Crystal structure of mammalian purple acid phosphatase

Crystal structure of mammalian purple acid phosphatase

Luke W Guddat[1]*, Alan S McAlpine[1], David Hume[1], Susan Hamilton[1], John de Jersey[1] and Jennifer L Martin[2]*

Background: Mammalian purple acid phosphatases are highly conserved binuclear metal-containing enzymes produced by osteoclasts, the cells that resorb bone. The enzyme is a target for drug design because there is strong evidence that it is involved in bone resorption.

Results: The 1.55 Å resolution structure of pig purple acid phosphatase has been solved by multiple isomorphous replacement. The enzyme comprises two sandwiched β sheets flanked by α-helical segments. The molecule shows internal symmetry, with the metal ions bound at the interface between the two halves.

Conclusions: Despite less than 15% sequence identity, the protein fold resembles that of the catalytic domain of plant purple acid phosphatase and some serine/threonine protein phosphatases. The active-site regions of the mammalian and plant purple acid phosphatases differ significantly, however. The internal symmetry suggests that the binuclear centre evolved as a result of the combination of mononuclear ancestors. The structure of the mammalian enzyme provides a basis for antiosteoporotic drug design.

Addresses: [1]Department of Biochemistry, The University of Queensland, St Lucia, Qld 4072, Australia and [2]Centre for Drug Design and Development, The University of Queensland, St Lucia, Qld 4072, Australia.

*Corresponding authors.
E-mail: guddat@biosci.uq.edu.au
j.martin@mailbox.uq.edu.au

Key words: metalloenzyme, protein structure, purple acid phosphatase, tartrate-resistant acid phosphatase, uteroferrin

Received: 1 March 1999
Revisions requested: 29 March 1999
Revisions received: 19 April 1999
Accepted: 19 April 1999

Published: 22 June 1999

Structure July 1999, 7:757–767
http://biomednet.com/elecref/0969212600700757

 ISSN 0969-2126

Introduction

Purple acid phosphatases (PAPs) are a group of nonspecific phosphomonoesterases, members of which have been characterised from animal, plant and fungal sources (for a recent review see [1]). The enzymes are distinguished from other acid phosphatases by their characteristic purple colour in concentrated solution. The animal enzymes are also distinguished from lysosomal and prostatic acid phosphatases by their resistance to inhibition by L(+)tartrate, and are commonly referred to as 'tartrate-resistant acid phosphatases' (TRAPs). TRAP activity is the most commonly used histochemical marker for osteoclasts, the cells that resorb bone. The TRAP gene is expressed at highest levels in these cells and is required for normal bone resorption. A transgenic knockout mouse lacking the enzyme is mildly osteopetrotic and also has defects in endochondral ossification [2]. We have shown that both pyrophosphate and phosphotyrosine-containing peptides are excellent substrates for mammalian PAPs, especially at the low pH of the osteoclastic space [3]. Pyrophosphate is a known inhibitor of bone resorption and the phosphotyrosine-containing protein osteopontin has been implicated in the attachment of osteoclasts to bone. Experiments *in vitro* have shown that when osteopontin is partially dephosphorylated by PAP it no longer binds osteoclasts [4]. Determination of the structure of mammalian PAP is therefore important for probing the substrate specificity of the enzyme and ultimately in designing potent and selective inhibitors that may prove useful in the treatment of bone resorptive diseases.

All of the PAPs characterised to date contain a binuclear metal centre (FeIII–MeII). A charge-transfer interaction between a tyrosine and the FeIII atom is responsible for the visible absorption. Mammalian PAPs have been characterised from human [3,5], pig [6], rat [7], mouse [3] and cow [6]. The enzymes are closely related: all are monomeric glycoproteins of ~35 kDa with amino acid sequences exhibiting >80% identity [7–9]. The active form of mammalian PAP is pink (λ_{max} = 550 nm), and contains a binuclear FeIII–FeII metal centre in which the iron atoms are antiferromagnetically coupled [10,11]. Oxidation converts the enzyme to the inactive, purple FeIII–FeIII form (λ_{max} ~550 nm) [12]. In addition to its phosphatase activity, the redox-active FeIII–FeII form of the mammalian enzyme can act as a catalyst of hydroxyl radical production in a Fenton-like reaction [5,13], using hydrogen peroxide and a reducing agent such as superoxide as substrates:

$$\text{Enz–FeII–FeIII} + H_2O_2 \rightarrow \text{Enz–FeIII–FeIII} + {}^{\cdot}OH + {}^{-}OH$$

$$\text{Enz–FeIII–FeIII} + O_2^{\cdot -} \rightarrow \text{Enz–FeII–FeIII} + O_2$$

Both osteoclasts and macrophages produce hydroxyl radicals, and whereas the Fenton catalyst has not been identified in either cell type, PAP is a candidate.

Members of another class of phosphatase, the serine/threonine protein phosphatases that include protein phosphatases

1 and 2B (calcineurin), also contain binuclear metal centres and appear to be related structurally to the PAPs [14]. A conserved sequence motif which encompasses the metal ligands in the protein phosphatases and the PAPs has been designated a 'metalloesterase signature motif' and has been identified in a range of other less well characterised phosphoesterases that are presumed to contain binuclear metal centres [15]. These metal-containing phosphatases are in turn part of a much larger and more diverse family of proteins which contain binuclear metal centres. Functions for these proteins include oxygen transport and activation, catalysis of isomerisation, phosphoryl transfer, hydration and a broad range of hydrolytic reactions [16,17].

Until now, the only available structure for a PAP has been that of the red kidney bean enzyme, determined at 2.65 Å resolution by Klabunde and colleagues [15,18]. In contrast to the monomeric 35 kDa mammalian enzyme, the plant PAP is a homodimer with subunits of ~55 kDa. Each subunit of the plant enzyme incorporates a binuclear FeIII–ZnII centre. Despite overall low sequence identity between the plant and animal PAPs (<20% within the catalytic domains), the amino acid residues which provide the metal ligands to the iron (Tyr167, His325 and Asp135) and zinc (His286, His323 and Asn201) and the bridging ligand (Asp164) (red kidney bean numbering) are conserved across all known members of the PAP family [1]. Furthermore, substitution of zinc with FeII in the red kidney bean enzyme active site generates a redox-active form of the enzyme with spectroscopic properties similar to those of the mammalian enzymes [19]. A tentative model for the mammalian enzyme was developed on the basis of secondary structure prediction and the crystal structure of the red kidney bean enzyme [20].

