Structure
Article

Insights into the Catalytic Mechanism of PPM Ser/Thr Phosphatases from the Atomic Resolution Structures of a Mycobacterial Enzyme

Marco Bellinzoni,[1,3] Annemarie Wehenkel,[1,3] William Shepard,[2] and Pedro M. Alzari[1,*]
[1] Unité de Biochimie Structurale, CNRS-URA 2185, Institut Pasteur, 75724 Paris Cedex 15, France
[2] Synchrotron Soleil, L'Orme des Merisiers, Saint Aubin, 91192 Gif-sur-Yvette, France
[3] These authors contributed equally to this work.
*Correspondence: alzari@pasteur.fr
DOI 10.1016/j.str.2007.06.002

SUMMARY

Serine/threonine-specific phosphatases (PPs) represent, after protein tyrosine phosphatases, the second major class of enzymes that catalyze the dephosphorylation of proteins. They are classed in two large families, known as PPP and PPM, on the basis of sequence similarities, metal ion dependence, and inhibitor sensitivity. Despite their wide species distribution and broad physiological roles, the catalytic mechanism of PPM phosphatases has been primarily inferred from studies of a single enzyme, human PP2Cα. Here, we report the biochemical characterization and the atomic resolution structures of a soluble PPM phosphatase from the saprophyte *Mycobacterium smegmatis* in complex with different ligands. The structures provide putative snapshots along the catalytic cycle, which support an associative reaction mechanism that differs in some important aspects from the currently accepted model and reinforces the hypothesis of convergent evolution in PPs.

INTRODUCTION

Phospho-Ser/Thr protein phosphatases (PPs) are dinuclear metalloenzymes that remove phosphate from serine or threonine residues, which accounts for over 98% of reversibly protein-bound phosphate in eukaryotic cells (Olsen et al., 2006). PPs are divided into two large gene families, PPP and PPM, which can be further categorized in subfamilies based upon regulatory and targeting domains that are associated with the catalytic domain, their sensitivity to a variety of different inhibitors, distinct metal requirements, and genetic homology (Barford et al., 1998; Cohen, 1989; Jackson and Denu, 2001). Members of the PPP family are normally composed of a catalytic subunit in association with a regulatory subunit or domain. Human PP1, PP2A, PP2B, and PP5 are the most representative eukaryotic members of this family (Gallego and Virshup, 2005; Rusnak and Mertz, 2000). The PPM family was identified on the basis of the strict requirement for an exogenous divalent ion (Mg^{2+}, Mn^{2+}) for activity, as well as for the insensitivity to known PPP inhibitors such as okadaic acid. PPP and PPM phosphatases display unrelated amino acid sequences, although the structure of the alpha isoform of human PP2C (PP2Cα), which is considered the defining member of the PPM family, displays some similarities with PPP enzymes in their overall fold and dinuclear metal center (Das et al., 1996; Jackson and Denu, 2001).

Long thought to be restricted to eukaryotes, PPM phosphatases are also widely distributed in eubacterial and archaeal genomes (Kennelly, 2002, 2003), where their physiological roles are only now starting to be unveiled. The crystal structure of the catalytic domain from membrane-associated *Mycobacterium tuberculosis* PstP confirmed the overall resemblance to PP2Cα in protein architecture and catalytic machinery (Pullen et al., 2004), although it also showed the presence of a third metal ion in the active center, whose functional role remains to be determined.

Despite their wide species distribution and diverse physiological roles, relatively little information is available at the molecular level on the catalytic mechanism of PPM phosphatases, and most of these data come indeed from structural and enzymological studies of a single protein, human PP2Cα (Das et al., 1996; Fjeld and Denu, 1999; Jackson and Denu, 2001; Jackson et al., 2003). In the present study, we report the biochemical characterization and the crystal structure at atomic resolution of MspP, a soluble PPM phosphatase identified in the saprophyte *M. smegmatis*. The protein structure has been determined in complex with a cacodylate ion (1.4 Å resolution), a sulfate ion (1.1 Å), and inorganic phosphate (0.83 Å) in different crystal forms. The atomic models of MspP suggest a reaction mechanism that differs in some important aspects from that proposed for PP2Cα and provide novel structural insights into the mode of action of this large family of enzymes.

Table 1. Kinetic Parameters of MspP Using *p*NPP as Substrate

Source of Metal	k_{cat} (s^{-1})	K_m (*p*NPP) (mM)	K_m (Metal) (mM)	k_{cat}/K_m (*p*NPP) (M^{-1} s^{-1})	k_{cat}/K_m (Metal) (M^{-1} s^{-1})
$MnCl_2$	4.19 ± 0.15	0.97 ± 0.04	0.81 ± 0.02	4,310 ± 94	4,995 ± 239
$MgCl_2$	0.125 ± 0.006	9.91 ± 0.63	10.62 ± 0.72	12.68 ± 0.26	11.97 ± 1.36
Specific activity of MspP as a function of pH					
pH	5.5	6	6.5	7	7.5
Specific activity	0.05 ± 0.006	0.11 ± 0.005	0.61 ± 0.01	2.02 ± 0.04	3.33 ± 0.1

The specific activity is defined as μM of product (*p*NP) produced per min and per mg of enzyme.

RESULTS AND DISCUSSION

A New Bacterial Member of the PPM Family of Phosphatases

A prokaryotic genomic survey revealed a large number of genes coding for PP2C-like phosphatases that may exist in long (putative transmembrane) and short (soluble) forms. Putative transmembrane phosphatases are almost exclusively found in Actinobacteria and Cyanobacteria, while short soluble phosphatases appear to be distributed more equally over all bacterial genomes. We identified an open reading frame coding for a soluble PP2C-like protein phosphatase, here named MspP, in the genome of *Mycobacterium smegmatis* mc^2155. MspP is a polypeptide of 233 residues that includes all conserved motifs and critical amino acids identified in PPM phosphatases (Bork et al., 1996). It shows low but significant (15%–35%) sequence similarities with other PPM enzymes, including the two proteins for which the structure is currently available, e.g., human PP2Cα (17% sequence identity) (Das et al., 1996) and the catalytic domain of *M. tuberculosis* PstP (*Mt*PstP, 37%) (Boitel et al., 2003; Pullen et al., 2004).

