76 Å for the main-chain C^{α} atoms, excluding one hairpin loop between β-strands S8 and S9 (Fig. 5*b*). The similarity of the fold in each domain indicated that the change between the two conformations occurred with a hinge motion. From analysis using the program *DynDom* (Hayward & Berendsen, 1998), the rotation angle of the hinge motion was shown to be about 19° around an axis that penetrates α-helix H5 and the loop region connecting the two domains.

In the closed conformation, a hairpin loop between β-strands S8 and S9 is bent toward the active site in the core domain (the region in the circle in Fig. 5*b*); therefore, we refer to this as the bending loop in the following discussion. The bending loop is stabilized by this positional shift, as indicated by the observation that the mean temperature factor of the bending loop (166–172) was lower in the closed conformation (22.1 $Å^2$ for main-chain atoms, 25.8 $Å^2$ for side-chain atoms and 24.1 $Å^2$ overall) than in the open conformation (38.7, 42.5 and 44.1 $Å^2$, respectively). These differences were significant in comparison with the slight differences in the mean temperature factors of all atoms in each conformation (22.4, 26.3 and 24.4 $Å^2$ in the closed conformation, and 21.1, 25.3 and 23.2 $Å^2$ in the open conformation, respectively).

The two crystal forms used in this study had slightly different unit-cell parameters (see Table 1). The difference coincided with a shift of the *Pho*MPGP molecule in addition to the conformation change. The distances between the core domains of the two molecules in the asymmetric unit were 34.5 Å in the apo form and 39.5 Å in the holo form. The two molecules in the asymmetric unit interacted with each other less in holo-form *Pho*MPGP. The variation of the interaction between the two *Pho*MPGP molecules in the asymmetric unit indicated that *Pho*MPGP acts as a monomer. This supposition

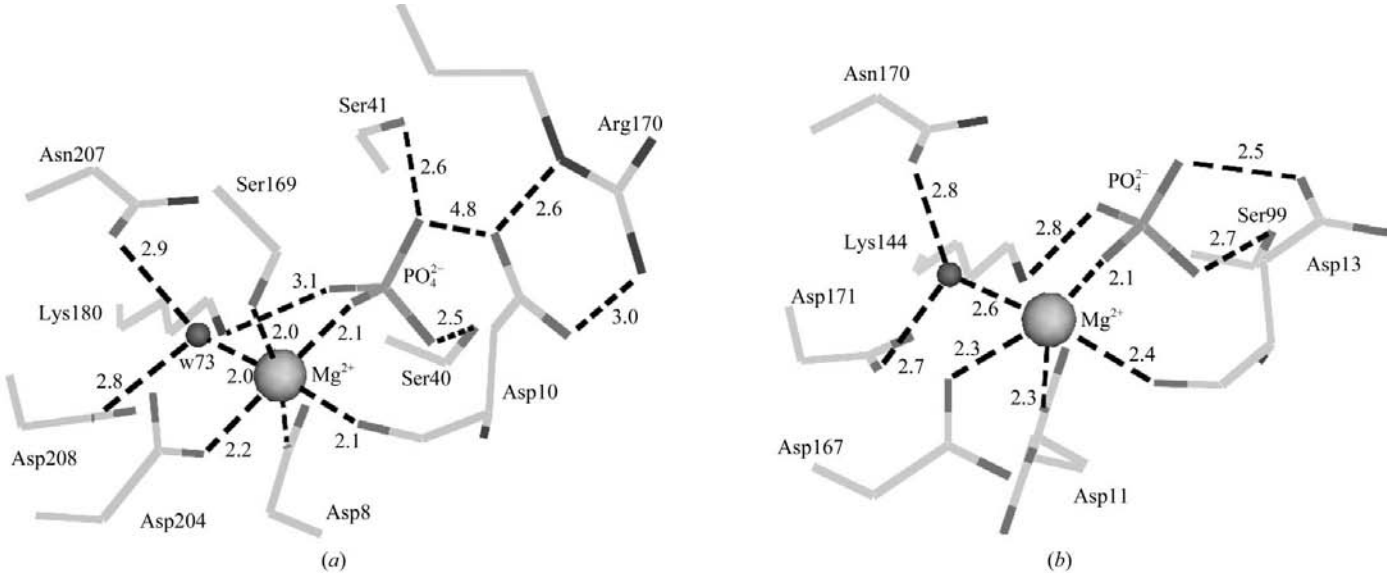

Figure 6
The active sites of (*a*) holo-form *Pho*MPGP and (*b*) *MJ*PSP (PDB code 1l7m). Electrostatic interactions are shown by dashed lines with distances in Å. Ser41, Ser169 and Arg170 of *Pho*MPGP have no counterparts in *MJ*PSP.

is consistent with the result of gel-filtration chromatography, in which the elution position corresponds to a monomer of *Pho*MPGP (29 kDa).