The most abundant source of the mammalian enzyme is pig allantoic fluid where the protein may have a novel function in iron transport [21]. We now report the experimentally determined three-dimensional structure of the pig allantoic fluid enzyme (uteroferrin) at 1.55 Å resolution. The enzyme is comprised of two sandwiched β sheets flanked by α-helical segments. The molecule shows internal symmetry, with the metal ions bound at the interface between the two halves. Despite less than 15% sequence identity, the protein fold resembles that of the catalytic domain of plant PAP and of some serine/threonine protein phosphatases.

Results and discussion

Structure of mammalian purple acid phosphatase

Previous studies in our laboratory showed that the addition of phosphate to reduced pig allantoic fluid purple acid phosphatase (PPAP) led to rapid oxidation of the enzyme

Table 1

Statistics for data collection and heavy-atom phasing.

	Native*	Native†	K_2Cl_6Os	K_2PtCl_4	$K_2Pt(NO)_4$	$K_2Pt(NO)_4$	K_2Cl_4Au	$HgCl_2$
Resolution (Å)	1.55	2.00	2.50	3.00	2.50	2.50	4.00	2.70
Resolution for outer shell (Å)	1.61–1.55	2.07–2.00	2.59–2.50	3.11–3.00	2.59–2.50	2.59–2.50	4.14–4.00	2.80–2.70
Total observations (I > 0σ(I))	140,370	69,826	40,174	20,921	51,966	25,367	5479	16,104
Unique reflections (I > 0σ(I))	46,161	24,836	12,179	7410	14,276	9047	2334	8978
Completeness (%)	89.4	96.4	91.9	94.7	96.0	68.6	68.9	82.1
Completeness for outer shell (%)	57.4	94.0	69.9	94.1	93.0	53.1	73.9	73.0
I/σ(I)	9.6	9.4	10.6	5.8	7.4	9.7	5.8	6.7
I/σ(I) for outer shell (%)	2.1	2.6	3.1	3.0	3.6	2.9	4.5	2.4
R_{merge} (%)‡	5.9	6.4	6.8	12.9	9.4	7.2	11.9	10.5
R_{merge} for outer shell (%)	27.1	31.3	22.7	26.8	27.1	22.7	19.6	27.8
Soak time (days)			2	1	1	2	1	7
Soak concentration (mM)			33	3	3	33	3	33
R_{iso} (%)§			11.1	10.5	9.1	15.9	13.1	17.9
Maximum resolution for phasing (Å)			2.8	3.0	3.5	2.8	4.5	4.5
R_{Cullis}¶			0.598	0.591	0.578	0.570	0.585	0.562
R_{Kraut}¥			0.169	0.146	0.109	0.220	0.173	0.247
Phasing power**			1.40	1.37	1.33	1.56	1.45	1.52
FOM			0.268	0.311	0.322	0.287	0.301	0.283
Number of heavy-atom sites			4	4	3	6	7	5
Overall FOM	0.475							
FOM after solvent flattening	0.809							

*Native data collected at 100K. †Native data and all heavy-atom data collected at room temperature. ‡$R_{merge} = \Sigma | I_i - < I > | / \Sigma I_i$.
§$R_{iso} = \Sigma | F_{PH} - Fp | / \Sigma | F_P |$.
¶$R_{Cullis} = \Sigma |\,|F_{PH}|_{obs} \pm |F_P|_{obs}| - |F_H|_{calc}| / \Sigma\,|F_{PH}|_{obs} \pm F_P|_{calc}|$, where F_P, F_{PH} and F_H are the native structure factor, the protein and heavy-atom structure factor, and the heavy-atom structure factor, respectively. ¥$R_{Kraut} = \Sigma | \,|F_{PH}|_{obs} - |F_P|_{calc}| / \Sigma |F_{PH}|_{obs}$. **Phasing power = $\Sigma < F_H > /E$, where $< F_H >$ is the rms heavy-atom structure factor and E is the residual lack of closure. FOM, figure of merit.

Table 2

Refinement statistics.

Resolution range (Å)	100–1.55
Reflections F > 1σ(F))	43,833
Protein atoms	2371
Carbohydrate atoms	28
Fe atoms	2
Solvent atoms	340
R factor (%)*	0.213
R_{free} (%)*	0.257
Rmsd bond lengths (Å)	0.007
Rmsd bond angles (°)	1.31
Average B factor (Å^2)	20.0

*$R = \Sigma||F_{obs}|-|F_{calc}||/\Sigma|F_{obs}|$, where the R factor is calculated based on the reflections used in the refinement (90% of the total data), and R_{free} [50] is calculated using the remaining 10% of the data.

and formation of a very stable oxidised enzyme–phosphate complex [22]. Crystals of this complex have been grown by hanging-drop vapour diffusion and microseeding. Native crystals, cryocooled to 100K, diffracted to 1.55 Å. Six heavy-atom derivatives (Table 1) were used to solve the structure, which has an R factor of 0.213 and R_{free} of 0.257 (Table 2). The final model includes residues 3–304 of the 313-residue mature protein, two iron atoms, the bridging oxygen atom, the bound phosphate, the first two sugar residues of the single oligosaccharide linked to Asn97, 339 water molecules and one isopropanol molecule (Figure 1). An example of the quality of the final electron density is shown in Figure 2.

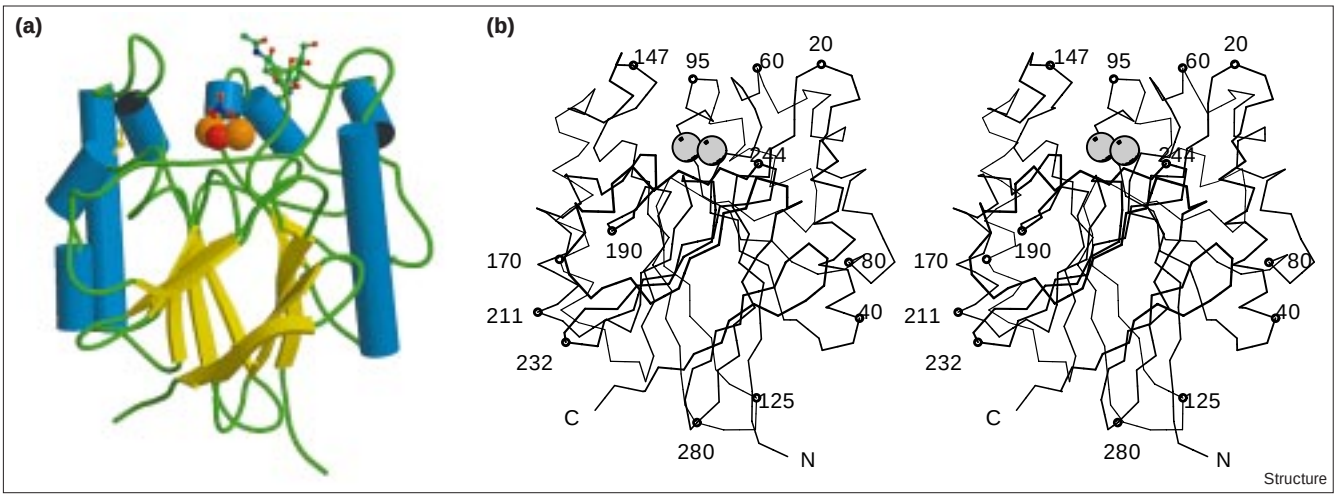
The structure is approximately spherical with dimensions of 45 × 45 × 40 Å. The core structure consists of two mixed β sheets packed together in a β sandwich which is flanked on either side by α-helical segments in a four-layer αββα sandwich architecture. Loops connecting the secondary structure elements provide all but one of the sidechain ligands for the binuclear metal centre. A disulphide bond with a left-hand spiral conformation links residues 142 and 200.