The gene coding for MspP was cloned, and the recombinant protein expressed in *Escherichia coli* for subsequent biochemical and structural studies. MspP is active against the surrogate substrate *p*NPP in the presence of $MnCl_2$ (Table 1). The specific activity (k_{cat}/K_m = 4310 $M^{-1}s^{-1}$) is comparable to that measured for human PP2Cα (Jackson et al., 2003) (k_{cat}/K_m = 1100 $M^{-1}s^{-1}$) but significantly higher than that observed for the catalytic domain of *Mt*PstP (Pullen et al., 2004) (k_{cat}/K_m = 60 $M^{-1}s^{-1}$). As shown in Table 1, MspP also behaves like PP2Cα (Fjeld and Denu, 1999) in that the specific activity decreases over 300-fold when $MnCl_2$ is replaced by $MgCl_2$ in the reaction buffer, indicating a strong preference for Mn^{2+} over Mg^{2+} ions, and the pH dependence of k_{cat}/K_m for the substrate *p*NPP revealed one critical ionizable group at pH ~7, which must be unprotonated for catalysis and could be assigned to a water nucleophile (Chen et al., 1997; Fjeld and Denu, 1999; Pohjanjoki et al., 1998).

Overall Structure

We have determined the atomic resolution structures of MspP in complex with phosphate (0.83 Å), sulfate (1.1 Å), and cacodylate (1.4 Å) from different crystal forms by molecular-replacement methods (Table 2). The final atomic models comprise the entire polypeptide chain, including an additional N-terminal glycine residue, left over upon cleavage of the affinity tag. The overall fold and topology of the enzyme closely resemble those observed for *Mt*PstP (Pullen et al., 2004) (Figure 1A) and human PP2Cα (Das et al., 1996) (Figure 1B). The common core is made up of two five-stranded antiparallel β sheets, on top of which the active site is located, flanked by a pair of antiparallel α helices on either side. Within the conserved core, the highest structural variability is observed for some connections between secondary structure elements as well as in the flap segment (MspP residues 130–160), immediately adjacent to the active site cleft (Figure 1C). The MspP flap is very conserved among prokaryotic PP2C-like phosphatases. It has the same length (50% sequence identity) and adopts the same fold as in *Mt*PstP (Figure 1A), suggesting that it is a structural element proper to bacterial enzymes. Furthermore, in 107 out of 114 bacterial sequences (94%), the flap segment differs in length by four or less residues over a sequence region (50 residues) delimited by conserved secondary structural elements. In contrast, the flap appears to be more variable in length and structure in eukaryotic enzymes, as illustrated by the structure of human PP2Cα (Figure 1B).

The Active Site

The canonical dinuclear metal center of PPM phosphatases is strictly conserved in MspP. Two octahedrally coordinated ions (M1, M2) separated by 3.8 Å are coordinated by the conserved aspartic acid residues Asp35, Asp185, and Asp223, the main-chain carbonyl of Gly36 and water molecules, one of which bridges the two-metal center (Figure 2). However, the structure of cacodylate-bound MspP revealed a third metal site (M3) close to the catalytic center (Figure 2A). The analysis of X-ray anomalous diffraction data just above and below the Mn K edge confirmed that all three metal sites are occupied by manganese (see Figure S1A in the Supplemental Data available with this article online). The Mn^{2+} at M3 is octahedrally coordinated and occupies the equivalent position as in *Mt*PstP, except that the imidazole ring of His153 replaces the equivalent residue Ser160 in *Mt*PstP (Figure 1A). While the direct interaction of M3 with phosphate in the active site (Figures 1C and 2) points to a role in substrate recognition, the precise function of the third Mn^{2+}

Table 2. Diffraction Data Collection and Refinement Statistics

Data Collection	Cacodylate Complex	Phosphate Complex	Sulfate Complex
Beamline	ESRF ID14-2	EMBL/DESY BW7A	ESRF ID14-4
Wavelength (Å)	0.933	0.805	0.939
Space group	$P2_12_12_1$	$P2_12_12$	$P2_12_12$
Cell dimensions a, b, c (Å)	36.568, 65.786, 110.059	76.382, 83.998, 33.478	76.547, 84.402, 33.609
Resolution[a] (Å)	30.0–1.45 (1.53–1.45)	10.0–0.83 (0.86–0.83)	20.0–1.08 (1.15–1.08)
Unique reflections	47,486	189,539	93,536
I/σ (I)[a]	15.6 (2.6)	15.8 (2.7)	9.9 (2.9)
Multiplicity[a]	3.8 (3.7)	3.9 (3.7)	3.8 (3.0)
Completeness[a] (%)	99.0 (97.7)	93.2 (88.0)	99.3 (99.1)
R_{merge}[a,b]	0.053 (0.434)	0.022 (0.477)	0.054 (0.380)
Refinement			
Used reflections[c]	45,070	179,726	88,321
R factor[d]	0.178	0.095	0.116
Free R factor[d]	0.191	0.124	0.160
Refined atoms			
Protein non-H atoms[e]	1,790	1,907	1,835
Water molecules	242	493	415
Rmsds			
Bond lengths (Å)	0.013	0.017	0.014
Bond angles (°)	1.459	1.996	2.076

[a] Numbers in parentheses correspond to the highest resolution shell.
[b] $R_{merge} = \sum_j |\hat{I}_h - I_{h,j}|/\sum_{h,j} I_{h,j}$ where $\hat{I}_h = (\sum_j I_{h,j})/n_h$ and n_h is the multiplicity of reflection h.
[c] This number does not include the free R set of reflections (5% of total reflections).
[d] R factor = $\sum_{hkl} ||F_o| \times k|F_c||/\sum_{hkl} |F_o|$; R_{free}: same for the test set (5% of the data).
[e] Including side chains refined in multiple conformations.

site in mycobacterial phosphatases remains to be established.

In the two other structures of MspP (phosphate- and sulfate-bound forms), we observed some heavy atom sites that were not consistent with Mn^{2+} as judged from anomalous diffraction data (Figures S1B and S1C). These positions were modeled as Mg^{2+} from the crystallization solutions. Thus, the final model of phosphate-bound MspP (crystallized in the presence of 250 mM $MgCl_2$) contains one Mn^{2+} (M1) and three Mg^{2+} (M2–M4) ions, while the sulfate-bound enzyme (crystallized in 200 mM $MgSO_4$) has one Mn^{2+} (M1) and four Mg^{2+} (M2–M5) ions. Traces of Mn^{2+} at the M2 site could still be detected in the phosphate-bound MspP structure from anomalous diffraction data, but it was not included in the final model. Despite the partial metal substitution, the catalytic binuclear center (M1, M2) is perfectly superposable and displays the same coordination geometry in all three structures (Figure 2). In contrast, the additional sites M3–M5 have relatively weak Mg^{2+} occupancies (15%–40%) and probably arise from the high concentration of magnesium ions present in the crystallization buffers. It should be noted that the M3 site occupied by Mg^{2+} in these two structures differ from the third Mn^{2+} site in the cacodylate-bound structure, mainly because His153 displays a different conformation, pointing outside toward the open solvent, and a water molecule now coordinates the Mg^{2+} ion (Figure 2B).

