Acta Crystallographica Section D
Biological Crystallography
ISSN 1399-0047

Structural elucidation of a dual-activity PAP phosphatase-1 from *Entamoeba histolytica* capable of hydrolysing both 3′-phosphoadenosine 5′-phosphate and inositol 1,4-bisphosphate

Khaja Faisal Tarique, Syed Arif Abdul Rehman and S. Gourinath*

School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067, India

Correspondence e-mail: samudralag@yahoo.com

Received 21 March 2014
Accepted 6 May 2014

PDB references: PAP phosphatase-1, 4o7i; complex with AMP, 4hxv

The enzyme 3′-phosphoadenosine 5′-phosphatase-1 (PAP phosphatase-1) is a member of the Li^+-sensitive Mg^{2+}-dependent phosphatase superfamily, or inositol monophosphatase (IMPase) superfamily, and is an important regulator of the sulfate-activation pathway in all living organisms. Inhibition of this enzyme leads to accumulation of the toxic byproduct 3′-phosphoadenosine 5′-phosphate (PAP), which could be lethal to the organism. Genomic analysis of *Entamoeba histolytica* suggests the presence of two isoforms of PAP phosphatase. The PAP phosphatase-1 isoform of this organism is shown to be active over wide ranges of pH and temperature. Interestingly, this enzyme is inhibited by sub-millimolar concentrations of Li^+, while being insensitive to Na^+. Interestingly, the enzyme showed activity towards both PAP and inositol 1,4-bisphosphate and behaved as an inositol polyphosphate 1-phosphatase. Crystal structures of this enzyme in its native form and in complex with adenosine 5′-monophosphate have been determined to 2.1 and 2.6 Å resolution, respectively. The PAP phosphatase-1 structure is divided into two domains, namely $\alpha+\beta$ and α/β, and the substrate and metal ions bind between them. This is a first structure of any PAP phosphatase to be determined from a human parasitic protozoan. This enzyme appears to function using a mechanism involving three-metal-ion assisted catalysis. Comparison with other structures indicates that the sensitivity to alkali-metal ions may depend on the orientation of a specific catalytic loop.

1. Introduction

Sulfur is mostly available to organisms in the form of inorganic sulfate, which is biologically inert and needs activation/fixation in order to enter into cellular metabolism. Activation begins with sulfate adenylation catalyzed by ATP sulfurylase (ATPS), leading to the production of adenosine 5′-phosphosulfate (APDS) (Nozaki *et al.*, 1998) (Supplementary Fig. S1[1]). This reaction is thermodynamically unfavourable, and the energy released from PP_i hydrolysis by inorganic pyrophosphatase drives the formation of the phospho-sulfuric anhydride bond of APS (Bradley *et al.*, 2009; Nozaki *et al.*, 1998; Hatzios *et al.*, 2008; Thomas *et al.*, 1992). APS is further phosphorylated by APS kinase (APSK) to produce 3′-phospho-APS (PAPS), which serves as the universal sulfuryl-group donor for sulfotransferases (Bradley *et al.*, 2009; Nozaki *et al.*, 1998; Hatzios *et al.*, 2008; Thomas *et al.*, 1992). APS and PAPS can also be reduced by the respective reductases, and the reduced sulfur (H_2S) is incorporated directly into metabolites by many terminal enzymes (Bradley *et al.*, 2009; Nozaki *et al.*, 1998;

[1] Supporting information has been deposited in the IUCr electronic archive (Reference: RR5071).

Hatzios *et al.*, 2008; Thomas *et al.*, 1992). 3′-Phosphoadenosine 5′-phosphate (PAP) is the end product of all of the sulfonation reactions catalyzed by sulfotransferases and of the reduction of PAPS to sulfite by PAPS reductase (Hatzios *et al.*, 2008). The enzyme 3′(2′),5′-bisphosphate nucleotidase (PAP phosphatase) catalyzes a reaction in which PAP is converted to adenosine 5′-monophosphate and inorganic phosphate (Ramaswamy & Jakoby, 1987; Murguía *et al.*, 1995). PAP acts as a competitive inhibitor for many sulfotransferases that utilize PAPS as the sulfate donor, and its accumulation has been found to produce a variety of toxic effects such as the inhibition of RNA-processing enzymes (Toledano *et al.*, 2012). Therefore, together with APS kinase, PAP phosphatase regulates and drives the sulfate-activation pathway, which would otherwise become futile in its absence. Mutations in the PAP phosphatase gene result in cysteine and sulfite auxotrophy in many organisms and these phosphatases are potential targets for lithium therapy in manic patients (Yenush *et al.*, 2000). In addition, the phosphatase activity is also required for tolerance to superoxide stress and to salt in some organisms (Murguía *et al.*, 1995; Zhang & Biswas, 2009). Taken together, PAP phosphatase plays a key regulatory role in the sulfate-assimilation pathway (Neuwald *et al.*, 1992). PAP phosphatase-1 belongs to the IMPase superfamily of phosphatases, which are inhibited at submillimolar concentrations of Li^+ and show an absolute requirement for a divalent ion, usually Mg^{2+}, for their activity. Other members of this superfamily include fructose-1,6-bisphosphatase (FBPase), inisitol monophosphatase (IMPase) and inositol polyphosphate 1-phosphatases (IPPases). The lengths of these proteins range from 252 to 400 amino-acid residues and they share low sequence identities with one another. The crystal structures of the IMPase superfamily that have been determined include those of human IMPase (Bone *et al.*, 1992), *Staphylococcus aureus* IMPase (Bhattacharyya *et al.*, 2011), bovine IPPase (York *et al.*, 1994), swine FBPase (Xue *et al.*, 1994) and PAP phosphatases from yeast (Albert *et al.*, 2000) and rat (Patel *et al.*, 2002). Together with other homologous and heterologous protein structures in the Protein Data Bank (PDB), they share a similar core structure consisting of about 138 residues with an r.m.s.d. of about 1 Å and contain conserved residues essential for metal binding and substrate hydrolysis (Ke *et al.*, 1989; York *et al.*, 1995; Spiegelberg *et al.*, 1999).

Here, we focus on the sulfate-activation pathway of *Entamoeba histolytica*, especially on its 3′(2′),5′-bisphosphate nucleotidase (PAP phosphatase-1) enzyme. This pathogen is anaerobic or microaerophilic in nature and cannot tolerate a high concentration of oxygen during infection. Cysteine and its derived molecules are the major sulfur-containing molecules that play essential roles in the antioxidative defence mechanism of *E. histolytica*, and are therefore important for its survival (Mehlotra, 1996; Chinthalapudi *et al.*, 2008; Kumar *et al.*, 2011; Raj *et al.*, 2012). Sulfate-activation pathway enzymes, which were identified in the mitosomal fraction of *E. histolytica*, are critical for the survival of this protozoan parasite (Mi-ichi *et al.*, 2009) and hence can serve as possible drug targets. Activated sulfur is mostly incorporated in sulfolipid-I, an abundant sulfated glycolipid and a virulence factor in *E. histolytica* (Mi-ichi *et al.*, 2011). *In silico* studies of the *E. histolytica* genome suggest the presence of a number of phosphatases (Anwar & Gourinath, 2013), including two isozymes of 3′(2′),5′-bisphosphate nucleotidase.

In the current study, we have cloned, overexpressed, purified and solved crystal structures of *E. histolytica* PAP phosphatase-1 in complex with Mg^{2+} and in complex with AMP and Na^+ as well as Mg^{2+}. Biochemical characterization of this enzyme has shown it to catalyse the hydrolysis of both 3′-phosphoadenosine 5′-phosphate (PAP) and inositol 1,4-bisphosphate [Ins(1,4)P_2]. The enzyme is stable over wide ranges of temperature, NaCl concentration and pH, but its activity is diminished by Li^+ at submillimolar concentrations. Immunolocalization studies of PAP phosphatase-1 show it to be present throughout the cytoplasm.

2. Materials and methods

2.1. Identification of the PAP phosphatase-1 gene in *E. histolytica*

Two isozymes of 3′(2′),5′-bisphosphate nucleotidase have been identified in the *E. histolytica* genome database, EHI_179820 and EHI_193350 (Brinkac *et al.*, 2010), corresponding to GenBank accession Nos. XP_655585.1 and XP_651950.1. The two corresponding putative protein isoforms, PAP phosphatase-1 (317 amino acids) and PAP phosphatase-2 (285 amino acids), share less than 25% sequence identity with each other; nevertheless, they both contain the conserved signature sequence motif for metal-ion binding, a characteristic feature of the inositol monophosphatase superfamily of proteins.

2.2. Preparation of recombinant vector

The PAP phosphatase-1 gene was PCR-amplified from the genomic DNA of *E. histolytica* HM-1-IMS with 5′-CATGCCATGGCAATGTCATTTGATAAAGAACTT-3′ (*Nco*I) as the forward primer and 5′-CGGCTCGAGTTTTAAATCAGAAAGTACATCT-3′ (*Xho*I) as the reverse primer. The amplified product of Pap phosphatase-1 (954 bp) was digested with *Nco*I and *Xho*I and ligated into pET-28b expression vector with a C-terminal His tag. The absence of any mutations was confirmed by gene sequencing of the recombinant pET-28 vector.