Figure 3 compares the three-dimensional structures of the pig enzyme and the red kidney bean enzyme (RKPAP) and provides a structure-based sequence alignment. As predicted by Klabunde *et al.* [20] the core Cα structures are very similar. The root mean square deviation (rmsd) for Cα atoms of 204 core residues (shown in green in Figure 3) is 1.4 Å. Apart from the N-terminal domain (blue; Figure 3) which is absent in the mammalian enzyme, there are three regions in which significant structural differences are observed (magenta; Figure 3): the dimer interface of the plant enzyme; the surface of the mammalian enzyme exposed by the absence of the N-terminal domain; and variable loops in the vicinity of the active site (see below).

Internal symmetry

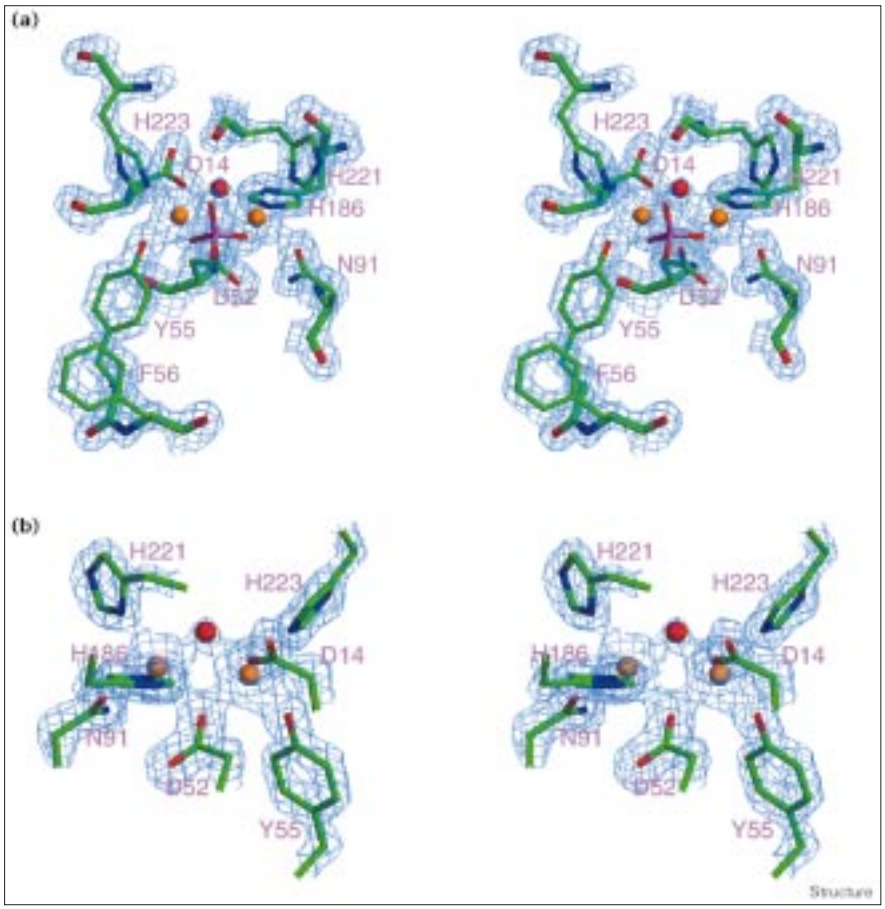
A striking feature of the PPAP structure is the pseudo-twofold symmetry, with the binuclear metal centre located on the twofold axis (Figure 4). A common structural motif (purple elements in Figure 4) is present in each half (designated motifs A and B). This motif comprises a core

Figure 1



Structure of pig purple acid phosphatase. (a) Schematic representation of the oxidised pig purple acid phosphatase–phosphate complex, showing the approximate twofold symmetry of the molecule. Helices are drawn as blue cylinders and strands as yellow arrows. Turns and random coils are shown in green. The Fe atoms are drawn as orange spheres and the bridging oxygen as a red sphere. The disulphide bond (in the top left corner), two carbohydrate residues N-linked to Asn97, and the phosphate molecule is drawn in ball-and-stick representation. (b) Stereoview Cα trace of pig purple acid phosphatase. The figure is in the same orientation as (a). The Fe atoms are identified by spheres.

Figure 2



Electron density at the active site of pig purple acid phosphatase. (a) Stereoview of the active site of the oxidised enzyme–phosphate complex, showing the Fe atoms, bridging oxygen, phosphate and metal-binding ligands. The *cis* peptide bond can also be seen between Tyr55 and Phe56. (b) Alternative view of the active site rotated by 180° around the y axis, as compared with (a), and with the phosphate removed to highlight the bridging oxygen and carboxylate groups. The $2F_o–2F_c$ electron density, contoured at 2.0σ, is overlaid in both images.