The holo-form *Pho*MPGP structure consists of two *Pho*MPGP molecules in the closed conformation which resemble that observed in apo-form *Pho*MPGP. The r.m.s. distance for main-chain C^{α} atoms is 0.66 Å on superposing closed conformations from the apo-form and holo-form structures. Binding the phosphate ion and the magnesium ion produces differences in the closed conformation of the holo form compared with that of the apo form; the loop between β-sheet S1 and α-helix H1, including the conserved sequence motif 1, is shifted. Newly observed interactions are hydrogen bonds between the phosphate ion and the main-chain amide groups of Ile9 and Asp10 and coordination between the magnesium ion and the main-chain carbonyl group of Asp10.

3.3. Comparison between holo-form *Pho*MPGP and *MJ*PSP

*Pho*MPGP has four sequence motifs that are conserved amongst HADSF enzymes and are involved in the active-site architecture. We compared the active site of the holo form of *Pho*MPGP with that of *MJ*PSP and the results indicated that the catalytic residues are spatially conserved. However, *Pho*MPGP and *MJ*PSP differ in the cap domain (*MJ*PSP has a cap domain consisting of a four-helix bundle fold that is typical of the C1 cap group) and therefore the manner in which the residues of the cap domain participate in the active site is different. In the following description, sequence motifs 3 and 4 found in *Pho*-MPGP according to Burroughs *et al.* (2006) correspond to sequence motif 3 reported by Wang *et al.* (2001, 2002).

Four sequence motifs in holo-form *Pho*MPGP and in the *MJ*PSP crystal structure with a phosphate ion (PDB code 1l7m) were compared (Fig. 6). In *Pho*MPGP, the Asp8–Thr12 sequence motif 1, corresponding to Asp11–Thr15 of *MJ*PSP, contains the Asp8 residue at the same position as the Asp11 residue in *MJ*PSP, which is phosphorylated during its reaction. In contrast, Asp10 in *Pho*MPGP has a different character from its counterpart in *MJ*PSP; Asp10 in *Pho*MPGP is distant from the phosphate ion and forms a salt bridge with the side chain of Arg170 in the bending loop. In contrast, Asp13 in *MJ*PSP is close to the phosphate ion and is a candidate for the general acid/base