Different Positions of Ligands within the MspP Active Site

The crystal form of MspP preincubated with 10 mM inorganic phosphate at pH 8.5 showed strong positive peaks for a tetrahedral arrangement of atoms forming a tripodal complex with the dinuclear metal center (Figure 2B). The enzyme preincubated with $MnCl_2$ was found to bind phosphate with an affinity of ~80 μM by using isothermal titration calorimetry (data not shown), suggesting that the tetrahedral arrangement should correspond to a bound phosphate molecule. However, upon unrestrained refinement at 0.83 Å, the central phosphorous atom was found to be partially occupied (~20%), whereas the four neighboring oxygens retained full occupancy. Nevertheless, the presence of the ligand at this site is strongly supported by the observation that total occupancies refined to nonphysical values significantly above unity (1.5) when the five positions were assigned to an alternate arrangement of water molecules. Therefore, the phosphate ion was

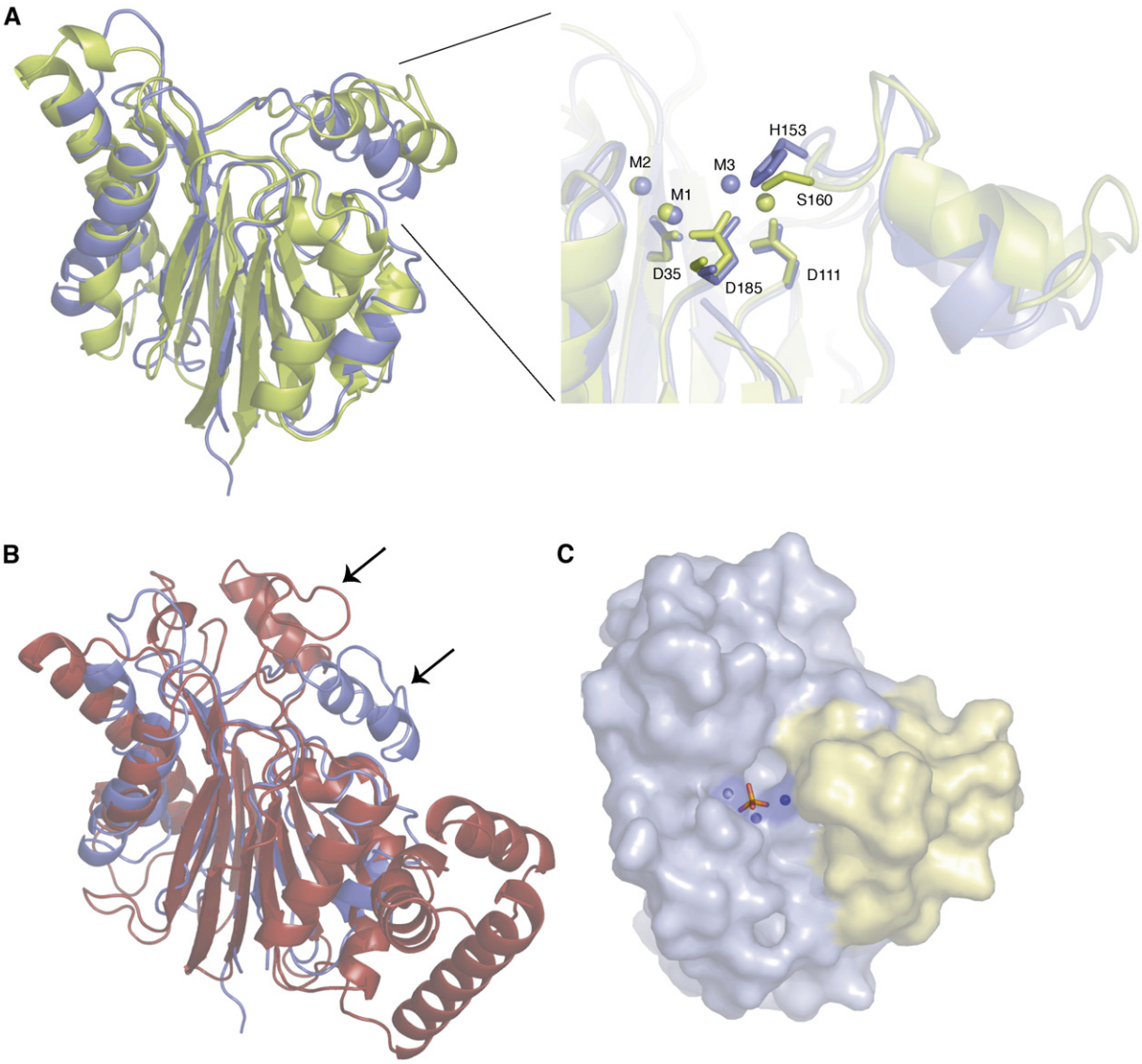

Figure 1. Overall Structure of MspP

(A) Superposition of MspP (in blue) and *Mt*PstP (PDB code: 1TXO, in green). The rmsd is 1.76 Å for 212 core residues. The inset shows a close view of the catalytic centers, with the three Mn^{2+} ions and some coordinating residues labeled.

(B) Superposition of MspP (in blue) and PP2Cα (PDB code: 1A6Q, in red), with a rmsd of 2.37 Å for 189 core residues. Note the different structure and orientation of the flap segment (shown by arrows).

(C) General view of the molecular surface of MspP in complex with phosphate. The flap segment is shown in yellow.

finally modeled in mutually exclusive occupation with four water molecules. Two phosphate oxygens bind respectively to M1 and M2, while a third oxygen atom, which points towards the bulk solvent, forms a hydrogen bond with the Arg17 guanidinium group. The fourth phosphate oxygen bridges the two metal ions and occupies the position of the water nucleophile (Das et al., 1996), suggesting that the structure of phosphate-bound MspP represents the enzyme-product complex.