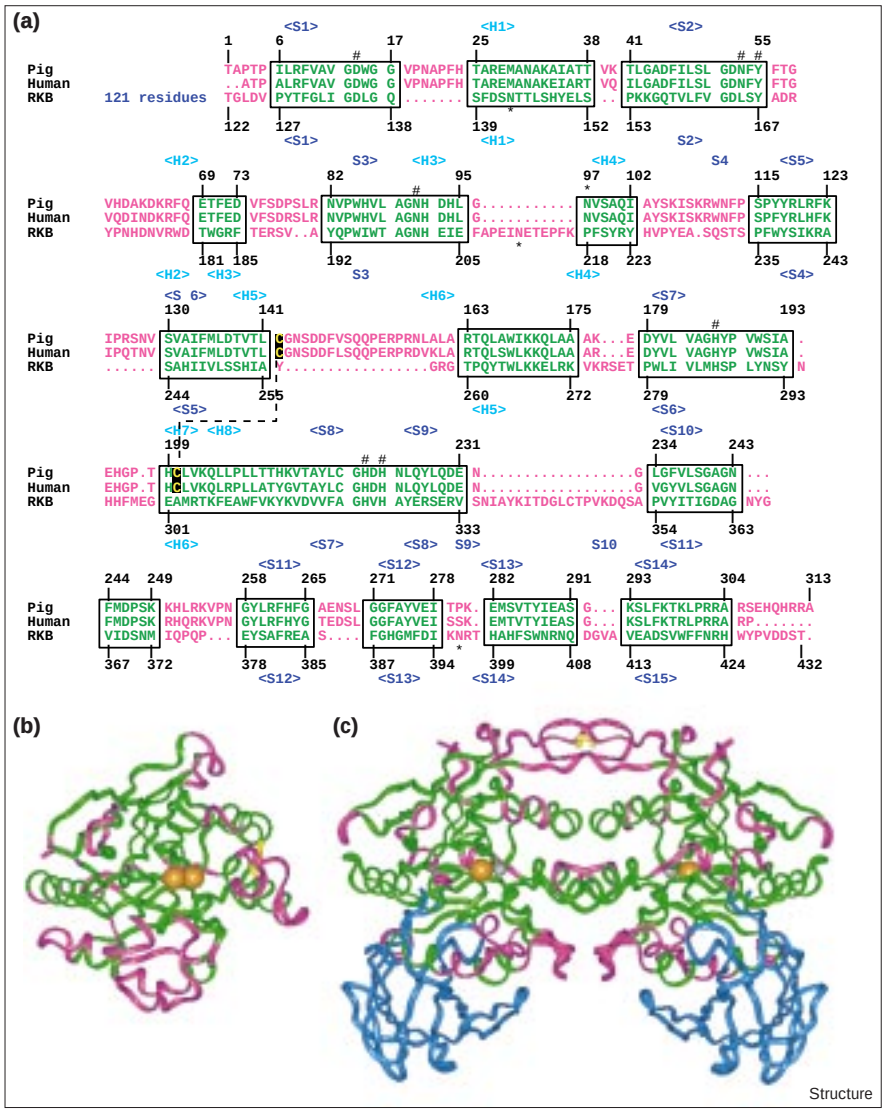
βαβαβ unit forming a parallel β sheet, plus two additional β strands, one on either side of the core β sheet. The Cα atoms of the 56 residues from each of the two motifs can be superimposed with an rmsd of 1.72 Å.

The finding of internal structural similarity in PAP suggests that the enzyme may have evolved by gene amplification. Motifs A and B were therefore analysed to identify whether there is any residual sequence homology. The only significant similarity is found in a short stretch of sequence following the fourth β strand in the motif, 91-**NH**DH-94 (single-letter amino acid code) in motif A and 220-G**H**D**H**-223 in motif B. This is nonetheless an important similarity as both of these sequences contribute two metal- or ligand-binding residues (in boldface type; also see below). Furthermore, the topologically equivalent loops between strand S2 and helix H2 of motif A (51-G**D**NF**Y**FTGVHDAK-63) and between strand S7 and helix H7 of motif B (185-G**H**YPVWSIAE**H**GP-197) each provide two metal- or ligand-binding residues. In this latter case, ligand-binding function appears to be conserved between the two motifs, but structural and sequence homology is not significant. In the pig gene the majority of the coding region is on a single exon, but in mouse and human genes an intron after the Trp111 codon is close to the boundary separating the two halves of the molecule [23]. This observation supports the view that the two halves constitute 'genetic domains' brought together during evolution. In this structure each motif contributes ligands to both metal ions, so that some exchange of ligands between metal ions must also have occurred.

The symmetry within PPAP indicates the possibility of symmetry in the RKPAP monomer. Indeed, in the paper describing the structure of the plant enzyme it was noted that the C-terminal domain incorporated two sandwiched βαβαβ motifs forming the core unit of the enzyme [18]. However, the plant enzyme is more complex than mammalian PAP (see comparison of the two structures in Figure 3) and the symmetry is less evident. Nevertheless, it is possible to superimpose the Cα atoms of 51 residues in the equivalent motifs A and B of the plant enzyme with an rmsd of 1.82 Å.

We speculate that there may be additional examples of binuclear metal ion complexes in proteins which have