2.3. Expression and purification of PAP phosphatase-1 in *Escherichia coli*

pET-28b-PAP phosphatase-1 plasmids were transformed into *E. coli* strain BL21(DE3). Positive colonies were selected from kanamycin-containing LB plates and inoculated into 50 ml TB medium for overnight growth. 1% of this culture was inoculated into a secondary culture containing 50 mg l^{-1} kanamycin. The culture was incubated at 37°C until the OD_{600} reached 0.6–0.8. Protein expression was initiated by the addition of isopropyl β-D-1-thiogalactopyranoside to a final concentration of 200 μM. The temperature was lowered to

16°C after induction and the culture was allowed to grow overnight. The cells were harvested at 4°C by centrifugation at 8000 rev min^{-1} for 5 min and the pellet was suspended in suspension buffer (1 g of pellet per 10 ml) consisting of 50 m*M* Tris pH 7.5, 150 m*M* NaCl, 1 m*M* PMSF, 0.1% Triton X-100, 0.2 m*M* EDTA, 10 m*M* imidazole, 0.05 mg ml^{-1} lysozyme, 5 m*M* β-mercaptoethanol. The lysed cells were sonicated and then centrifuged at 18 000 rev min^{-1} for 1 h. The cleared supernatant was applied onto an Ni Sepharose column and the column was then equilibrated with 50 ml washing buffer (50 m*M* Tris pH 7.5, 150 m*M* NaCl, 5 m*M* β-mercaptoethanol, 40 m*M* imidazole). The protein was eluted in buffer consisting of 50 m*M* Tris pH 7.5, 150 m*M* NaCl, 250 m*M* imidazole, 5% glycerol, 5 m*M* β-mercaptoethanol. For further purification and removal of imidazole, the concentrated protein was loaded onto a HiLoad G200 16/60 column and the flow rate was kept at 1 ml min^{-1}. The purity of the desired protein was checked on 12% SDS–PAGE and purified fractions were pooled and concentrated using a Centricon (Supplementary Fig. S2). The column was pre-equilibrated with running buffer consisting of 10 m*M* Tris pH 7.5, 150 m*M* NaCl, 5% glycerol, 5 m*M* β-mercaptoethanol. Glycerol was added to increase the solubility and stability of the protein. The identity of the PAP phosphatase-1 protein was confirmed by tryptic digestion and mass fingerprinting using a Bruker ESI Ion Trap mass spectrometer. Protein concentration was determined by measuring the UV absorption at 280 nm. An approximate extinction coefficient of 24 410 M^{-1} cm^{-1} was calculated using the *ExPASy* server (Wilkins *et al.*, 1999).

2.4. Enzyme assay and thermal denaturation studies

The nucleotidase activity of PAP phosphatase-1 was measured by the malachite green method (Baykov *et al.*, 1988) using PAP as the primary substrate. In the present work, a 100 µl reaction mixture consisting of 50 m*M* Tris pH 7.0, 100 µ*M* PAP, 5 m*M* $MgCl_2$ and 2 µg purified PAP phosphatase-1 protein was incubated for 5 min at room temperature and the reaction was then stopped by the addition of 50 µl malachite green dye solution. After 20 min, the blue-coloured phosphomolybdate complex was detected at 650 nm in a microplate reader. The absorbance and hence the activity is directly proportional to the amount of P_i released, which forms a green phosphomolybdate malachite complex. To determine the optimum temperature, the optimum pH, the salt sensitivity and the types of divalent metal ions acting as cofactors for the activity of PAP phosphatase-1, these factors were varied during the standard assay while keeping the components of the reaction mixture unchanged at standard conditions. To determine the thermal stability of PAP phosphatase-1, denaturation experiments were carried out by CD spectroscopy equipped with a Peltier-type temperature controller. The instrument was calibrated with D-10-camphosulfuric acid. Spectra in the far-UV range (222 nm) were measured with the concentration of PAP phosphatase-1 kept at 0.1 mg ml^{-1}. The path length of the cell was 0.1 cm. In order to reduce the noise owing to chloride ions, sodium chloride was replaced by sodium chlorate as a salt. In each step of protein purification, *e.g.* from affinity to gel-filtration chromatography, 20 m*M* phosphate buffer pH 8.0 was used. The results were interpreted in terms of the fraction of protein unfolded with increase in temperature.

2.5. Dynamic light-scattering measurements

Dynamic light-scattering (DLS) measurements were performed using SpectroSize300 DLS equipment (Molecular Dimensions) equipped with a temperature-controlled microsampler. The concentration of PAP phosphatase-1 used in this experiment was approximately 8 mg ml^{-1}. The protein was centrifuged at 13 000 rev min^{-1} for 30 min at 4°C and filtered through a 0.25 µm Whatman syringe filter prior to the experiment. For each experiment ten scans were taken and the duration of each scan was 20 s. The mean hydrodynamic radius, standard deviation, polydispersity and percentage of peak area were analyzed using the *Dynamics* 6.10 software using the optimized resolution.

2.6. Generation of antibodies and Western blotting

Western blotting was performed by transfer of protein that was separated by SDS–PAGE (Laemmli, 1970) to a polyvinylidene difluoride (PVDF) membrane using a semi-dry transfer system (Bio-Rad). SDS was included in the transfer buffer in order to increase the transfer efficiency of high-molecular-weight proteins. The membrane was blocked with 5% Blotto (Genotech Inc, USA) in Tris-buffered saline with Tween (TBST) for 1 h. Protein bands were probed with anti-PAP phosphatase-1 (1:5000) raised in mice, washed with TBST, incubated with a 1:10 000 dilution of HRP-conjugated secondary antibody (Bio-Rad) and finally detected by enhanced chemiluminescence.

2.7. Immunofluorescence labelling

E. histolytica HM-1-IMS cells were resuspended in TYI-S-33 incomplete medium and placed on a cover slip wiped with acetone for 10–20 min at 37°C. The media were decanted and the cells were fixed by keeping them in 3.7% paraformaldehyde (PFA) at 37°C for 30 min. PBS was used to remove PFA and the cells were permeabilized by adding 1 ml 0.1% Triton X-100 for 15 min at 37°C. They were washed again with PBS and neutralized with 50 m*M* NH_4Cl for 30 min at room temperature. Excess NH_4Cl was removed by washing with PBS. The cells were incubated with primary anti-PAP phosphatase-1 (1:200) at 37°C for 1 h. Prior to incubation with secondary antibodies, the slides were washed with 1% BSA-PBS. 100 µl of Hoechst dye was added onto the slide and incubated for 30 min at 37°C. Anti-rabbit Alexa Fluor 488 for PAP phosphatase-1 was used. The cover slips were sealed and an Olympus Fluoview FV100 laser scanning microscope was used to view confocal images.

2.8. Crystallization and X-ray data collection for PAP phosphatase-1

Purified PAP phosphatase-1 was concentrated to 30 mg ml^{-1} in 10 m*M* Tris pH 7.8 containing 2 m*M* $MgCl_2$, 150 m*M* NaCl, 1 m*M* ammonium phosphate, 5% glycerol, 5 m*M* β-mercaptoethanol and 1 m*M* AMP. Following extensive robotic (Mosquito) crystallization trials, the hanging-drop vapour-diffusion method yielded feather-shaped crystals using the Jena Bioscience Kinase Screen (condition C1). A 600 nl drop was set up with a 1:1 ratio of protein and precipitant and was equilibrated against 200 µl precipitant. Crystals appeared within 10 d in 28% PEG 4K, 200 m*M* lithium acetate, 100 m*M* Tris pH 7.5. The crystals were equilibrated with a cryoprotectant solution consisting of 30% PEG 4K, 100 m*M* Tris pH 7.5, 200 m*M* lithium acetate. Crystals were then mounted in cryoloops and flash-cooled in liquid nitrogen at 100 K. The crystals diffracted X-rays to 2.1 Å resolution for the native structure and 2.6 Å resolution for the AMP-complex structure using an in-house Bruker MicroStar generator (Cu $K\alpha$, 1.54 Å) and a MAR imaging plate at the Advanced Instrumentation Research Facility (AIRF), Jawaharlal Nehru University (JNU). The data were indexed, processed and scaled with *HKL*-2000.

2.9. Structure solution and refinement

The PAP phosphatase-1 structure was determined by molecular replacement using yeast Hal2P as a search model (PDB entry 1qgx; Albert *et al.*, 2000), with which it has 39% sequence identity. PAP phosphatase-1 crystallized in space group $P2_12_12_1$ with one molecule in the asymmetric unit. The output of *CHAINSAW* run on the aligned sequences of PAP phosphatase-1 and yeast Hal2P was used for molecular replacement by *Phaser* in *CCP*4 (McCoy *et al.*, 2007). The best solution had a TFZ score (translation-function *Z*-score) of 19.2 and an LLG (log-likelihood gain) of 389. An initial cycle of restrained refinement using *REFMAC*5 (Murshudov *et al.*, 2011) resulted in a decrease in the *R* factor to 36%. The model was then submitted to *ARP*/*wARP* for autobuilding (Langer *et al.*, 2008), which successfully built 90% of the side chains of the model into good-quality electron density. The remaining parts of the polypeptide were built manually with *Coot* (Emsley & Cowtan, 2004) and the phosphate ion, magnesium ion and water molecules were subsequently added according to the electron density. Multiple rounds of manual adjustment of the model in *Coot* and refinement with *REFMAC*5 were carried out (Murshudov *et al.*, 2011). The model at this stage had an *R* and R_{free} of 21 and 26%, respectively, and was subsequently submitted to the *PDB_REDO* web-based server for improvement of the refinement statistics, which decreased the *R* and R_{free} values to 18 and 23%, respectively, with no ambiguity in the electron density (Joosten *et al.*, 2012). The final model consists of 2485 protein atoms, 167 water molecules, one phosphate ion and one Mg^{2+} ion, with good refinement statistics (Table 1) and electron density. The structure factors and coordinates for PAP phosphatase-1 have been deposited in the PDB as entry 4o7i. Subsequently, this refined structure was chosen as a template to solve the structure of PAP phosphatase-1 complexed with AMP, Mg^{2+} and Na^{+} to 2.6 Å resolution (PDB entry 4hxv). These metal ions are coordinated by a network of water molecules and negatively charged amino-acid groups to neutralize their positive charges. The identity of the metal ions and AMP were confirmed by their binding geometries, coordination spheres, temperature factors and the difference Fourier electron-density map. The quality of the model was checked by *PROCHECK* (Laskowski *et al.*, 1993).

Table 1
Data-collection and refinement statistics for PAP phosphatase-1.

Values in parentheses are for the highest resolution shell.