MspP binds cacodylate (whose identification was confirmed by X-ray data collected about the As K edge; Figure S1A) close to the dinuclear metal center. The spatial positions of the two cacodylate oxygens (Figure 2A) match those of two phosphate oxygens in the previous complex (Figure 2B). The catalytic water bridging the two Mn^{2+} ions is now detached from the cacodylate group but remains close enough (3.6 Å) for nucleophilic attack, suggesting that this complex partially mimics the binding of the phospho-substrate. A water molecule that is hydrogen bonded to Arg17 occupies the site of the fourth phosphate oxygen, likely because this position is energetically unfavorable for the cacodylate methyl group. A small rotation of the ligand to match this water position would therefore bring about the putative position of the phosphate

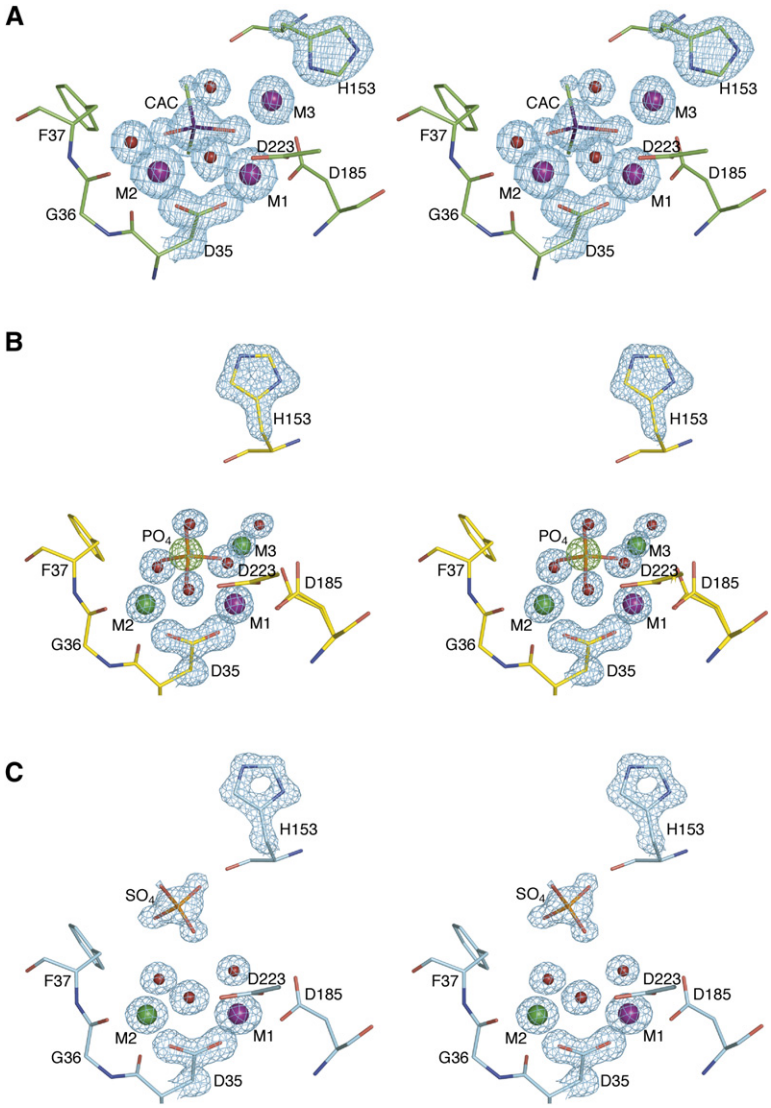

Figure 2. Final Electron Density Maps Contoured at 2 σ Level of the Ligands Bound to MspP

(A) MspP-cacodylate complex at 1.4 Å resolution.

(B) MspP-phosphate complex at 0.83 Å resolution. The (F_o-F_c) difference map (countoured at 5 σ level), calculated before the phosphate ion was added to the model, is shown in green.

(C) MspP-sulfate complex at 1.08 Å resolution. In all cases, Mn^{2+} ions are shown as large violet spheres, Mg^{2+} ions as large green spheres, and water molecules as small red spheres.

group in the enzyme-phospho-substrate complex (see below). The third crystal form of MspP, which was obtained in the absence of phosphate or cacodylate, revealed a tetrahedral molecule close to the active site that was identified as a sulfate ion from the crystallization buffer. The ligand refined to nearly full occupancy (0.9) and is stabilized by hydrogen bonding interactions with the guanidinium group of Arg17 and three water molecules, including the catalytic nucleophile (Figure 2C).

Structural Insights into the Catalytic Mechanism of PPM Phosphatases

The atomic models of MspP provide strong arguments to support a general acid-base mechanism of catalysis, in which an activated water molecule in form of an hydroxide anion serves as a nucleophile to attack the phosphorous atom in a S_N2-like mechanism (Figure 3). As observed for PP2Cα (Fjeld and Denu, 1999) and other dinuclear metallohydrolases (Chen et al., 1997; Pohjanjoki et al., 1998), the pH-dependent activity of MspP revealed an ionizable group with a pK_a of approximately 7, which might correspond to the activation of the water molecule bridging the dinuclear metal center. The deprotonation of this water molecule is further assisted by Asp223 (Figure 3B), in agreement with structural and biochemical data for the equivalent residue Asp282 in PP2Cα (Jackson et al., 2003). The role of this metal-bridging water as the catalytic nucleophile in MspP is further confirmed by the 0.83 Å resolution structure of the phosphate-enzyme complex, in which the solvent site is partially occupied by a phosphate oxygen. Another critical residue is Arg17, which plays an important role in substrate binding and transition state stabilization, since its guanidinium group is seen to interact with the incoming phosphate in all three MspP structures. This again agrees with biochemical data from PP2Cα showing that a mutant lacking the equivalent