Figure 3



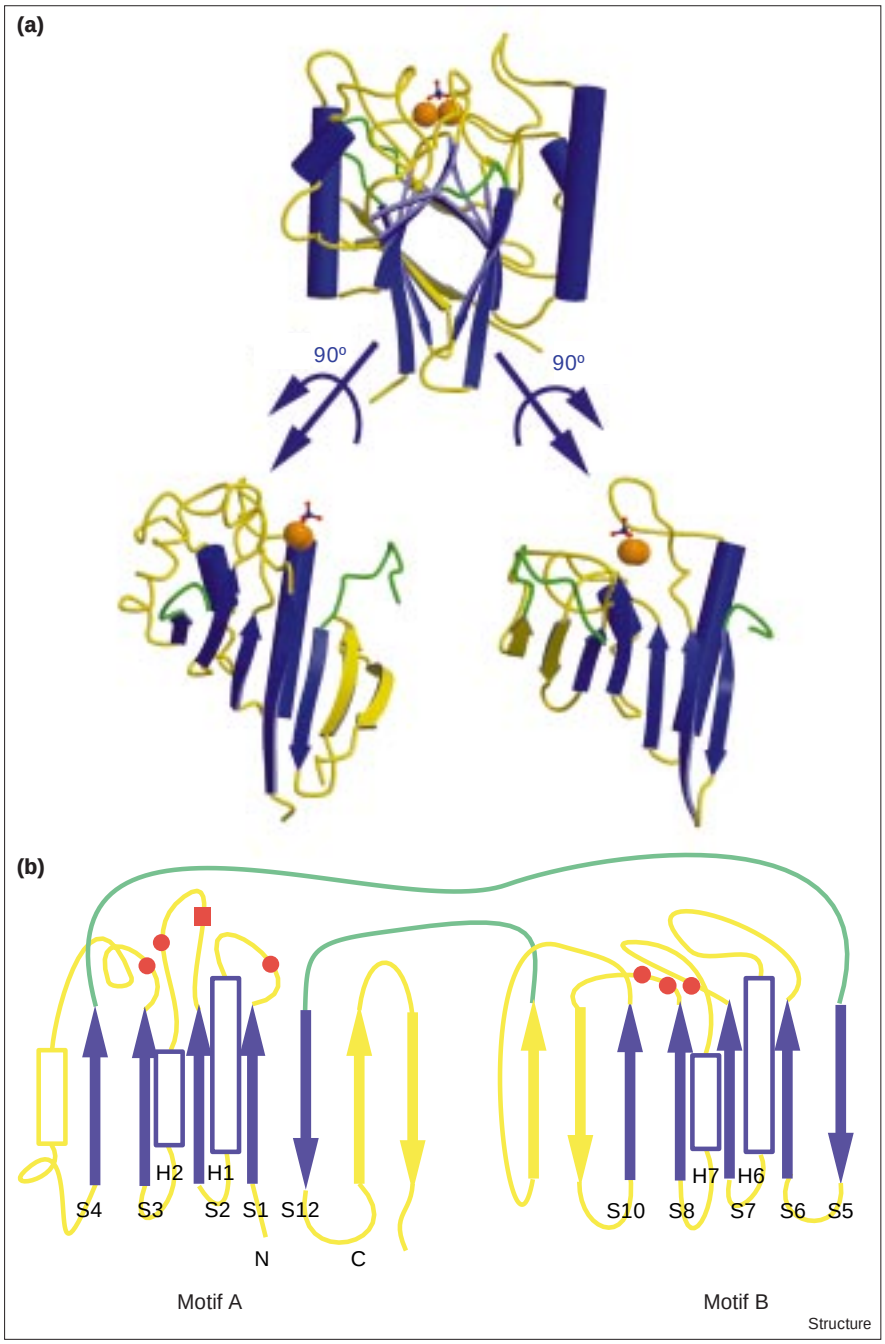
Comparison of pig and red kidney bean purple acid phosphatases. (a) Structure-based sequence alignment of red kidney bean (RKB) and pig purple acid phosphatase. The sequence of the human enzyme is also included and is aligned with the pig enzyme on the basis of sequence homology. Residues in green were aligned using the LSQ features in the program O [43]. Secondary structure elements are indicated and labelled H for helix and S for strand, and are shown above (pig) and below (red kidney bean) the sequence, respectively. These elements were defined using the program DSSP [51]. Glycosylation sites and metal-binding ligands are identified by an asterisk and hash symbol, respectively. The disulphide bond connecting residues 142 and 200 in the pig enzyme is depicted by a dashed line. Ribbon diagrams are also shown that compare the structures of (b) pig and (c) red kidney bean purple acid phosphatases. The Fe atoms are drawn as orange spheres and the Zn atoms as silver spheres. Disulphide bonds are in yellow. The Cα atoms are colour-coded such that the N-terminal domain in the red kidney bean structure is in blue, structurally conserved or 'core' residues in the two enzymes are in green, and the remaining residues are in magenta. (The same colouring is used for the sequence alignment above.)

evolved by combining protein domains for mononuclear complexes. An alternative pathway for the evolution of such centres has been proposed for the enzyme urease, which contains two nickel ions. The structure of the catalytic domain of urease is very similar to that of the mononuclear zinc enzyme adenosine deaminase. It has been suggested on this basis that urease has evolved from a single ancestral adenosine deaminase-like protein by way of mutation to provide additional ligands such as the bridging carbamylated lysine ligand [24]. Further analysis of the structures of other binuclear metal-containing proteins and their corresponding genes would allow comment on the generality of these mechanisms.

Active-site structure and catalytic mechanism

The structure of the active site of the enzyme–phosphate complex in the vicinity of the metal centre is shown in Figure 2 and the important bond distances are shown in Tables 3 and 4. The FeIII–FeIII distance is 3.31 Å, which is consistent with that found by extended X-ray absorption fine structure (EXAFS) studies (3.2–3.3 Å) [25,26] and the crystal structure of RKPAP (3.26 Å). The coordination spheres of both metals are close to octahedral. Phosphate bridges the two metals symmetrically. Strong electron density is observed for a bridging oxo/hydroxo group (Figure 2). The bridging oxygen is within hydrogen-bonding distance of the carbonyl oxygen of His221. In RKPAP the bridge could not be unambiguously located from the electron density but was modelled into a similar position to that found in the pig enzyme [15]. The modelled hydroxo group was within hydrogen-bonding distance of the carbonyl oxygen of His323, equivalent to His221 in the pig enzyme. The observation of the bridging solvent molecule in the oxidised enzyme–phosphate complex counters previous suggestions

Figure 4



Internal symmetry in the pig purple acid phosphatase structure. The pseudo-symmetry in the pig purple acid phosphatase structure is emphasised by depicting the secondary structure elements of the two motifs in purple. One half (left) incorporates residues from both the N terminus (1–115) and the C terminus (270–303) and the other half incorporates the consecutive residues 116–269. Each half comprises a seven-stranded mixed β sheet flanked on one surface by α helices and loops. The two loops that connect motifs A and B are in green and other structural elements are in yellow. The Fe atoms are depicted as orange spheres and the bound phosphate is shown in ball-and-stick representation. **(a)** The structure of the pig enzyme, with the pseudo-symmetry axis aligned approximately in the direction of the y axis. Shown below are the two halves of the structure opened out like the pages of a book, by rotating each half by 90° or –90° around the y axis. **(b)** The topology of the pig enzyme, again highlighting the pseudo-symmetry in the structure. The purple secondary structure elements within each motif are labelled as in Figure 3. The approximate position of the metal-binding ligands (red spheres) and the bridging ligand (red square) are also shown.

that phosphate might displace the bridging solvent molecule upon binding to the enzyme [27]. The iron–oxygen bond lengths and the Fe–O–Fe bond angle (2.08 Å, 105.7°) are consistent with values determined for the oxidised pig enzyme–phosphate complex by EXAFS (1.94 Å, 109°), and are more consistent with a bridging hydroxo than an oxo group based on a comparison with model oxo-bridged complexes [28]. A range of other spectroscopic evidence supports this view (reviewed in [1]).

A role for the bridging hydroxo group as the attacking nucleophile in the hydrolytic mechanism of PAP and the serine/threonine protein phosphatases has been suggested [14,27], and is supported by ^{18}O-labelling studies of a model complex in which a phosphate ester is coordinated to two Co(III) centres with bridging hydroxo groups [29]. Alternatively, Aquino *et al.* [30] have obtained kinetic data which favour a mechanism in which monodentate binding of phosphate esters to the divalent metal is followed by

Table 3

Metal–ligand distances.