Data set	PAP phosphatase-1	PAP phosphatase-1–AMP
Crystallographic data		
X-ray source	AIRF, JNU	AIRF, JNU
Wavelength (Å)	1.54	1.54
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)	a = 39.4, b = 64.5, c = 117.2	a = 339.6, b = 63.9, c = 117.8
R_{sym} or R_{merge} (%)	13 (47)	18 (43)
Completeness (%)	95.5 (93.6)	98.9 (88.8)
Multiplicity	8.2 (6.3)	11.6 (10.8)
Average $I/\sigma(I)$	13.7 (3.7)	14.8 (5.3)
No. of molecules in asymmetric unit	1	1
Refinement		
Resolution (Å)	50–2.1 (2.18–2.10)	50–2.6 (2.68–2.59)
Total No. of observations	141135	112782
No. of unique observations	17183	9698
R_{work}/R_{free} (%)	18/23	23/28
Mean B factor (Å^2)	23	22.2
No. of atoms		
Protein	2453	2460
Mg	1	1
Na	0	1
AMP	0	23
Phosphate	5	0
Water	110	49
R.m.s. deviations		
Bonds (Å)	0.010	0.011
Bond angles (°)	1.34	1.44
Ramachandran plot (%)		
Most favoured region	92.9	91.9
Additionally allowed region	7.1	7.1
Generously allowed region	0.0	1.1
Disallowed region	0.0	0.0
PDB code	4o7i	4hxv

3. Results

3.1. PAP phosphatase-1 is present in the cytoplasm and is active over wide ranges of temperature, pH and metal-ion conditions

Purified PAP phosphatase-1 hydrolyses 3′-adenosine 5′-phosphate (PAP), and the amount of P_i released by the enzyme was measured at different pH values and temperatures, as well as with different metal ions, to determine the optimal activity of the enzyme and for biochemical characterization. The pH-dependent activity profile was determined by measuring the ability of this enzyme to hydrolyze PAP in the presence of $MgCl_2$ at pH values of between 4 and

10. PAP phosphatase-1 was stable and >80% active between pH 6 and 8.5 and was less than 40% active in acidic (below pH 5.0) and alkaline (above pH 9.0) conditions (Fig. 1*a*). PAP phosphatase-1 was most active at pH 7.0, which was thus used as the standard pH for further enzymatic studies. PAP phosphatase-1 was also found to be stable and active over a fairly wide range of temperatures; it is >80% active between 10 and 60°C, but with a steep drop off to <20% of the optimum activity at higher temperatures. The maximum activity of the enzyme was found to be at 50°C (Fig. 1*b*). All subsequent experiments for this enzyme were performed at 30°C unless otherwise indicated. In order to determine what type of divalent metal ion is acting as a cofactor, a standard enzyme assay was performed in the presence of 50 m*M* Tris pH 7.0 and 5 m*M* each of $CoCl_2$, $MgCl_2$, $ZnCl_2$, LiCl, $CaCl_2$ and $MnCl_2$. PAP phosphatase-1 showed activity not only with its natural cofactor Mg^{2+} but also with Mn^{2+}, Co^{2+} and, to some extent, Zn^{2+}. There was less than 20% activity (and in some cases no activity) in the presence of metal ions such as Ca^{2+} and Li^{+} (Fig. 1*c*). Interestingly, the nucleotidase activity of PAP phosphatase-1 was found to be higher with Co^{2+} than with the natural Mg^{2+} cofactor; the activity with Mg^{2+} in turn was found to be higher than that with Mn^{2+} or with Zn^{2+}. Mg^{2+}, being a natural cofactor for this enzyme, was chosen for carrying out all of the enzymatic studies unless stated otherwise. The optimum concentration of Mg^{2+} was found to be near 5 m*M*. In addition to the kinetic studies at elevated temperature, the thermostability of PAP phosphatase-1 was confirmed by circular-dichroism studies at 222 nm (Supplementary Fig. S3), which showed the melting temperature (T_m) of PAP phosphatase-1 to be about 61°C. Similar to SAL1 from *Arabidopsis* (Quintero *et al.*, 1996) and RnPIP from rat (López-Coronado *et al.*, 1999, PAP phosphatase-1 can hydrolyze both PAP and Ins(1,4)P_2 and therefore acts as a inositol-polyphosphate-1-phosphatase with dual substrate specificity. The efficiency of hydrolyzing Ins(1,4)P_2 is however about 15% of that of PAP (Fig. 1*d*). Taken together, these results reveal PAP phosphatase-1 to be a robust enzyme which unlike other enzymes of

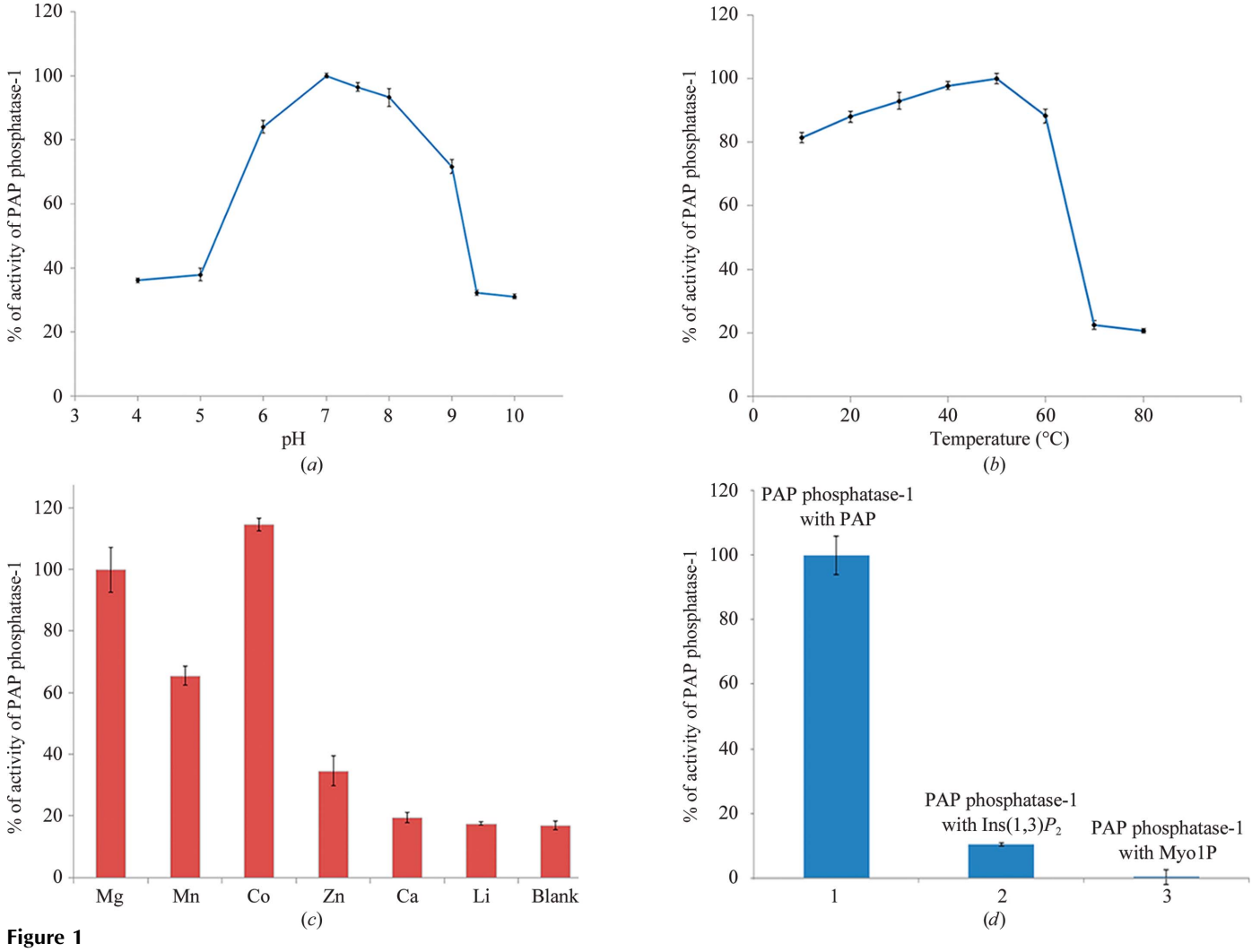

Figure 1
PAP phosphatase-1 activity under varying conditions. Phosphatase activity of PAP phosphatase-1 as a function of (*a*) pH, (*b*) temperature, (*c*) various metal ions and (*d*) different substrates. For PAP phosphatase-1 the standard 100 µl reaction mixture volume contained 50 m*M* Tris pH 7.0, 100 µ*M* PAP, 5 m*M* $MgCl_2$ and 2 µg purified PAP phosphatase-1 protein. Activities are expressed as a percentage of the maximum observed activities. The data shown represent the average of two independent experiments. Error bars represent the corresponding standard deviations.

sulfate activation, which are present in the mitosomes, was found to be present throughout the cytoplasm of *E. histolytica* (Supplementary Fig. S4). Anti-PAP phosphatase-1 stains endogenous PAP phosphatase-1 at 35 kDa (Supplementary Fig. S4).

3.2. Overall structure of PAP phosphatase-1

Statistics for the 2.1 Å resolution crystal structure of the native PAP phosphatase-1 and for the 2.6 Å resolution structure of the PAP phosphatase-1–AMP complex are shown in Table 1. In each crystal structure, the asymmetric unit is made up of one PAP phosphatase-1 monomer, which is consistent with our DLS and gel-filtration chromatography experiments (Supplementary Fig. S5), which also showed PAP phosphatase-1 to be monomeric as well as monodisperse. The conformations of the PAP phosphatase-1 protein with and without AMP bound are nearly identical: in each case the protein consists of nine α-helices, 14 β-strands and three short 3_{10}-helices (Fig. 2), and alignment of all 317 C^{α} atoms yields an r.m.s.d. of 0.3 Å. Similar to other members of the IMPase superfamily, the structure of PAP phosphatase-1 is divided into two domains. The N-terminal domain resembles the α+β structural class, and contains generally segregated α and β regions, with the latter consisting of mostly antiparallel β-sheets. Residues 1–184 in this domain contain helices H1–H5 and antiparallel β-sheets that include strands B1–B9. The C-terminal domain is an α/β structure, in which the α-helices and β-strands for the most part alternate along the sequence; residues 186–317 contain a parallel β-sheet that includes strands B10–B14 sandwiched between helices H6–H9 (Fig. 2). The two domains are connected to each other by a single 5–6-residue loop.