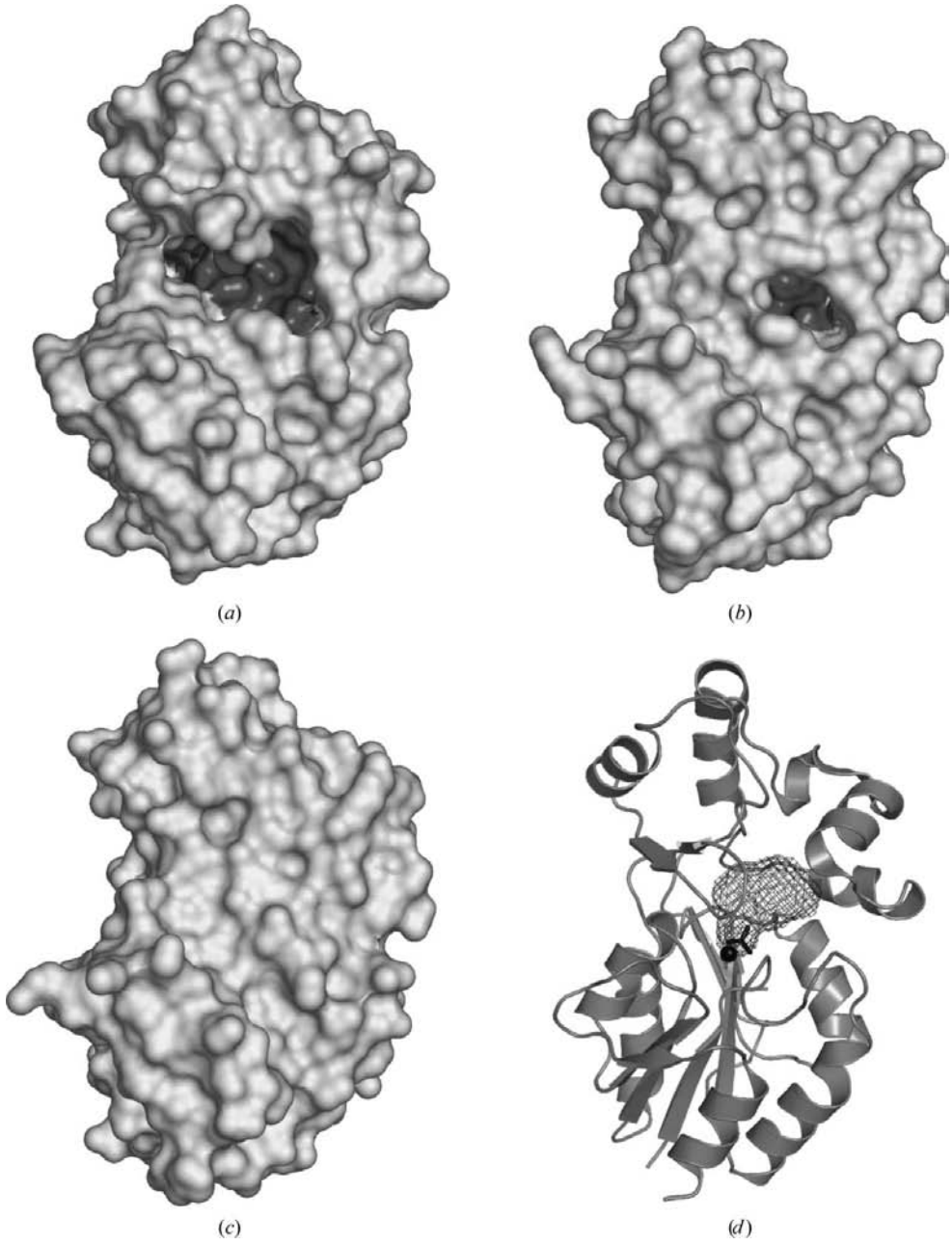

Figure 7
Molecular surfaces of (*a*) apo-form *Pho*MPGP open conformation, (*b*) apo-form *Pho*MPGP closed conformation and (*c*) holo-form *Pho*MPGP. The sequestered surface in (*c*) is indicated in dark grey in (*a*) and (*b*). (*d*) Ribbon model of holo-form *Pho*MPGP viewed from the same angle as in (*c*). In (*d*), the phosphate and magnesium ions are indicated by a stick model and as a sphere, respectively. The mesh indicates the cavity detected by the program *VOIDOO* (Kleywegt & Jones, 1994) with a probe 1.4 Å in radius. The cavity had an ovoid-like shape with one protrusion where the phosphate ion binds in the active site.

involved in its catalytic reaction. Lys180 in motif 3 may stabilize the phosphate ion in the reaction, as does Lys144 in *MJ*PSP. The distances between the O atom of the phosphate ion and the N atom of the lysine residue are 3.1 Å in *Pho*MPGP and 2.7 Å in *MJ*PSP. The D*XX*ND sequence in motif 4 corresponds to residues 167–171 in *MJ*PSP motif 3; Asp204 (Asp167 in *MJ*PSP) stabilizes the magnesium ion and Asp208 and Asn207 (Asp171 and Asn170 in *MJ*PSP) make hydrogen bonds to the water molecule w73 coordinated to the magnesium ion (corresponding to water molecule w222 in the report of Wang *et al.*, 2001). Ser40 in motif 2 (Ser99 in *MJ*PSP) forms a hydrogen bond to the phosphate ion. Ser41 in *Pho*MPGP is a characteristic residue that forms a hydrogen bond to the phosphate ion from the opposite side from where Asp11 is coordinated in *MJ*PSP. The phosphate ion sits on the active site in the same way as in the *MJ*PSP structure, except for Asp10, which forms a salt bridge with Arg170.