Metal ion	Ligand atom	Distance (Å)
FeIII*	Asp14 Oδ2	2.11
	Asp52 Oδ2	2.27
	Tyr55 OH	1.98
	His223 Nε2	2.32
	μ-oxo O	2.08
	PO_4 O3	2.26
FeII*†	Asp52 Oδ2	2.40
	Asn91 Oδ1	2.24
	His186 Nε2	2.23
	His221 Nδ1	2.37
	μ-oxo O	2.08
	PO_4 O4	2.45

*The Fe–O–Fe angle is 105.7°. †The Fe atom that is FeII in the active form of the enzyme.

Table 4

Selected active-site distances.

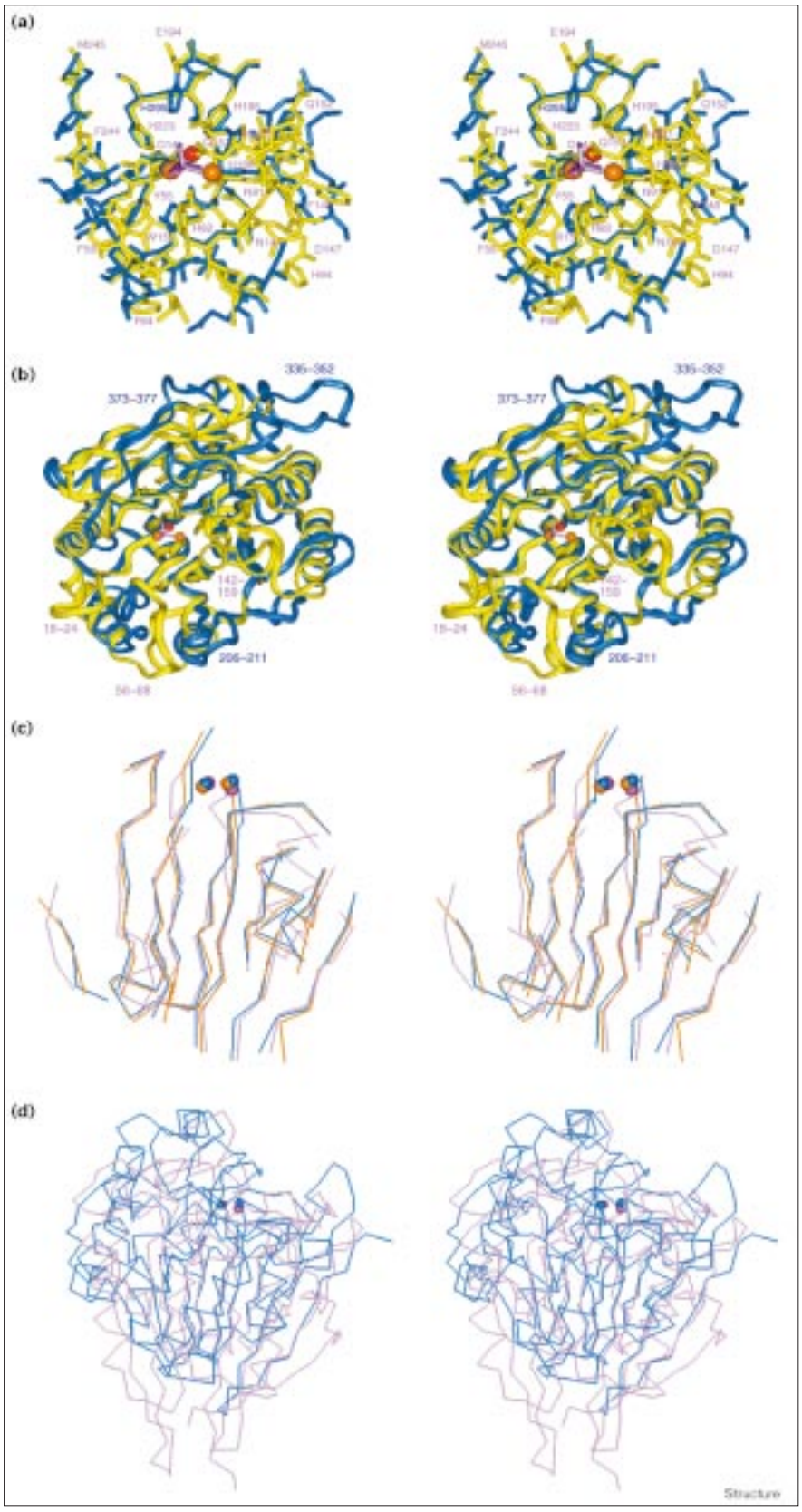
From	To	Distance (Å)
μ-oxo	Asp52 Oδ2	2.90
	His221 Nδ1	3.19
	His223 Nε2	3.13
	Asp14 Oδ2	2.77
	PO_4 O1	2.88
	PO_4 O3	2.68
	PO_4 O4	2.77
	His221 C=O	2.46
PO_4 O1	His195 Nε2	2.76
PO_4 O2	His92 Nε2	2.68
	PO_4 O4	2.45
PO_4 O3	His92 Nε2	3.17
	Asp52 Oδ2	3.04
	His223 Nε2	3.13
PO_4 O4	Asn91 Nδ2	2.73

attack by the hydroxide ligand coordinated to ferric ion. Furthermore, they propose that bridging of the phosphate ester between the metals results in a nonproductive complex. In the serine/threonine protein phosphatases, bidentate binding of the phosphate ester and a coordinated hydroxyl group may both be accommodated in the enzyme–substrate complex. However, in PAP, this coordinated hydroxyl group is replaced by the tyrosine sidechain responsible for the purple colour. Labelling of the bridging oxygen with ^{18}O should in principle allow a distinction between these mechanisms.

The spatial arrangement of the metal ligands in the pig and red kidney bean enzymes is very similar (Figure 5). The Cα atoms of the seven metal-binding residues superimpose with an rmsd of 0.18 Å. Two other residues in the vicinity of the metal centre, His92 and His195, which are hydrogen bonded to phosphate, also occupy equivalent positions in the two structures. As can be seen from Figure 5, however, these are the only sidechains which may be overlaid between the two proteins in the vicinity of the binuclear centre. A significant difference is that a third histidine (His295), which in the RKPAP interacts with phosphate, is replaced by glutamic acid in PPAP (Glu194) and is oriented away from the phosphate in the oxidised enzyme–phosphate complex. A further point of interest is that two of the metal-binding ligands are associated with unusual mainchain conformations: His221 (the carbonyl oxygen of which is within hydrogen-bonding distance of the bridging hydroxyl) is part of a γ turn and Tyr55 has a *cis* peptide bond to Phe56. The γ turn is also present in RKPAP but there is no evidence for a *cis* peptide bond (presumably because of the lower resolution). The polypeptide chain in this region of RKPAP forms a 3_{10} helix. These unusual conformations may be stabilised by the geometric requirements of the metal ion complex, and it will be interesting to see whether they are retained in the apoenzyme structure.