3.3. The binding of substrate and metal ions at the active site

The active site of the PAP phosphatase-1 enzyme lies at the junction of the N- and C-terminal domains and is identified in the structure of the complex by the presence of metal ions (Mg^{2+} and Na^{+}) and adenosine monophosphate, one of the byproducts of the reaction catalysed by PAP phosphatase-1. The identities of the metal ions were confirmed by their binding geometry, coordination sphere, temperature factors and the difference Fourier electron-density map. The structure of PAP phosphatase-1 bound with metal ions and AMP represents the stage after PAP (substrate) hydrolysis in which the product of the hydrolysis was found to be trapped. The crystal structures show numerous residues of PAP phosphatase-1 that take part in hydrogen bonding as well as ionic and hydrophobic interactions with the metal ions and AMP. These residues are conserved, or replaced with similar residues, in homologous structures (Albert *et al.*, 2000; Patel *et al.*, 2002). Amino-acid residues interacting with the metal ions and AMP at the active site and their atomic distances are also described in Supplementary Table S1.

Inspection of the PAP phosphatase-1–AMP complex structure indicates that AMP makes numerous stabilizing interactions with the protein. The adenine ring of AMP is stacked between the aromatic ring of Tyr251 and the imidazole ring of His203 in a coplanar manner (Fig. 3*a*). Although, as described above, the structures of PAP phosphatase-1 with and without AMP bound are very similar, and their overall *B* factors are nearly identical (22.1 $Å^2$

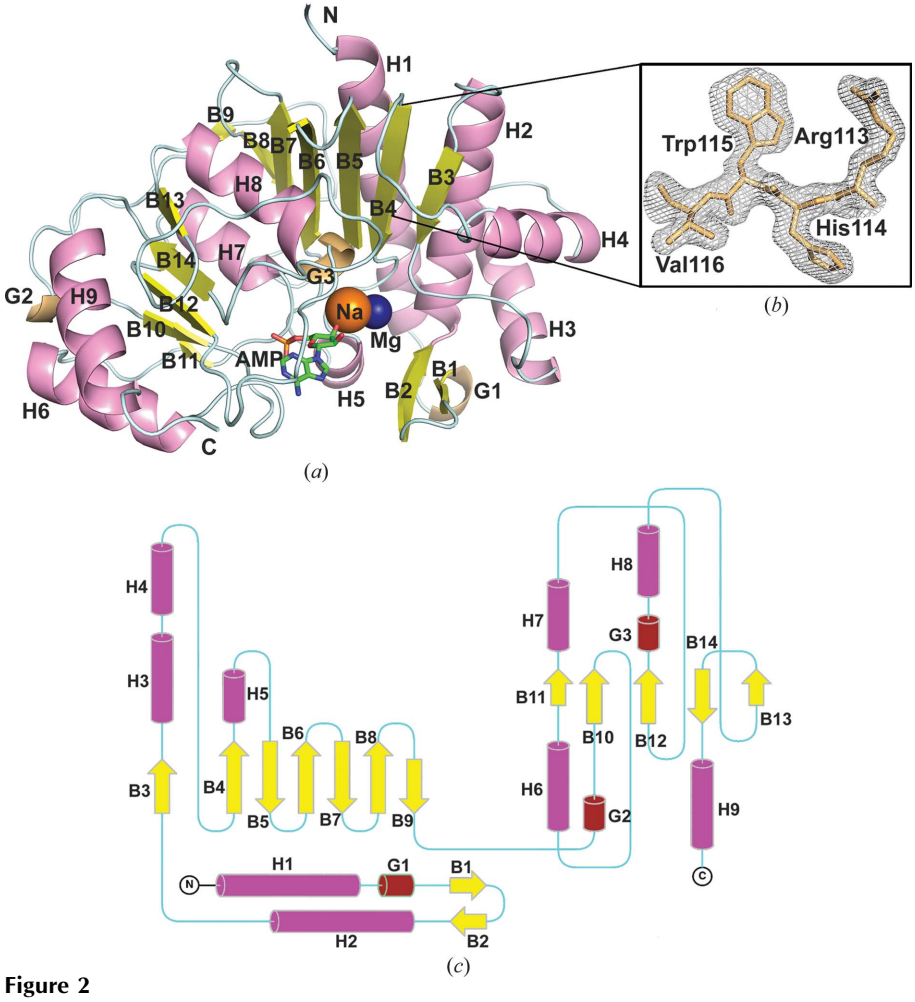

Figure 2
Structure of PAP phosphatase-1. (*a*) Ribbon representation of PAP phosphatase-1 with bound metal ions and AMP at the active site. (*b*) $2F_o - F_c$ electron density at a 1σ cutoff for a portion of B4. (*c*) Topology of the secondary-structural elements of PAP phosphatase-1 prepared by *TopDraw* (Bond, 2003). The N-terminal domain forms an approximate α+β fold, while the C-terminal domains belongs to the α/β class. Yellow arrows represent β-strands, pink cylinders represent α-helices and short brown cylinders labelled with a G are 3_{10}-helices.

for the native structure and 21.7 Å^2 for the complex), Tyr251 and His203 have significantly higher temperature factors in the native structure (38.8 and 25.5 Å^2, respectively) than when AMP is bound (19.7 and 21.8 Å^2, respectively). Moreover, the distance between Tyr251 and His203 in the AMP-complex structure is decreased by around 0.5 Å. In addition, the 5′-phosphate group of AMP makes ionic and hydrogen-bond interactions with polar amino-acid residues such as Lys230, Ser227, Ser198 and His203 (Fig. 3*b*). This phosphate group, as well as the ribose sugar and adenosine ring of the AMP molecule, interact with Arg244 of the protein; AMP also makes hydrophobic interactions with Gly122, Asp226, Ser202, Glu253 and Tyr251. Taken together, these observations indicate that the AMP-binding site becomes more compact and stable upon nucleotide binding. The active sites of Li^+-sensitive/Mg^{2+}-dependent phosphatases possess three metal-binding sites, denoted M1, M2 and M3, in a hydrophilic cavity (Albert *et al.*, 2000; Patel *et al.*, 2002; Johnson *et al.*, 2001). It has been proposed that the binding affinity of Mg^{2+} is highest for M1, intermediate for M2 and lowest for M3 (M1 > M2 > M3;

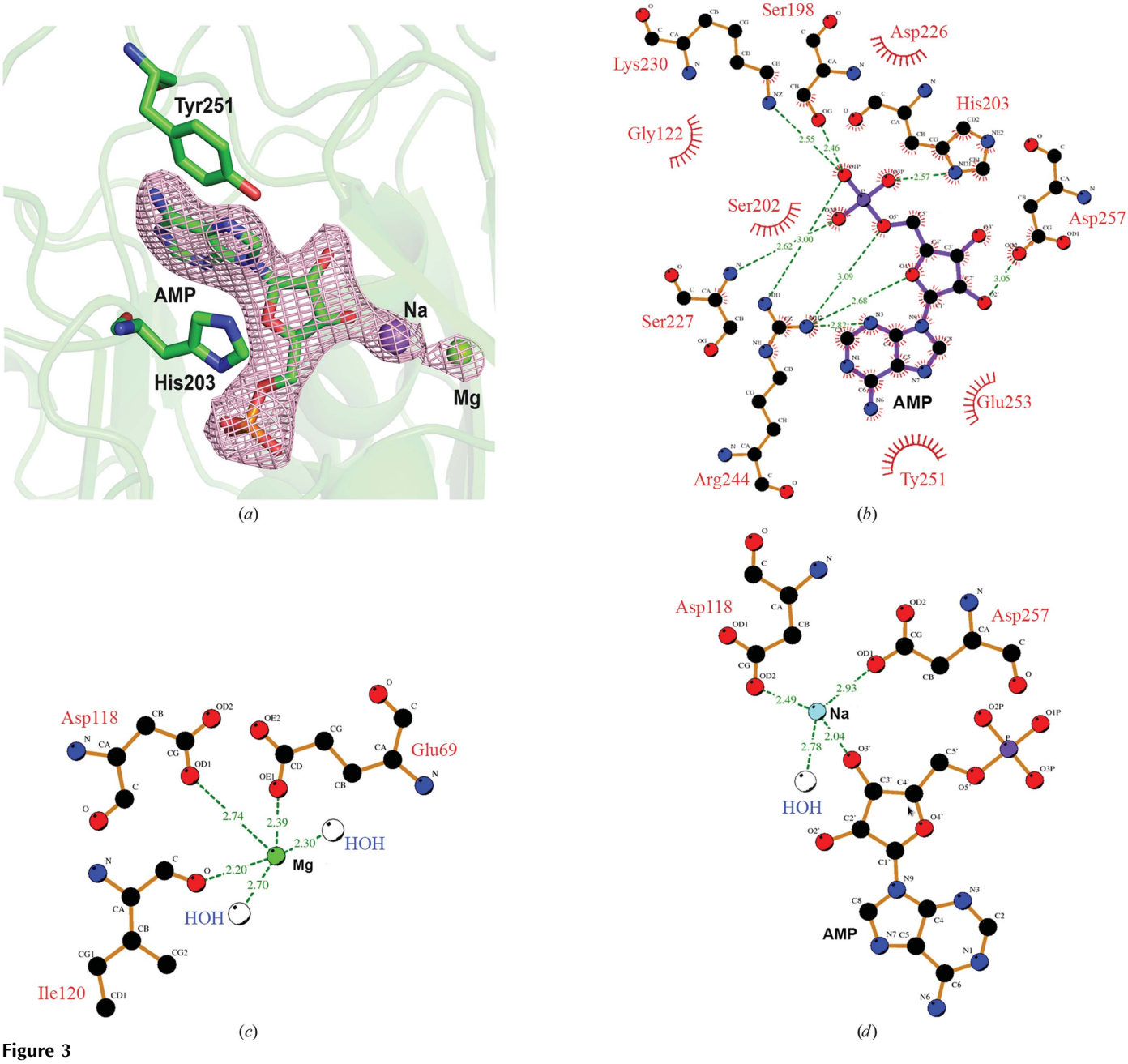

Figure 3
Ligand interaction with AMP and metal ions. (*a*) $2F_o - F_c$ electron-density map of the AMP molecule at 1σ cutoff. At the active site, the adenine ring of AMP is stacked between the aromatic rings of His203 and Tyr251. The bound Na^+ ion (purple) and Mg^{2+} ion (green) are also shown. (*b*, *c*, *d*) The interactions of the PAP phosphatase-1 protein with bound metal ions and AMP at the active site. The plots were generated by *LigPlot+* (Laskowski & Swindells, 2011). Hydrogen bonds are shown as green dotted lines, while arcs represent residues making nonbonded contacts with the ligands. The bound metal ions are shown in green (Mn) and cyan (Na) and waters are shown in white.