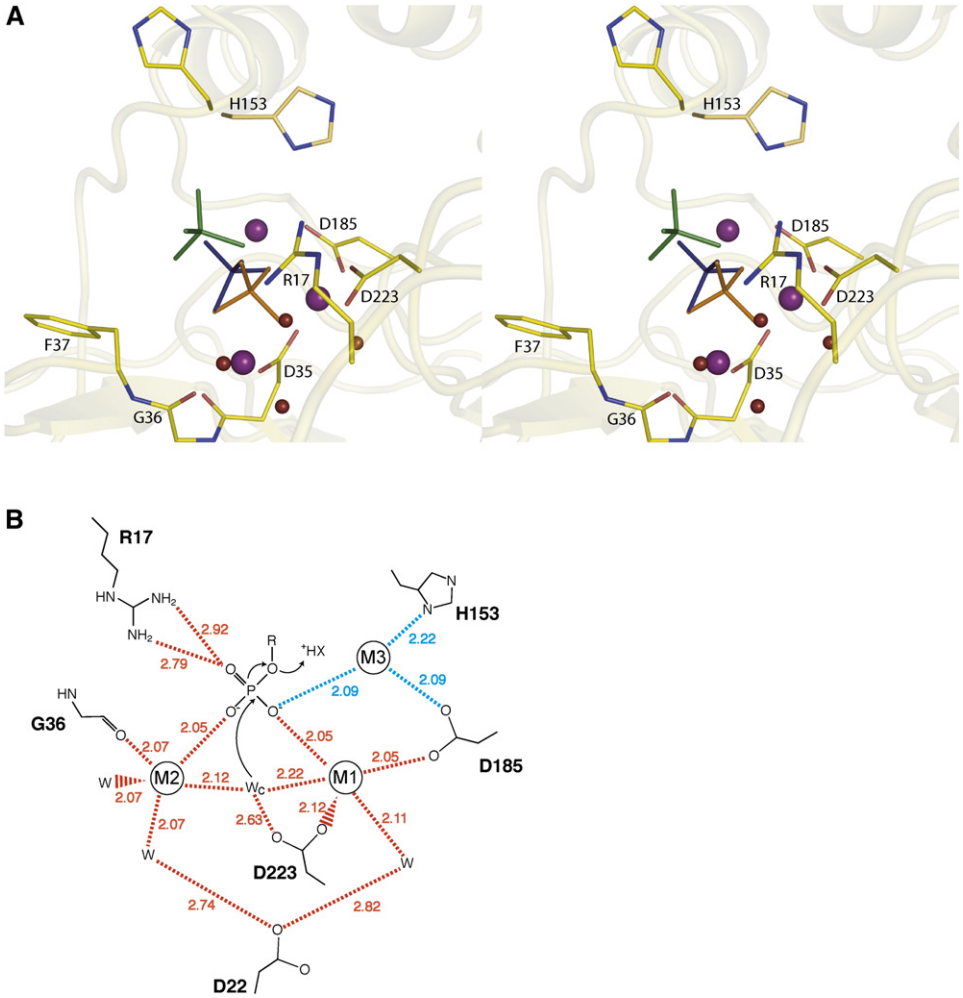

Figure 3. The Catalytic Mechanism of MspP

(A) Stereoview of the active site showing the trinuclear metal center and selected amino acid residues (His153 is shown in two conformations as observed in the structures of the enzyme-cacodylate and enzyme-phosphate complexes, respectively). The three positions of phosphate, as deduced from the corresponding crystal structures (see text), are shown in different colors: competent substrate (blue), reaction product (orange), and incoming substrate or outgoing product (green).

(B) Schematic view illustrating the proposed catalytic mechanism. Interatomic distances shown in red correspond to the MspP-phosphate complex at 0.83 Å resolution, those in blue to the MspP-cacodylate complex at 1.4 Å resolution. The octahedral coordination of metals M1 and M2 is completed by the carboxylate group of Asp 35 (not shown for clarity), with distances M1 – $O_{\delta 2}$ = 2.10 Å and M2 – $O_{\delta 1}$ = 2.00 Å. Water molecules (not shown) complete the octahedral coordination at site M3. The putative proton donor (XH^+) may correspond to either a water molecule or a substrate group.

residue, Arg33, displays a 7-fold higher K_m value for the surrogate substrate *p*NPP and a 2-fold reduction in k_{cat} (Jackson et al., 2003).

The proposed mechanism for human PP2Cα suggests that the phosphate group from the substrate binds the dinuclear metal center through water-mediated interactions (Das et al., 1996). Since the phosphate-binding site observed in PP2Cα is closely equivalent to the sulfate-binding site in the MspP-sulfate complex (Figure 4), the latter might represent a model of the competent enzyme-substrate (Michaelis) complex for MspP. If this were the case, the different positions occupied by the bound ligand (phosphate or cacodylate) in the two other MspP structures may have no functional relevance and be a consequence of the different crystallization conditions (pH, metal content). A possible argument against this interpretation is that the tripodal complex between inorganic phosphate and the dinuclear metal center in the MspP-phosphate complex is very similar to that previously described for other metallohydrolases, such as *Bacillus pasteurii* urease (Benini et al., 2001) and purple acid phosphatase from sweet potato (Schenk et al., 2005), as well as for sulfate bound to the bacteriophage lambda PPP phosphatase (Voegtli et al., 2000).

An alternative scenario, more consistent with our structural data, is one in which the phosphate group from the substrate directly interacts with the metal ions in a bidentate mode during catalysis, in the same way as

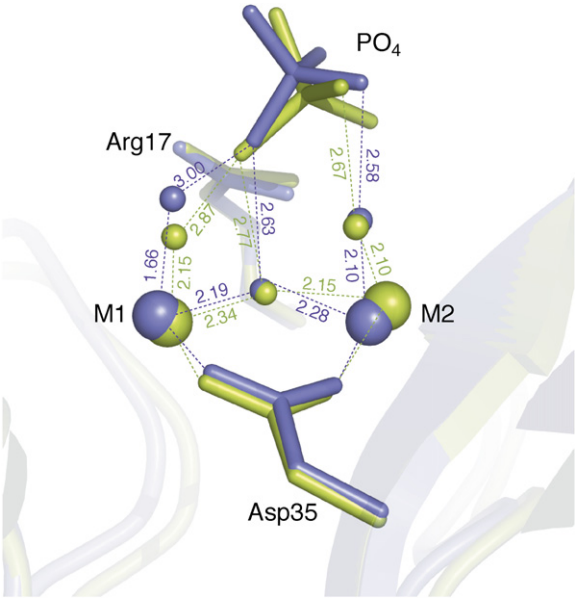

Figure 4. Structural Superposition of Phosphate Bound to PP2Cα, in Blue, and Sulfate Bound to MspP, in Green
Distances involving the water molecules that mediate ligand-metal interactions (represented as small spheres) are indicated in both cases. MspP residues Asp 35 and Arg 17 correspond to PP2Cα residues Asp 60 and Arg 33, respectively.

documented for PPP phosphatases (Egloff et al., 1995; Swingle et al., 2004). According to this interpretation, the MspP-cacodylate and MspP-phosphate complexes (in which two metal-bound oxygens from each ligand occupy identical positions) would mimic the competent enzyme-substrate and enzyme-product complexes, respectively, as illustrated in Figure 3A. The sulfate position could then correspond to that of the incoming phospho-protein substrate (or the outgoing inorganic phosphate), rather than a catalytically competent state. This hypothesis is reinforced by the unfavorable position and orientation of sulfate for nucleophilic attack by the catalytic water, which is at a distance of 4.1 Å from the sulfur atom (Figure 2C). It further suggests that the PP2Cα-phosphate complex (structurally equivalent to the MspP-sulfate complex, see Figure 4) might not represent a functionally competent state, as first noted by Fjeld and Denu (1999). Indeed, the PP2Cα crystals were obtained at pH 5.0 (Das et al., 1996), a value for which the enzyme is improperly protonated and has only 0.1% of the optimal activity (Fjeld and Denu, 1999; Jackson et al., 2003). It is therefore plausible that, in the PP2Cα crystals at acidic pH, the monoanionic state of phosphate could have favored a water-mediated interaction, rather than a direct association, with the positively charged metal ions.