There are significant differences between the pig and red kidney bean enzymes beyond the immediate vicinity of the metal ions. The most obvious difference (Figure 5) is the loop (residues 142–159), which is an insert in the PPAP sequence compared with that of RKPAP (Figure 3). The loop is tethered to motif B through a disulphide bond. It has long been known that some mammalian PAPs can be cleaved within this loop by trypsin or α-chymotrypsin, yielding two disulphide-bonded fragments. The nicked form of the bovine enzyme has a k_{cat} value approximately fivefold higher than that of the intact enzyme for *p*-nitrophenyl phosphate, whereas the K_m values for the two forms are similar [31]. An increase in specific activity of the human, mouse and pig enzymes is also seen upon nicking [3,31]. As well as having an effect on the catalytic efficiency, it seems likely that a loop in this position could affect the specificity of the enzyme towards high molecular weight substrates such as phosphoproteins. There is another loop in the corresponding position in motif A (Figure 5; residues 18–24) which may also be involved in substrate binding. Other polar residues in the active-site region which have been tentatively implicated in substrate binding in RKPAP [18] are not conserved in PPAP. For example, Arg258 is replaced by Leu161, Glu299 and Tyr365 are deleted, and His253 is replaced by Val139. Additional differences in the substrate-binding sites of the two enzymes may exist if, as has been suggested, residues on one monomer of the dimeric RKPAP contribute to the active site on the opposing monomer (e.g. Lys306) [15,18]. Overall, this structural comparison shows that the specificities of these enzymes, especially for large substrates, are likely to be

Figure 5



Structural comparison of pig purple acid phosphatase, red kidney bean purple acid phosphatase and PPP-type serine/threonine protein phosphatases. (a,b) Stereoview diagrams showing the superimposition of pig (yellow) and red kidney bean (blue) purple acid phosphatases. (a) All-atom superimposition of the region in the vicinity of the active site. (b) Ribbon trace of the whole of the pig purple acid phosphatase superimposed onto the C-terminal domain of the red kidney bean enzyme. The bridging oxygen (red sphere) and Fe atoms (orange spheres) are shown for the pig enzyme. Phosphate is shown as a purple stick model and residue numbers are labelled in purple (pig) or blue (red kidney bean). Major differences occur at loops 142–159 and 18–24, which are insertions in pig purple acid phosphatase compared with the red kidney bean enzyme. (c) Stereoview diagram showing the superimposition of Cα atoms from a core region of the structures of mammalian purple acid phosphatase (purple), protein phosphatase 1 (blue) [38] and calcineurin (orange) [39]. The metal atoms are drawn as CPK spheres. (d) Stereoview diagram showing the superimposition of the complete Cα trace of mammalian purple acid phosphatase (purple) and protein phosphatase 1 (blue).

quite different. Even for small substrates, differences in substrate specificity have been observed between RKPAP and PPAP. For example, RKPAP exhibits a marked preference for ATP over *p*-nitrophenyl phosphate, whereas the reverse is the case for PPAP [32,33].

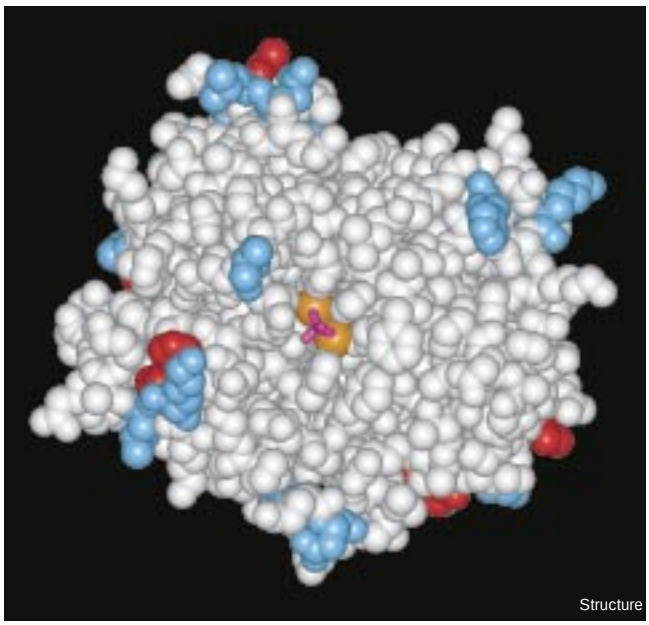
Both enzymes exhibit a strong preference for activated versus unactivated phosphate esters, which distinguishes them from the serine/threonine protein phosphatases. In particular, we have shown that the pig and human PAPs catalyse the hydrolysis of phosphotyrosine-containing peptides very efficiently [3,34]. The difference might be attributable to the ability of the serine/threonine phosphatase to protonate the leaving group of unactivated substrates more effectively than does PAP.

Comparison with other protein structures

In the CATH classification system, proteins are grouped according to class, architecture, topology and homologous superfamily [35] (http://www.biochem.ucl.ac.uk/bsm/cath). PPAP is in the α/β class of proteins that have a four-layer αββα sandwich architecture. Within this architecture, six different CATH topologies have been described. The C-terminal domain of RKPAP corresponds to one of these. As described above, the pig enzyme has the same topology as the RKPAP catalytic domain. Furthermore, their sequence homology also groups them within the same homologous superfamily. The other five CATH topologies of the four-layer αββα sandwich architecture are note worthy because they all bind two or more metal ions: the serine/threonine phosphatases, such as protein phosphatase 1 and protein phosphatase 2B (calcineurin); aldehyde ferredoxin oxidoreductase; deoxyribonuclease I; metallo-β-lactamase; and glutamine phosphoribosylpyrophosphate amidotransferase.