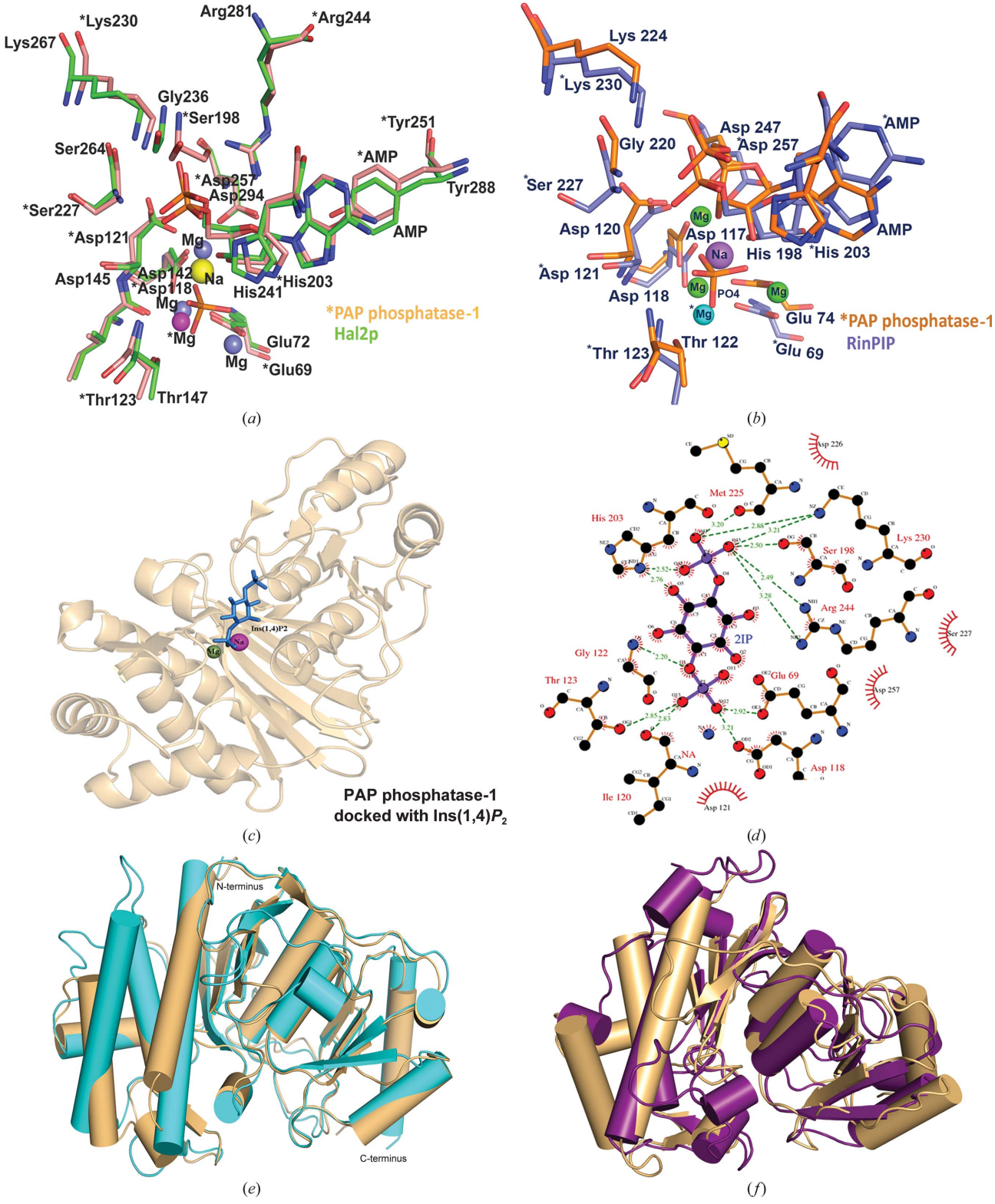

Figure 4
Structural comparison with homologous enzymes. (*a*, *b*) Superposition of the active site of PAP phosphatase-1 (orange) with those of the homologous structures from yeast (Hal2p; green) and rat (RnPIP; purple). Both the Hal2p and RnPIP structures have three metal ions and one phosphate ion at the active site. The amino-acid constellations and their specific orientation required for the binding of these metal ions and phosphate are conserved and superpose well on each other. Thus, the overall geometry of the residues coordinating the metal ions at the active site is same in all three structures. The web-based *FATCAT* server was used for all types of structural superposition (Ye & Godzik, 2004). (*c*) Modelling of Ins(1,4)P_2 into the active site of PAP phosphatase-1 was performed by the web-based *PatchDock* server (Schneidman-Duhovny *et al.*, 2005). (*d*) *LigPlot* representation of hydrogen-bond interaction between amino-acid residues and docked Ins(1,4)P_2 at the active site. (*e*, *f*) Superposition of PAP phosphatase-1 (orange) with Hal2P (cyan) and RnPIP (purple).

Gill *et al.*, 2005; Pollack *et al.*, 1994). (Binding of Mg^{2+} at M1 leads to cooperative binding of a second Mg^{2+} at M2, leading in turn to the binding of substrate at the active site; binding of a third Mg^{2+} is weakest and is noncooperative in nature.) The current crystallographic studies are consistent with these proposals. Electron density for a bound metal is observed only at M1 in the native PAP phosphatase-1 structure and only at M1 and M2 in the PAP phosphatase-1–AMP complex structure. (Density is also observed for a phosphate ion in the native structure.) The density in M1 is interpreted as Mg^{2+} owing to its octahedral coordination by Glu69 OE1, Asp118 OD1, Ile120 O, and two water molecules (Fig. 3*c*), while that of M2 is most probably Na^{+} according to the observed tetrahedral coordination by Asp118 OD2, Asp257 OD1, AMP and one water molecule (Fig. 3*d*). The lack of significant electron density for a bound metal at M3 of either PAP phosphatase-1 structure can be interpreted as M3 being unoccupied or even nonexistent in this protein, but we instead propose that this metal-binding site does exist in PAP phosphatase-1 and is likely to be occupied by Li^{+}. Unlike most metals, which are heavier than the common protein atoms, Li^{+} contains only two electrons (only twice the number in an H atom) and hence is a relatively poor scatterer of X-rays even when bound to the protein. In our studies, a high concentration of Li^{+} salt was used for the crystallization of PAP phosphatase-1 (see §2), and Li^{+} is known to often displace Mg^{2+} at M3 owing to the relatively low affinity of Mg^{2+} for M3 and the similar ionic radii of Li^{+} and Mg^{2+} (0.060 and 0.065 nm, respectively; Bhattacharyya *et al.*, 2012). (Binding of Li^{+} at M3 finally leads to noncompetitive inhibition for members of the IMPase superfamily.) Additional evidence for the presence of three metal-binding sites in PAP phosphatase-1, despite the lack of observed electron density at M3, is provided by a superposition with homologous structures (see also §3.4). The Mg^{2+}, Na^{+} and the probable Li^{+} ions at M1, M2 and M3 of PAP phosphatase-1 superpose onto the three Mg^{2+} ions at M1, M2 and M3 in the active sites of Hal2p and RnPIP. The amino-acid coordination of the metal ions at M1, M2 and M3, including the specific orientations that they adopt, are conserved and superpose well on one another in these three structures (Figs. 4*a* and 4*b*). Since metal-ion occupancy depends upon the special orientation of amino-acid residues acting as ligands, induced orientation of ligands owing to the availability of substrates and the concentration of metal ions used in the crystallization process (Gill *et al.*, 2005), the expectancy for the third Mg^{2+} at M3 would be high in the absence of Li^{+} and in the presence of other chelating agents such as phosphate ion. This is evident in the crystal structures of Hal2P and RnPIP, where three Mg^{2+} ions are found in the active site in the presence of AMP and phosphate ion. PAP phosphatase-1, however, required Li^{+} salt for crystallization, and this electron-poor, nearly X-ray-invisible metal is likely to have displaced Mg^{2+} in M3. The ability of PAP phosphatase-1 to catalyse the hydrolysis of Ins(1,4)P_2 can be explained by modelling it into the active site of the enzyme. Molecular docking of Ins(1,4)P_2 produced satisfactory hydrogen-bond networks with critical amino-acid residues necessary for catalysis (Fig. 4*c*). The 4′-phosphate group of Ins(1,4)P_2 maintains a similar type of ligand interactions at the active site as seen for the 5′-phosphate group of AMP (Fig. 4*d*). This clearly shows that PAP phosphatase-1 can accommodate both

Table 2
Structural superposition of members of the Li^{+}-sensitive/Mg^{2+}-dependent phosphatase superfamily with PAP phosphatase-1.