A direct interaction between phosphate and metals would have important functional implications. The two Mn^{2+} ions may serve as Lewis acids to activate the catalytic water and to increase the electrophilicity of the phosphorous atom. Furthermore, the direct phosphate-metal interactions support an associative character of MspP catalysis (Cleland and Hengge, 2006). As inferred from the MspP-cacodylate and MspP-phosphate structures, the two phosphate groups (from the substrate and product of the reaction) would be positioned with the oxygens corresponding to the nucleophile and the leaving group in the apical positions with respect to the plane defined by the three equatorial oxygens of the putative phosphorane intermediate (Figure 3A). Furthermore, the two Mn^{2+} and the guanidinium group of Arg17 would stabilize the three equatorial oxyanions from the pentacoordinated transition state intermediate to a larger extent than the ground state phosphate with only two oxyanions (the opposite would be the case for a dissociative mechanism since the metaphosphate intermediate has a single negative charge). This is also consistent with previous observations about the catalytic activity of dinuclear Co(III) complexes, which point to an associative character for the hydrolysis of phosphate monoesters directly bound to a dinuclear metal center (Williams, 2004).

The Nature of the Proton Donor

Dephosphorylation of physiological substrates usually produces poor leaving groups, due to the high pK_a of the conjugated acids (e.g., ~13.6 for serine and threonine; Hengge, 2001), and a water molecule or protein residue is therefore expected to act as a general acid in the reaction (Mertz et al., 1997). His62 was proposed to fulfill this role in human PP2Cα (Jackson et al., 2003), but this residue is not conserved in MspP (where it is replaced by phenylalanine) nor in other bacterial PPs. A close inspection of the MspP structure suggests that His153 would be the only residue favorably positioned with respect to the bound phosphate to donate a proton in the reaction. We produced a His153-Ala mutant (H153A) and found that the enzymatic activity of this mutant on *p*NPP (k_{cat}/K_m = 3307 ± 154 $M^{-1}s^{-1}$) was only marginally affected compared to the wild-type enzyme (Table 1), strongly arguing against a catalytic role of His153 as a general acid in the reaction. Thus, unless MspP undergoes significant conformational changes upon substrate binding, it appears that a water molecule or a functional group from the protein substrate should fulfill the role of general acid in the reaction mechanism.

Concluding Remarks

The proposed mechanism of MspP catalysis (Figure 3), involving the direct binding of the phosphate group to the dinuclear metal center, closely resembles that described for the PPP family of phosphatases (Egloff et al., 1995; Jackson and Denu, 2001), but differs from the proposed model for human PP2Cα, in which the interaction between the phosphate group and the catalytic metals is mediated by water molecules. Although further experimental work is still required to elucidate the mechanistic similarities and differences between bacterial and eukaryotic PPM phosphatases, our results support a common associative catalytic mechanism and are suggestive of convergent evolution with PPP phosphatases, despite a very different

environment surrounding the active site in these two families of phospho-Ser/Thr phosphatases (Barford, 1996). The observed structural and mechanistic variations in PPM phosphatases, such as the different nature of the proton donor, the presence of additional metal-binding sites or the variations in the flap segment, may reflect the functional promiscuity of these enzymes, in which each member of the family is able to deal with a large variety of protein substrates.

EXPERIMENTAL PROCEDURES

Cloning of *M. smegmatis* MspP

The *M. smegmatis* gene MSMEG_1928 coding for MspP was amplified by PCR with the following primers, which introduce an EcoRI site at both ends of the gene: 5′-TTTTAGAATTCGAGAATCTTTATTTTCAGGGC ATGGCATCGGTGTTGAGT-3′ (forward), also carrying a 24 nt sequence (underlined) that codes the optimal tobacco etch virus protease (TEV) cutting site Glu-Asn-Leu-Tyr-Phe-Gln-Gly, and 5′-TTTTA GAATTCAGCCGAGGTCGATGAC-3′ (reverse). The amplification of the gene was performed on *M. smegmatis* genomic DNA with 0.5 U Pfx polymerase (Invitrogen), with the following PCR temperature profile: 30 s at 95°C, 30 s at 55°C, 1 min 30 s at 68°C for 30 cycles. The 0.7 kb reaction product was purified with Qiaquick spin columns (QIAGEN), digested with EcoRI, and ligated into the corresponding site of the pET-28a vector (Novagen), leading to the expression plasmid pM144. The point mutant His153-Ala was produced with the QuikChange Site-Directed mutagenesis kit (Stratagene), according to the manufacturer's instructions, leading to the plasmid pM170. The correct sequence of all constructs was verified by DNA sequencing on both strands.

Protein Expression and Purification

The same protocol was followed for the expression and purification of wild-type MspP and the H153A mutant. *E. coli* BL21(DE3) cells (Novagen) were transformed with pM144 or pM170 and grown on Luria Broth (LB) medium, supplemented with kanamycin (50 $\mu g\ ml^{-1}$). Cultures for protein purification were grown at 20°C to the optical density (OD_{600}) of 0.8, and the expression of the MspP was induced by the addition of 0.5 mM IPTG (isopropyl-β-thio-galactopyranoside); growth was then continued for further 15 hr at 20°C. Bacteria were harvested by centrifugation at 5,000 × g for 20 min, washed with PBS (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 [pH 7.3]), and stored at −80°C.