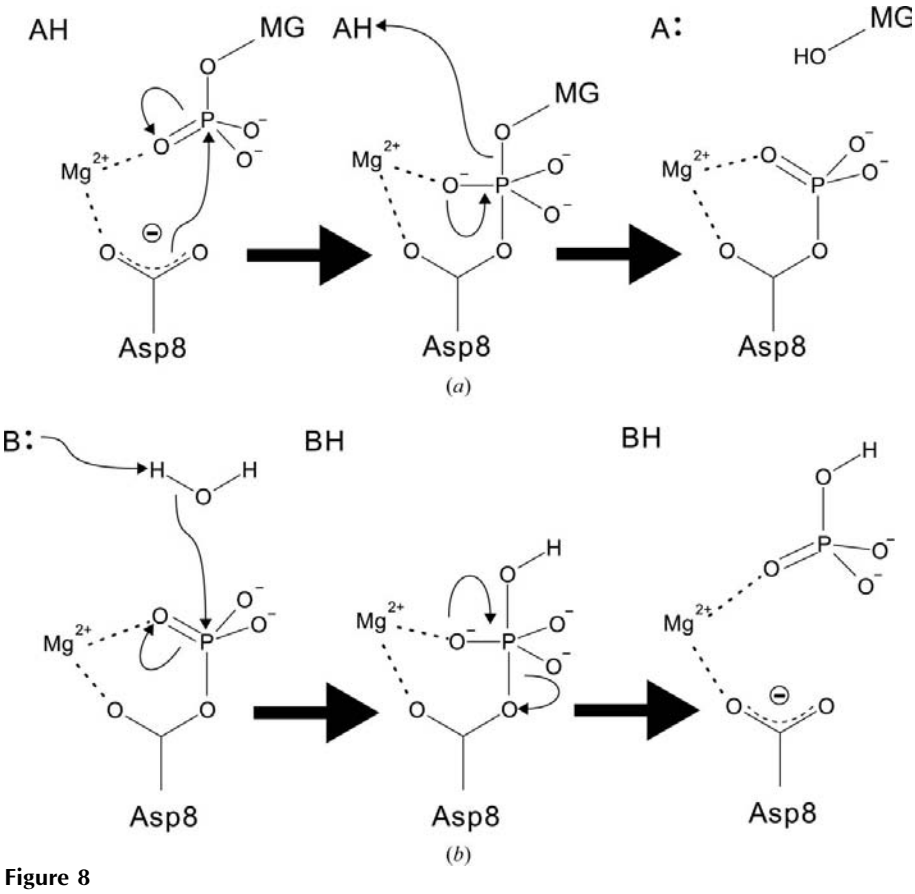

Figure 8
Catalytic mechanism of the first (*a*) and the second (*b*) reactions. In the first reaction, the phosphate group of MPG is transferred to the side chain of Asp8 by a nucleophilic reaction. A general acid (labelled A) provides a proton to the released MG. In the second reaction, the phosphate group of the side chain of Asp8 is transferred by a nucleophilic reaction to a water molecule activated by a general base (B).

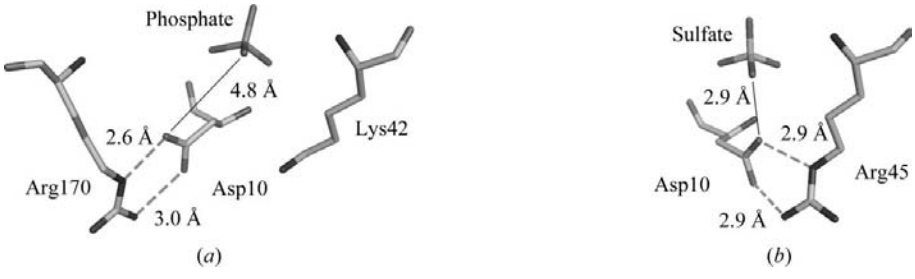

Figure 9
Salt bridge of the general acid/base of (*a*) *Pho*MPGP and (*b*) BT4131 from the C2B cap subclass (PDB code 1ymq). The side chains of the general acid/base (Asp10 in both) and its counterpart in the salt bridge (Arg170 and Arg45) are shown as stick models. The phosphate ion and the sulfate ion whose O atoms are hydrogen donors in the general base reaction are also shown. Lys42 of *Pho*MPGP, shown as a stick model, is the spatial counterpart of Arg45 of BT4131. Salt bridges are shown by dashed lines with their distances (Å). Distances between hydrogen donors and the general acid/base in the general base reaction are also shown.