PDB code	Sequence identity (%)	No. of C^{α} atoms aligned	R.m.s.d. (Å)	Reference
1k9y	39	313	1.384	Albert *et al.* (2000)
2hhm	23	260	2.668	Bone *et al.* (1992)
2fvz	25	236	2.371	Structural Genomics Consortium (unpublished work)
1jp4	23	269	3.045	Patel *et al.* (2002)
2wef	22	271	3.134	Structural Genomics Consortium (unpublished work)
1inp	18	283	3.190	York *et al.* (1994)
1lbx	19	248	2.836	Stieglitz *et al.* (2002)
1frp	15	246	3.060	Xue *et al.* (1994)
1dk4	21	249	2.850	Stec *et al.* (2000)

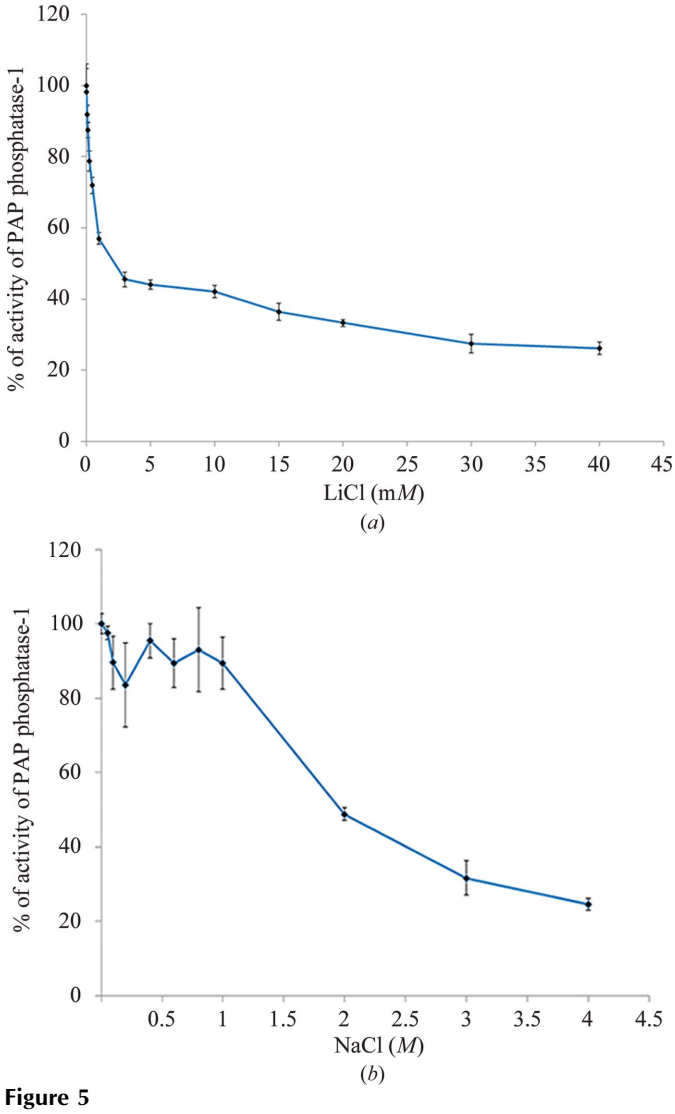

Figure 5
PAP phosphatase-1 activity in the presence of salts. (*a*) LiCl, (*b*) NaCl.

PAP and Ins(1,4)P_2 as substrates; however, the stabilizing interaction of PAP owing to the aromatic amino-acid residues (His203 and Tyr251) enhances its specificity and hydrolysis compared with Ins(1,4)P_2.

3.4. Comparisons with other structures

X-ray crystal structures of homologous proteins from the IMPase superfamily have been determined and, as explained earlier, they all share a similar fold and contain conserved signature sequence motifs for metal binding, *i.e.* D-X_n-EE-X_n-DP(I/L)DG(S/T)-X_n-WD-X_n-GG (Supplementary Fig. S6). Pairwise alignment and three-dimensional structure superposition of the PAP phosphatase-1 with these homologous proteins were carried out to identify the core secondary-structural elements and for structure comparison. Table 2 displays the number of aligned C^{α} atoms and the r.m.s.d.s between structures of some of the members of this superfamily taken from the PDB. The PAP phosphatase-1 structure superimposes well with yeast Hal2p (r.m.s.d. of 1.38 Å), while there were much larger differences with the rest of the homologues (r.m.s.d.s from 2.37 to 3.19 Å). Despite sharing maximum sequence identity (39%) and structure homology, Hal2p is only capable of hydrolysing PAP. Therefore, PAP phosphatase-1 is functionally more similar to RnPIP (22% sequence identity), which has dual specificity for both PAP as well as Ins(1,4)P_2. The majority of the structural differences between PAP phosphatase-1, Hal2p and RnPIP are owing to different lengths and orientations of some of the secondary-structural elements which are quite visible (Figs. 4*e* and 4*f*).

Table 3
IC_{50} of Li^+ inhibition for different members of the IMPase superfamily.

Enzyme	IC_{50} (m*M*)
PAP phosphatase-1	1.0
Hal2p (Murguía *et al.*, 1995)	0.1
CysQ from *Mycobacterium tuberculosis* (Hatzios *et al.*, 2008)	0.5
Human IMPase (Yenush *et al.*, 2000)	0.3
Sal1 from *Arabidopsis* (Quintero *et al.*, 1996)	0.2
Bovine inisitol polyphosphatase (Inhorn & Majerus, 1988)	0.3
IMPase from *Methanocaldococcus jannaschii* (MJ0109) (Stec *et al.*, 2000)	250
IMPase/FBPase from *Archaeoglobus fulgidus* (Stieglitz *et al.*, 2002)	290
DHAL2 from *Debaryomyces hansenii* (Aggarwal *et al.*, 2005)	2.4
PAP phosphatase from *Escherichia coli* (Fukuda *et al.*, 2007)	0.05
Porcine liver FBPase (Zhang *et al.*, 1996)	1–3
RnPIP (IPPase/Papase) (Patel *et al.*, 2002)	0.8
PAP phosphatase from *Arthrospira platensis* (Zhang *et al.*, 2006)	3.6

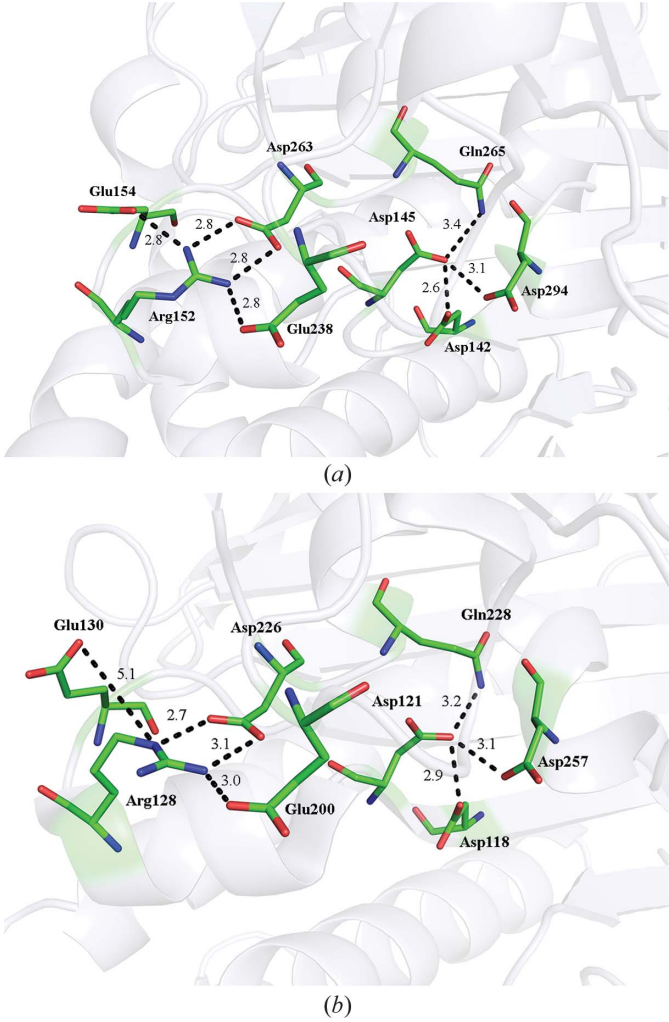

Figure 6
Ionic networks responsible for Li^+/Na^+ ion sensitivity. (*a*) In Hal2p, a set of salt bridges between Arg152, Glu154, Gln238 and Asp263, and the formation of an ionic network between Asp145, Asp142, Asp294 and Gln265, influence the binding and sensitivity of Li^+/Na^+ ions at the metal-binding site M2. (*b*) Although these amino acids are the same or are chemically similar in the corresponding positions of PAP phosphatase-1, *i.e.* Arg128, Glu130, Glu200, Asp226, Asp121, Asp118 and Gln228, and make similar ionic interactions, this enzyme, unlike Hal2p, is sodium insensitive.

4. Discussion

4.1. Mechanism of action

Based on the similarity of the overall conformations and active-site structures of PAP phosphatase-1 to those of previously determined homologous structures (see above; Patel *et al.*, 2002; Johnson *et al.*, 2001; Bhattacharyya *et al.*, 2012; Stieglitz *et al.*, 2002), we propose that the mechanism of the hydrolysis reaction should be similar. In PAP phosphatase-1, the reaction would start with the interaction of Glu69 with the Mg^{2+} ions at M1 and M3. A nucleophilic water molecule coordinated by these metal ions and activated by a conserved threonine at position 123 then carries out an in-line attack on the phosphorus moiety of PAP, leading in turn to the formation of a trigonal bipyramidal transition state. This state is stabilized by another Mg^{2+} at M2 (Na^+ is placed at M2 in the current PAP phosphatase-1 structure), which disperses the developed negative charge on the phosphorus group. Hydrolysis of the substrate then leads to inversion of the phosphate configuration at the active site, thereby releasing the product (Supplementary Fig. S7).