Cells were resuspended in lysis buffer (25 mM HEPES [pH 8.0], 500 mM NaCl, 25 mM imidazole, 5% glycerol) and lysed by sonication on ice. The lysate obtained from 2 l of bacterial culture was centrifuged at 26,800 × g for 1 hr, filtered (0.45 μm), and loaded onto a 5 ml HisTrap Ni^{2+}-IMAC column (GE Healthcare). Recombinant MspP was eluted applying a 25–400 mM imidazole gradient in the same buffer; the fractions containing the recombinant protein, as confirmed by SDS-PAGE, were pooled and dyalised overnight at 4°C against 25 mM HEPES (pH 7.5), 500 mM NaCl, 5% glycerol. To cleave the affinity tag, recombinant TEV protease was prepared as described (Lucast et al., 2001), added to a final mass ratio of 1:35, and the reaction mixture was left at 18°C for 48 hr in the presence of 1 mM DTT. The digestion mixture was passed through 1 ml of Ni-NTA resin (QIAGEN), and the flow through (containing cleaved MspP) was concentrated and further purified on a HiLoad 26/60 Superdex 75 gel filtration column (GE Healthcare), equilibrated in 25 mM HEPES (pH 7.5), 500 mM NaCl, 5% glycerol, 2 mM $MnCl_2$. The fractions corresponding to the MspP peak, as confirmed by SDS-PAGE, were pooled and concentrated to 50 mg/ml with 10 kDa cutoff Vivaspin concentrators (Sartorius). The concentrated protein was flash frozen in liquid nitrogen and stored at −80°C.

Crystallization and Data Collection

Purified MspP at 50 mg/ml was crystallized by the hanging-drop vapor-diffusion method. Crystallization solutions were: 18% (w/v) PEG12000, 85 mM Na cacodylate (pH 6.5), 120 mM Na acetate, 13% (v/v) glycerol for the enzyme-cacodylate complex; 30% (w/v) PEG3350, 100 mM Tris-HCl (pH 8.5), 250 mM $MgCl_2$ for the enzyme-phosphate complex; 20% (w/v) PEG4000, 100 mM Tris-HCl (pH 8.0), 200 mM $MgSO_4$, 10% glycerol for the enzyme-sulfate complex. To obtain crystals of the enzyme-phosphate complex, inorganic potassium phosphate (KH_2PO_4) at pH 7.0 was added at 10 mM to the 50 mg/ml MspP solution prior to crystallization. Rod-like crystals (0.2 × 0.2 × 0.5 mm) grew within a week at 18°C. The crystals were transferred to cryoprotectant solutions containing 80% crystallization solution and 20% (v/v) glycerol and flash frozen in liquid nitrogen. Data sets from the enzyme-cacodylate and enzyme-sulfate complex were collected on the ESRF beamlines ID14-2 and ID14-4, respectively; the data set from the enzyme-phosphate complex was collected on the EMBL/DESY beamline BW7A. Diffraction data were collected at 100 K from a single frozen crystal in each case. Data processing and reduction were carried out with either the programs MOSFLM, SCALA, and TRUNCATE from the CCP4 software package (CCP4, 1994) or the XDS/XSCALE suite (Kabsch, 1993). Data collection statistics are provided in Table 2.

Structure Refinement

The 3D structure was first solved from the MspP-cacodylate complex by molecular replacement with the program AMoRe (Navaza, 1994), and the coordinates of the catalytic domain of *M. tuberculosis* PstP solved in our laboratory (PDB code: 2CM1; A.W., M.B., and P.M.A., unpublished data) as the search model. The MspP model was first built from the best molecular replacement solution with ARP/wARP (Perrakis et al., 1999) and subsequently refined by alternate cycles of restrained refinement at 1.45 Å resolution with the program Refmac5 (Murshudov et al., 1999) and manual model building with the programs O (Jones et al., 1991) and Coot (Emsley and Cowtan, 2004). The refined model without ligands was used to solve the phosphate-bound MspP structure by molecular replacement methods with Amore. This model was rebuilt with ARP/wARP and first refined isotropically at 1.0 Å resolution with Refmac5. The model was then improved by several cycles of visual inspection and manual modeling of the solvent and disordered parts with Coot alternated with conjugate gradient least-squares refinement with the program SHELXL97 (Sheldrick and Schneider, 1997). Anisotropic refinement was performed with default restraints; the effective resolution was stepwise increased from 1.00 Å up to 0.83 Å. Hydrogen atoms were added to the riding positions, except on solvent and protein hydroxyl groups. A partially refined model at 1.0 Å resolution (without ligands) was taken as the starting structure for the refinement of the isomorphous enzyme-sulfate complex. The structure was refined anisotropically with SHELXL97 at 1.08 Å resolution, and model inspection and building was carried out with Coot. Final refinement statistics are shown in Table 2.

Phosphatase Assay Conditions

All enzymatic assays were performed with the substrate *p*-nitrophenyl phosphate (*p*NPP), and the absorbance at 405 nm was read in a 96-well GENios (Tecan) microplate reader. Each point corresponded to at least triplicate measures and values for k_{cat} and K_m were deduced from Lineweaver-Burke plots. The reaction buffer was 5 mM $MnCl_2$, 100 mM Tris-HCl (pH 7.5), and initial linear rates were determined with an extinction coefficient of 12.5 $mM^{-1}\ cm^{-1}$. To determine the kinetic constants, MspP was incubated with 5 mM $MnCl_2$ and varying amounts of *p*NPP. To determine the dependance on metal, both proteins were incubated with 5 mM *p*NPP and varying amounts of $MnCl_2$ or 50 mM *p*NPP and varying amounts of $MgCl_2$.

The specific activity of wild-type MspP (0.07 μM) was measured as a function of pH in 100 mM BisTris, 5 mM $MnCl_2$, and 5 mN *p*NPP. The molar extinction coefficients (ε) of *p*NP were determined for each

pH value by measuring the absorbance at 405 nm for solutions of 0.1 mM *p*NP in the same buffer, and each point corresponds to at least triplicate measurements.

Supplemental Data

Supplemental data include double difference anomalous maps (Figure S1) calculated with diffraction data collected above and below the Mn and As K edges (Table S1) and are available at http://www.structure.org/cgi/content/full/15/7/863/DC1/.

ACKNOWLEDGMENTS

We thank A. Haouz for help with robotic crystallization, the staff at the European Synchrotron Radiation Facility and Deutsches Elektronen-Synchrotron/European Molecular Biology Laboratory for assistance during data collection, and F. Schaeffer and S.T. Cole for helpful discussions. This project was supported by grants from the Institut Pasteur (GPH-Tuberculose) and the European Commission (NM4TB, contract LSHP-CT-2005-018923).

Received: March 7, 2007
Revised: June 7, 2007
Accepted: June 10, 2007
Published: July 17, 2007

REFERENCES

Barford, D. (1996). Molecular mechanisms of the protein serine/threonine phosphatases. Trends Biochem. Sci. *21*, 407–412.

Barford, D., Das, A.K., and Egloff, M.P. (1998). The structure and mechanism of protein phosphatases: insights into catalysis and regulation. Annu. Rev. Biophys. Biomol. Struct. *27*, 133–164.