The contribution of the cap domain differs between *Pho*MPGP and *MJ*PSP; the cap domain of *MJ*PSP is involved in substrate recognition of the carboxyl group of phospho-L-serine, while that of *Pho*MPGP has the bending loop bent toward the active site. Notably, the side chain of Ser169 is coordinated to the magnesium ion and the side chain of Arg170 forms a salt bridge with the side chain of Asp10 in sequence motif 1. Of these, Ser169 is less highly conserved in MPGPs from thermophiles and hyperthermophiles (data not shown). Arg170 is a more intriguing residue because it is conserved in MPGPs and a similar salt bridge to the general acid/base Asp is also observed in the crystal structures of the HADSF C2B and C0 cap group phosphatases (Lu *et al.*, 2005).

3.4. Cavity between the two domains

*Pho*MPGP has a cavity located between the core domain and the cap domain that includes the active site. In the open conformation of apo-form *Pho*MPGP this cavity is accessible from the outside (Fig. 7*a*), but it is less accessible in the closed conformation (Fig. 7*b*). In holo-form *Pho*MPGP the cavity is completely sequestered (Fig. 7*c*) and the phosphate ion is confined within it (Fig. 7*d*). The cavity is observed to have a hydrophilic nature and water molecules occupy the cavity in the holo-form and apo-form *Pho*MPGP crystal structures. The same cavity is observed in the crystal structure of the HADSF enzyme BT4131, in which the cavity is called the 'solvent cage' to indicate a potential substrate-binding site (Lu *et al.*, 2005).

In contrast to the open conformation of *Pho*MPGP, the closed conformations of apo-form *Pho*MPGP and holo-form *Pho*MPGP are similar in some respects:

these include stabilization of the bending loop as indicated by its lowered temperature factor, lower accessibility of the cavity and the domain conformation. Without the magnesium ion, *Pho*MPGP adopts a closed conformation similar to that of the holo-form *Pho*MPGP. The differences are that the holo-form *Pho*MPGP has an Asp10–Arg170 salt bridge and Ser169 is coordinated to the magnesium ion between the cap domain and the core domain; these are not observed in the closed conformation of apo-form *Pho*MPGP.

4. Discussion

4.1. Catalytic mechanism of dephosphorylation

The similarity in active-site architecture between the holo form of *Pho*MPGP and *MJ*PSP (Fig. 6) suggests that they share a similar catalytic mechanism of dephosphorylation. In analogy to the catalytic mechanism of *MJ*PSP (Wang *et al.*, 2002), the catalytic mechanism of dephosphorylation of *Pho*MPGP is proposed to take place as follows (Fig. 8). Firstly, the P atom of bound MPG is attacked by the carboxyl group of Asp8 and the phosphate group is released from MG. Secondly, the phosphate group is released from Asp8 by nucleophilic attack by the activated water molecule. In this series of reactions, the magnesium ion acts as the substrate activator and Asp10 is the candidate for the general acid/base.

The role of the salt bridge between Asp10 and Arg170 in the reaction mechanism is intriguing. Usually, the formation of a salt bridge by the general acid/base is considered to be beneficial for the general acid/base reaction by stabilizing the hydrogen bond to the reaction counterpart (Fig. 9*b*; Lu *et al.*, 2005). However, the Asp10–Arg170 salt bridge in *Pho*MPGP keeps Asp10 too distant (4.8 Å) to allow the formation of a hydrogen bond to the O atom of the phosphate ion (Fig. 9*a*). One suggested role of the Asp10–Arg170 salt bridge is inhibition of the reverse reaction by removing Asp10 from the stabilized architecture which would be accomplished by a salt bridge to Lys42 (Asp10–Lys42). In the closed conformation, where the product is immobilized, this role is important. Thus, the change between the open and closed conformations is important in two respects: accessibility of the active site for substrate binding and product release and inhibition of reverse reaction.