4.2. PAP phosphatase-1 is sensitive to Li^+ and insensitive to Na^+

Phosphatases of the IMPase superfamily, to which PAP phosphatase-1 belongs, are universally affected by submillimolar concentrations of Li^+. Depending upon lithium sensitivity, *i.e.* the IC_{50} for Li^+ inhibition, the superfamily is divided into three classes: (i) slightly sensitive, (ii) moderately sensitive and (iii) highly sensitive (Table 3). In the current studies, PAP phosphatase-1 was shown to be highly sensitive to Li^+, being inhibited at submillimolar levels (IC_{50} LiCl = 1 m*M*), but was insensitive to Na^+ up to relatively high concentrations (Fig. 5), a property similar to its animal homologue (López-Coronado *et al.*, 1999). As noted above, Li^+, being small and having only two electrons, is a poor scatterer of X-rays. It is thus difficult to obtain information about the molecular mechanism of Li^+ inhibition of a protein by X-ray crystallography, and to date such information is not available. We therefore propose a mechanism for the inhibition of PAP phosphatase-1 based on previous studies and indirect evidence. Hal2p is inhibited by and sensitive to both Li^+ and Na^+ ions. It was proposed that the formation of salt bridges between Arg152, Asp 263 and Glu238 causes the displacement of a helix (Thr147–Arg152), enabling the formation of a hydrogen-bond network between Asp145, Asp142, Asp294 and Gln265. This changes the binding geometry and physical size of the metal-binding site at M2, which influences the sensitivity of Hal2P towards Li^+ and Na^+ (Fig. 6*a*; Albert *et al.*, 2000). If M2 is large, it would accommodate both Li^+ and Na^+, the enzyme would be sensitive to both ions, and Na^+ and Li^+ at M2 would produce a dead-end complex with AMP (Albert *et al.*, 2000). If M2, on the other hand, is smaller than a certain value such that only the smaller Li^+ is able to enter and coordinate stably with the amino-acid residues, the enzyme would be sensitive to Li^+ and not to Na^+. The set of salt bridges between Arg152, Asp263 and Glu200 was proposed to be absent in the sodium-sensitive enzymes (Albert *et al.*, 2000). Interestingly, these charged/polar amino-acid residues

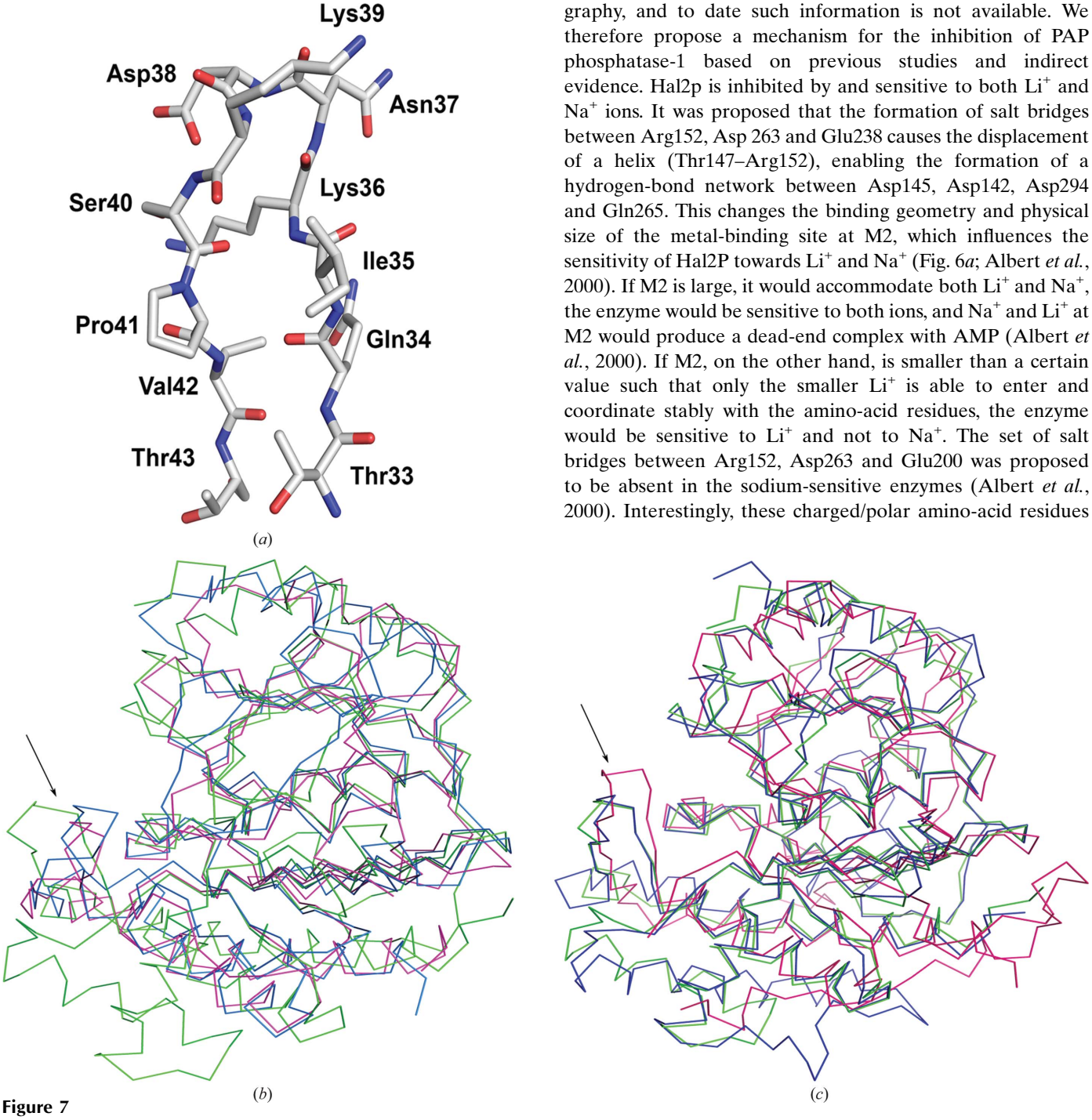

Figure 7
Analysis of the mobile-loop conformation across the IMPase superfamily. (*a*) Residues Thr33–Thr43 constitute the active mobile loop in PAP phosphatase-1. (*b*) Superposition of PAP phosphatase-1 (green) with homologous Li^+-insensitive enzymes from *A. fulgidus* and *M. jannaschii* (magenta and blue, respectively). (*c*) Superposition of PAP phosphatase-1 (green) with homologous Li^+-sensitive enzymes from yeast and rat (blue and pink, respectively). The conformation of the loop is shown with an arrow.

are conserved in Hal2P and PAP phosphatase-1 and are also known to be essential for increasing the thermostability of these enzymes (Albert *et al.*, 2000; Goldman, 1995). In PAP phosphatase-1 they are numbered Arg128, Asp226, Glu200, Asp257, Asp121, Asp118 and Gln228, and make interactions similar to those in Hal2p to generate a metal-binding site at M2 that is large enough to accommodate both Li^+ and Na^+ (Fig. 6*b*). On the basis of this argument, we might have expected PAP phosphatase-1 to be inhibited by both ions, but the enzyme was found to be insensitive to sodium. Moreover, a re-inspection of the physical size of M2 shows that it is very similar in the sodium-sensitive Hal2p from yeast and its sodium-insensitive homologues from human and rat. These comparisons, together with the findings for PAP phosphatase-1, appear to reveal a flaw in the hypothesis that the physical size of M2 in the IMPase superfamily determines the sensitivity towards Na^+. In PAP phosphatase-1 the mode of Li^+ inhibition and insensitivity towards Na^+ could be better explained by an alternative hypothesis based on three-metal-assisted catalysis and the involvement of a mobile catalytic loop in binding and localization of metal ions at M3 (Patel *et al.*, 2002; Johnson *et al.*, 2001; Bhattacharyya *et al.*, 2012; Stieglitz *et al.*, 2002). This site is a binding site for both Mg^{2+} and Li^+, and owing to their similar ionic radii they compete with each other for this site. According to this hypothesis, the size and conformation of this loop determine the sensitivity of the members of IMPase superfamily towards Li^+ (Johnson *et al.*, 2001; Stieglitz *et al.*, 2002). The conformation of the mobile loop also determines the affinity of Mg^{2+} for M3. In enzymes that are insensitive to Li^+ (*e.g.* IMPase/FBPase from *Archaeoglobus fulgidus* and *Methanocaldococcus jannaschii*), the loop is short and is situated relatively near the active site; this increases the affinity of Mg^{2+} for M3, and Li^+ cannot in this case compete easily with Mg^{2+} at this site. In contrast, in those enzymes that are lithium sensitive, the loop is longer, more mobile and situated away from the active site (*e.g.* Hal2p from yeast and RnPIP from rat); here, the affinity of Mg^{2+} for M3 is low and Mg^{2+} can easily be replaced by Li^+. PAP phosphatase-1 was superposed on these structures and the length and conformation of the equivalent mobile loop (Thr33–Thr43; Fig. 7*a*) was compared and found to be similar to those in Hal2p and RnPIP. On this basis, PAP phosphatase-1 would be predicted to be Li^+ sensitive (Figs. 7*b* and 7*c*). According to our enzyme-kinetics studies as described above, submillimolar concentrations of Li^+ do inhibit PAP phosphatase-1 activity, confirming the hypothesis that this mobile catalytic loop determines the sensitivity towards Li^+ across the IMPase superfamily (Johnson *et al.*, 2001; Stieglitz *et al.*, 2002). The size of the sodium ion with respect to M3 does help to explain the sodium insensitivity of PAP phosphatase-1. Note that in this enzyme the binding of Mg^{2+} to M3 creates an appropriate environment for the other Mg^{2+} ion at site M1 to generate a water-based nucleophile for in-line attack on the phosphorus moiety of the substrate, leading to hydrolysis. Li^+ can structurally substitute for Mg^{2+} at M3, but since it possesses little charge it cannot generate the nucleophilic water molecule essential for catalysis and thus the enzyme is inhibited. Na^+, being larger than both Mg^{2+} and Li^+, cannot enter M3 and hence cannot affect the activity of the enzyme, rendering the enzyme insensitive to Na^+.