Benini, S., Rypniewski, W.R., Wilson, K.S., Ciurli, S., and Mangani, S. (2001). Structure-based rationalization of urease inhibition by phosphate: novel insights into the enzyme mechanism. J. Biol. Inorg. Chem. *6*, 778–790.

Boitel, B., Ortiz-Lombardia, M., Duran, R., Pompeo, F., Cole, S.T., Cerveñansky, C., and Alzari, P.M. (2003). PknB kinase activity is regulated by phosphorylation in two Thr residues and dephosphorylation by PstP, the cognate phospho-Ser/Thr phosphatase, in *Mycobacterium tuberculosis*. Mol. Microbiol. *49*, 1493–1508.

Bork, P., Brown, N.P., Hegyi, H., and Schultz, J. (1996). The protein phosphatase 2C (PP2C) superfamily: detection of bacterial homologues. Protein Sci. *5*, 1421–1425.

CCP4 (Collaborative Computational Project, Number 4) (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.

Chen, G., Edwards, T., D'Souza, V.M., and Holz, R.C. (1997). Mechanistic studies on the aminopeptidase from *Aeromonas proteolytica*: a two-metal ion mechanism for peptide hydrolysis. Biochemistry *36*, 4278–4286.

Cleland, W.W., and Hengge, A.C. (2006). Enzymatic mechanisms of phosphate and sulfate transfer. Chem. Rev. *106*, 3252–3278.

Cohen, P. (1989). The structure and regulation of protein phosphatases. Annu. Rev. Biochem. *58*, 453–508.

Das, A.K., Helps, N.R., Cohen, P.T., and Barford, D. (1996). Crystal structure of the protein serine/threonine phosphatase 2C at 2.0 A resolution. EMBO J. *15*, 6798–6809.

Egloff, M.P., Cohen, P.T., Reinemer, P., and Barford, D. (1995). Crystal structure of the catalytic subunit of human protein phosphatase 1 and its complex with tungstate. J. Mol. Biol. *254*, 942–959.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Fjeld, C.C., and Denu, J.M. (1999). Kinetic analysis of human serine threonine protein phosphatase 2C alpha. J. Biol. Chem. *274*, 20336–20343.

Gallego, M., and Virshup, D.M. (2005). Protein serine/threonine phosphatases: life, death, and sleeping. Curr. Opin. Cell Biol. *17*, 197–202.

Hengge, A.C. (2001). Isotope effects in the study of enzymatic phosphoryl transfer reactions. FEBS Lett. *501*, 99–102.

Jackson, M.D., and Denu, J.M. (2001). Molecular reactions of protein phosphatases–insights from structure and chemistry. Chem. Rev. *101*, 2313–2340.

Jackson, M.D., Fjeld, C.C., and Denu, J.M. (2003). Probing the function of conserved residues in the serine/threonine phosphatase PP2C alpha. Biochemistry *42*, 8513–8521.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.

Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J. Appl. Cryst. *26*, 795–800.

Kennelly, P.J. (2002). Protein kinases and protein phosphatases in prokaryotes: a genomic perspective. FEMS Microbiol. Lett. *206*, 1–8.

Kennelly, P.J. (2003). Archaeal protein kinases and protein phosphatases: insights from genomics and biochemistry. Biochem. J. *370*, 373–389.

Lucast, L.J., Batey, R.T., and Doudna, J.A. (2001). Large-scale purification of a stable form of recombinant tobacco etch virus protease. Biotechniques *30*, 544–550.

Mertz, P., Yu, L., Sikkink, R., and Rusnak, F. (1997). Kinetic and spectroscopic analyses of mutants of a conserved histidine in the metallophosphatases calcineurin and lambda protein phosphatase. J. Biol. Chem. *272*, 21296–21302.

Murshudov, G.N., Vagin, A.A., Lebedev, A., Wilson, K.S., and Dodson, E.J. (1999). Efficient anisotropic refinement of macromolecular structures using FFT. Acta Crystallogr. D Biol. Crystallogr. *55*, 247–255.

Navaza, J. (1994). AMoRe: an automated package for molecular replacement. Acta Crystallogr. A *50*, 157–163.

Olsen, J.V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006). Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. Cell *127*, 635–648.

Perrakis, A., Morris, R., and Lamzin, V.S. (1999). Automated protein model building combined with iterative structure refinement. Nat. Struct. Biol. *6*, 458–463.

Pohjanjoki, P., Lahti, R., Goldman, A., and Cooperman, B.S. (1998). Evolutionary conservation of enzymatic catalysis: quantitative comparison of the effects of mutation of aligned residues in *Saccharomyces cerevisiae* and *Escherichia coli* inorganic pyrophosphatases on enzymatic activity. Biochemistry *37*, 1754–1761.

Pullen, K.E., Ng, H.L., Sung, P.Y., Good, M.C., Smith, S.M., and Alber, T. (2004). An alternate conformation and a third metal in PstP/Ppp, the *M. tuberculosis* PP2C-family Ser/Thr protein phosphatase. Structure *12*, 1947–1954.

Rusnak, F., and Mertz, P. (2000). Calcineurin: form and function. Physiol. Rev. *80*, 1483–1521.

Schenk, G., Gahan, L.R., Carrington, L.E., Mitic, N., Valizadeh, M., Hamilton, S.E., de Jersey, J., and Guddat, L.W. (2005). Phosphate forms an unusual tripodal complex with the Fe-Mn center of sweet potato purple acid phosphatase. Proc. Natl. Acad. Sci. USA *102*, 273–278.

Sheldrick, G.M., and Schneider, T.R. (1997). SHELXL: high resolution refinement. Methods Enzymol. *277*, 319–343.

Swingle, M.R., Honkanen, R.E., and Ciszak, E.M. (2004). Structural basis for the catalytic activity of human serine/threonine protein phosphatase-5. J. Biol. Chem. *279*, 33992–33999.

Voegtli, W.C., White, D.J., Reiter, N.J., Rusnak, F., and Rosenzweig, A.C. (2000). Structure of the bacteriophage lambda Ser/Thr protein phosphatase with sulfate ion bound in two coordination modes. Biochemistry *39*, 15365–15374.

Williams, N.H. (2004). Models for biological phosphoryl transfer. Biochim. Biophys. Acta *1697*, 279–287.

Accession Numbers

Atomic coordinates and structure factors have been deposited in the PDB with accession codes 2JFS (MspP-cacodylate complex), 2JFR (MspP-phosphate complex), and 2JFT (MspP-sulfate complex).