4.2. Substrate binding and conformational change

The cavity between the two domains is the substrate-binding site in which the MPG molecule is accommodated such that the phosphate group is bound to the active site as in the holo form. The size of the cavity is 269 Å^3, which is slightly larger than the volume of the substrate MPG (253 Å^3). In the same manner as HADSF enzymes such as *MJ*PSP and sucrose-6-phosphate phosphatase (Fieulaine *et al.*, 2005), the catalytic reaction of *Pho*MPGP is considered to proceed in the closed conformation when the substrate binds. When the substrate MPG is bound or the product MG and phosphate ion are released, *Pho*MPGP must be in the open conformation. We found that both apo forms and holo forms of *Pho*MPGP can adopt the closed conformation, in which the bending loop of the cap domain containing Arg170 is moved to cover the cavity. This is beneficial for the catalytic reaction because the substrate is firmly immobilized. Indeed, half of the activity remains without the magnesium ion, as reported by Empadinhas *et al.* (2001).

We thank N. Igarashi, N. Matsugaki and Y. Yamada of KEK-PF for their assistance in our synchrotron experiments. This work was supported by a grant for the National Project on Protein Structural and Functional Analyses from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

Bond, C. S. (2003). *Bioinformatics*, **19**, 311–312.
Borges, N., Marugg, J. D., Empadinhas, N., da Costa, M. S. & Santos, H. (2004). *J. Biol. Chem.* **279**, 9892–9898.
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.
Burroughs, A. M., Allen, K. N., Dunaway-Mariano, D. & Aravind, L. (2006). *J. Mol. Biol.* **361**, 1003–1034.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cowtan, K. (1994). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **31**, 34–38.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* DeLano Scientific, San Carlos, USA. http://www.pymol.org.
Empadinhas, N., Albuquerque, L., Costa, J., Zinder, S. H., Santos, M. A. S., Santos, H. & da Costa, M. S. (2004). *J. Bacteriol.* **186**, 4075–4084.
Empadinhas, N., Albuquerque, L., Henne, A., Santos, H. & da Costa, M. S. (2003). *Appl. Environ. Microbiol.* **69**, 3272–3279.
Empadinhas, N. & da Costa, M. S. (2006). *Int. Microbiol.* **9**, 199–206.
Empadinhas, N., Marugg, J. D., Borges, N., Santos, H. & da Costa, M. S. (2001). *J. Biol. Chem.* **276**, 43580–43588.
Faria, T. Q., Knapp, S., Ladenstein, R., Maçanita, A. L. & Santos, H. (2003). *Chembiochem*, **4**, 734–741.
Fieulaine, S., Lunn, J. E., Borel, F. & Ferrer, J.-L. (2005). *Plant Cell*, **17**, 2049–2058.
French, S. & Wilson, K. (1978). *Acta Cryst.* A**34**, 517–525.
Hayward, S. & Berendsen, H. J. C. (1998). *Proteins*, **30**, 144–154.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kabsch, W. (1976). *Acta Cryst.* A**32**, 922–923.
Kleywegt, G. J. & Jones, T. A. (1994). *Acta Cryst.* D**50**, 178–185.
La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lu, Z., Dunaway-Mariano, D. & Allen, K. N. (2005). *Biochemistry*, **44**, 8684–8696.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Perrakis, A., Sixma, T. K., Wilson, K. S. & Lamzin, V. S. (1997). *Acta Cryst.* D**53**, 448–455.
Sheldrick, G. M., Hauptman, H. A., Weeks, C. M., Miller, M. & Usón, I. (2001). In *International Tables for Crystallography*, Vol. F, edited by M. G. Rossmann & E. Arnold, pp. 333–351. Dordrecht: Kluwer Academic Publishers.

Terwilliger, T. C. (2000). *Acta Cryst.* D**56**, 965–972.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Wang, W., Cho, H. S., Kim, R., Jancarik, J., Yokota, H., Nguyen, H. H., Grigoriev, I. V., Wemmer, D. E. & Kim, S.-H. (2002). *J. Mol. Biol.* **319**, 421–431.
Wang, W., Kim, R., Jancarik, J., Yokota, H. & Kim, S.-H. (2001). *Structure*, **9**, 65–71.