5. Conclusion

The sulfate-activation pathway is crucial for the survival of *E. histolytica*. Accumulation of PAP, however, inhibits sulfate assimilation and RNA processing. PAP therefore needs to be hydrolyzed for proper growth. The enzyme 3′(2′)5′-bisphosphate nucleotidase (PAP phosphatase-1) showed dual substrate specificity by hydrolyzing PAP to AMP and phosphate ion and hydrolyzing Ins(1,4)P_2 to myoinositol 1-phosphate and phosphate ion. In the present work, we have reported the crystal structure and characterized the functions of PAP phosphatase-1 from *E. histolytica*. This is the first structure of any member of the IMPase superfamily from a protozoan to be determined. The structural findings have been complemented by kinetics studies. PAP phosphatase-1 is found to be active over wide ranges of temperature, pH and divalent metal ions. The involvement of large ion-pair networks imparting thermostability to the molecule is evident in the structure. Immunofluorescence studies showed this enzyme to be widely spread throughout the cytoplasm of *E. histolytica*. Interestingly, the previously proposed mechanism of Li^+/Na^+ sensitivity based on the size and the geometry of the metal-binding site M2 does not hold true in the case of PAP phosphatase-1. The cationic sensitivity may be better explained by open and closed conformations of the catalytic loop. Comparison of the mobile loop in PAP phosphatase-1 with those of different homologous structures of the IMPase superfamily clearly placed it in the group of Li^+-sensitive enzymes. The number of metal ions and their binding geometry at the active site is highly dependent upon the concentration of metal ions used during crystallization and the availability of substrate or product. To better understand the catalytic mechanism of action of PAP phosphatase-1, the involvement of three metal ions in catalysis has been proposed. In addition, the inositol-polyphosphate 1-phosphatase activity of PAP phosphatase-1 might play a role in the phosphatidylinositol signalling pathway and its connectivity to sulfur metabolism has yet to be explored.

6. Related literature

The following references are cited in the Supporting Information for this article: Gouet *et al.* (2003) and McWilliam *et al.* (2013).

We thank the Department of Biotechnology, Government of India for funding. KFT and SAAR thanks the CSIR for fellowship. We thank Dr Manish Kumar at Advanced Instrumentation Research Facility (AIRF), Jawaharlal Nehru University (JNU) for helping in data collection.

References

Aggarwal, M., Bansal, P. K. & Mondal, A. K. (2005). *Yeast*, **22**, 457–470.

Albert, A., Yenush, L., Gil-Mascarell, M. R., Rodriguez, P. L., Patel, S., Martínez-Ripoll, M., Blundell, T. L. & Serrano, R. (2000). *J. Mol. Biol.* **295**, 927–938.

Anwar, T. & Gourinath, S. (2013). *PLoS One*, **8**, e78714.

Baykov, A. A., Evtushenko, O. A. & Avaeva, S. M. (1988). *Anal. Biochem.* **171**, 266–270.

Bhattacharyya, S., Dutta, D., Ghosh, A. K. & Das, A. K. (2011). *Acta Cryst.* F**67**, 471–474.

Bhattacharyya, S., Dutta, D., Saha, B., Ghosh, A. K. & Das, A. K. (2012). *Biochimie*, **94**, 879–890.

Bond, C. S. (2003). *Bioinformatics*, **19**, 311–312.

Bone, R., Springer, J. P. & Atack, J. R. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 10031–10035.

Bradley, M. E., Rest, J. S., Li, W.-H. & Schwartz, N. B. (2009). *J. Mol. Evol.* **68**, 1–13.

Brinkac, L. M. *et al.* (2010). *Nucleic Acids Res.* **38**, D408–D414.

Chinthalapudi, K., Kumar, M., Kumar, S., Jain, S., Alam, N. & Gourinath, S. (2008). *Proteins*, **72**, 1222–1232.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Fukuda, C., Kawai, S. & Murata, K. (2007). *Appl. Environ. Microbiol.* **73**, 5447–5452.

Gill, R., Mohammed, F., Badyal, R., Coates, L., Erskine, P., Thompson, D., Cooper, J., Gore, M. & Wood, S. (2005). *Acta Cryst.* D**61**, 545–555.

Goldman, A. (1995). *Structure*, **3**, 1277–1279.

Gouet, P., Robert, X. & Courcelle, E. (2003). *Nucleic Acids Res.* **31**, 3320–3323.

Hatzios, S. K., Iavarone, A. T. & Bertozzi, C. R. (2008). *Biochemistry*, **47**, 5823–5831.

Inhorn, R. C. & Majerus, P. W. (1988). *J. Biol. Chem.* **263**, 14559–14565.

Johnson, K. A., Chen, L., Yang, H., Roberts, M. F. & Stec, B. (2001). *Biochemistry*, **40**, 618–630.

Joosten, R. P., Joosten, K., Murshudov, G. N. & Perrakis, A. (2012). *Acta Cryst.* D**68**, 484–496.

Ke, H., Thorpe, C. M., Seaton, B. A., Marcus, F. & Lipscomb, W. N. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 1475–1479.

Kumar, S., Raj, I., Nagpal, I., Subbarao, N. & Gourinath, S. (2011). *J. Biol. Chem.* **286**, 12533–12541.

Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.

Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). *Nature Protoc.* **3**, 1171–1179.

Laskowski, R. A., Moss, D. S. & Thornton, J. M. (1993). *J. Mol. Biol.* **231**, 1049–1067.

Laskowski, R. A. & Swindells, M. B. (2011). *J. Chem. Inf. Model.* **51**, 2778–2786.

López-Coronado, J. M., Bellés, J. M., Lesage, F., Serrano, R. & Rodríguez, P. L. (1999). *J. Biol. Chem.* **274**, 16034–16039.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., Cowley, A. P. & Lopez, R. (2013). *Nucleic Acids Res.* **41**, W597–W600.

Mehlotra, R. K. (1996). *Crit. Rev. Microbiol.* **22**, 295–314.

Mi-ichi, F., Abu Yousuf, M., Nakada-Tsukui, K. & Nozaki, T. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 21731–21736.

Mi-ichi, F., Makiuchi, T., Furukawa, A., Sato, D. & Nozaki, T. (2011). *PLoS Negl. Trop. Dis.* **5**, e1263.

Murguía, J. R., Bellés, J. M. & Serrano, R. (1995). *Science*, **267**, 232–234.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Neuwald, A. F., Krishnan, B. R., Brikun, I., Kulakauskas, S., Suziedelis, K., Tomcsanyi, T., Leyh, T. S. & Berg, D. E. (1992). *J. Bacteriol.* **174**, 415–425.

Nozaki, T., Arase, T., Shigeta, Y., Asai, T., Leustek, T. & Takeuchi, T. (1998). *Biochim. Biophys. Acta*, **1429**, 284–291.

Patel, S., Yenush, L., Rodríguez, P. L., Serrano, R. & Blundell, T. L. (2002). *J. Mol. Biol.* **315**, 677–685.

Pollack, S. J., Atack, J. R., Knowles, M. R., McAllister, G., Ragan, C. I., Baker, R., Fletcher, S. R., Iversen, L. L. & Broughton, H. B. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 5766–5770.

Quintero, F. J., Garciadeblás, B. & Rodríguez-Navarro, A. (1996). *Plant Cell*, **8**, 529–537.

Raj, I., Kumar, S. & Gourinath, S. (2012). *Acta Cryst.* D**68**, 909–919.

Ramaswamy, S. G. & Jakoby, W. B. (1987). *J. Biol. Chem.* **262**, 10044–10047.

Schneidman-Duhovny, D., Inbar, Y., Nussinov, R. & Wolfson, H. J. (2005). *Nucleic Acids Res.* **33**, W363–W367.

Spiegelberg, B. D., Xiong, J.-P., Smith, J. J., Gu, R. F. & York, J. D. (1999). *J. Biol. Chem.* **274**, 13619–13628.

Stec, B., Yang, H., Johnson, K. A., Chen, L. & Roberts, M. F. (2000). *Nature Struct. Biol.* **7**, 1046–1050.

Stieglitz, K. A., Johnson, K. A., Yang, H., Roberts, M. F., Seaton, B. A., Head, J. F. & Stec, B. (2002). *J. Biol. Chem.* **277**, 22863–22874.

Thomas, D., Barbey, R., Henry, D. & Surdin-Kerjan, Y. (1992). *J. Gen. Microbiol.* **138**, 2021–2028.

Toledano, E., Ogryzko, V., Danchin, A., Ladant, D. & Mechold, U. (2012). *Biochem. J.* **443**, 485–490.

Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D. & Hochstrasser, D. F. (1999). *Methods Mol. Biol.* **112**, 531–552.

Xue, Y., Huang, S., Liang, J.-Y., Zhang, Y. & Lipscomb, W. N. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 12482–12486.

Ye, Y. & Godzik, A. (2004). *Nucleic Acids Res.* **32**, W582–W585.

Yenush, L., Bellés, J. M., López-Coronado, J. M., Gil-Mascarell, R., Serrano, R. & Rodríguez, P. L. (2000). *FEBS Lett.* **467**, 321–325.

York, J. D., Ponder, J. W., Chen, Z., Mathews, F. S. & Majerus, P. W. (1994). *Biochemistry*, **33**, 13164–13171.

York, J. D., Ponder, J. W. & Majerus, P. W. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 5149–5153.

Zhang, J. & Biswas, I. (2009). *J. Bacteriol.* **191**, 4330–4340.

Zhang, J.-Y., Zou, J., Bao, Q., Chen, W.-L., Wang, L., Yang, H. & Zhang, C.-C. (2006). *Appl. Environ. Microbiol.* **72**, 245–251.

Zhang, R., Villeret, V., Lipscomb, W. N. & Fromm, H. J. (1996). *Biochemistry*, **35**, 3038–3043.