The similarity between RKPAP and the serine/threonine protein phosphatases of the PPP gene family has been noted previously. Indeed, in the SCOP protein classification scheme [36] (http://scop.mrc-lmb.cam.ac.uk/scop/), these two protein classes are grouped within the same superfamily and within the same fold. This definition for the 'metallo-dependent phosphatase' fold would also encompass the structure of protein phosphatase 2C, a PPM-type metal-dependent serine/threonine protein phosphatase. However, the PPM topology is clearly different from the PPP topology and is suggestive of convergent evolution [37], whereas the topologies of PAPs and the PPP phosphatases are very similar. The Cα atoms of 102 residues of PPAP can be superimposed onto the structure of protein phosphatase 1 (PDB accession code 1fjm [38]) with an rmsd of 1.9 Å (Figure 5). There is, however, little structural similarity outside this region (Figure 5). The Cα atoms of 111 residues of PPAP can also be superimposed onto the structure of calcineurin (PDB accession code 1aui [39]) with an rmsd of 2.1 Å. Furthermore, the PPP fold has just two cross-overs between the two halves of the phosphatase domain, and these are at exactly the same points in the topology as the two cross-overs of PAP. By comparison, the PPM fold has four cross-over points between the two halves. Finally, as noted earlier, there is a conserved sequence motif spanning the metal ligands in the PPP and PAP families, which is not present in the PPM family. It has been suggested that this conserved sequence motif indicates divergent evolution from an ancestral metallophosphoesterase [14]. Mammalian PAP, with its simple single-domain structure and striking internal pseudo-symmetry, may more closely resemble an ancestral metallophosphoesterase than the multidomain PPP phosphatases and the multidomain dimeric RKPAP. Furthermore, the internal symmetry of the mammalian PAP suggests that an ancestral metallophosphoesterase could, in turn, have evolved from a gene amplification event from a metal-binding (or phosphate-binding) protein, possibly having a two-layer αβ sandwich architecture. It will be interesting to see if any proteins that have this simpler architecture are known to bind metal and/or phosphate.

Figure 6



CPK representation of the surface of pig purple acid phosphatase. White atoms are identical to those in the human enzyme. Residues coloured blue represent conservative changes in the human enzyme and those in red are not conserved (see also Figure 3). The Fe atoms are shown as orange spheres and the phosphate is drawn in magenta.

Comparison of pig and human PAPs

As noted earlier, one impetus to solving the structure of mammalian PAP is the potential for structure-based design of active-site antagonists of the human enzyme in osteoclasts. The sequences of human and pig enzymes are

~88% identical with 36 amino acid differences, only seven of which are not conservative. A preliminary homology model of the human enzyme, constructed from PPAP, shows that 29 of the 36 different residues are exposed to the surface and seven are buried. Figure 6 shows a CPK model of the PPAP structure with differences between the enzymes highlighted. It is evident that few of the changes are near the active site. In summary, the pig enzyme structure is a good model for human PAP.

Biological implications

Mammalian purple acid phosphatases, or tartrate-resistant acid phosphatases, are expressed by osteoclasts, the cells that resorb bone, and by activated macrophages. There is strong evidence that the osteoclast enzyme is required for normal bone resorption, and elevated serum levels are associated with osteoporosis and other metabolic bone diseases and malignancies. Mammalian purple acid phosphatase has therefore been identified as a possible target for antiosteoporotic drugs. The high-resolution crystal structure of the pig enzyme provides a basis for the structure-based design of such compounds.

Mammalian purple acid phosphatase catalyses two diverse reactions: the hydrolysis of phosphate esters (including phosphoproteins) and the production of hydroxyl radicals. The significance of the phosphatase activity and hydroxyl radical production in the function of osteoclasts and macrophages is not known. The mammalian enzyme has a similar protein fold and binuclear metal-binding site to that of the plant enzyme (and to the catalytic domain of PPP-type serine/threonine phosphatases), suggesting similar catalytic mechanisms. However, the mammalian and plant sequences are only ~15% identical and the plant enzyme has an N-terminal domain that is absent in the mammalian enzyme. In addition, the mammalian enzyme is monomeric and the plant enzyme is dimeric. Of most importance, the protein architectures surrounding the active sites of the two enzymes are significantly different. These differences suggest that the substrate(s) and function(s) of the mammalian and plant enzymes are likely to be quite distinct.

The structure of pig purple acid phosphatase exhibits striking internal symmetry that was not apparent in the more complex plant enzyme and PPP-type phosphatase structures. The symmetry suggests that the binuclear metal centre evolved through the combination of single metal-binding centres. We postulate that this might represent a general mechanism for the evolution of binuclear metal-containing proteins.

Materials and methods

Crystallisation

Pig allantoic fluid purple acid phosphatase (PPAP) was purified as described previously [6] and the oxidised enzyme–phosphate complex prepared as described [22]. The oxidised enzyme–phosphate complex was crystallised by the hanging-drop method. The well solution for crystallisation consisted of 0.1 M sodium citrate pH 5.0, 25% (w/v) PEG 3300, 0.1 M LiCl and 10% (v/v) isopropanol. A detailed description of the crystallisation will be published elsewhere. Native cryocooled crystals diffracted strongly, allowing a data set to 1.55 Å resolution to be collected. Cryocooling did not always yield isomorphous crystals of the native enzyme; therefore the search for heavy-atom derivatives was carried out at room temperature. However, at room temperature, crystals were stable for only 18 h in the X-ray beam, resulting in data sets of 2.0 Å resolution or less.

Data collection and structure determination

Crystals belonged to the space group $P2_12_12_1$ and had one molecule per asymmetric unit and the cryocooled crystal had a unit cell of a = 64.61 Å, b = 70.01 Å, c = 77.11 Å, $\alpha = 90.0°$, $\beta = 90.0°$ and $\gamma = 90.0°$. Room temperature crystals had a slightly larger unit cell of a = 66.85 Å, b = 70.12 Å, c = 78.52 Å, $\alpha = 90.0°$, $\beta = 90.0°$ and $\gamma = 90.0°$. All X-ray data were measured on an R-Axis IIC area detector and processed using DENZO and Scalepack [40]. Heavy-atom positions for four of the derivatives were found independently by direct methods using SHELXS [41]. Difference Fourier techniques were used to identify metal-binding sites for the gold and mercury derivatives and additional heavy-atom sites for the first four derivatives. Heavy-atom positions and occupancies were refined using the PHASES program package [42]. MIR phases were improved by solvent flattening using BC Wang's procedure as implemented in PHASES. The program O [43] was used for model building and inspection of electron-density maps. The highest features in the experimental map were attributed to the Fe–Fe dimetal centre and phosphate. With this information a Cα trace of the RKPAP crystal structure was overlaid onto the PPAP electron density. This established the hand of the molecule and facilitated the tracing of ~85% of the polypeptide chain. At that point, data from the cryocooled crystal (1.55 Å resolution) were used to construct $2F_o–F_c$ and $F_o–F_c$ electron-density maps. Multiple stages of model building and crystallographic refinement, with a bulk-solvent correction, were carried out using X-PLOR Version 3.851 [44], resulting in a final model that included residues 3–304, two Fe atoms, the bridging oxygen, bound phosphate, two *N*-acetylglucosamine residues, 339 water molecules and one isopropanol molecule. The electron density for the sidechains of Lys109, Arg155, Lys177 and Arg253 was weak and they were therefore modelled as alanine in the final structure. Sidechains for Val12, Thr25, His199, Glu282 and Thr298 were modelled in two alternate conformations. The final R_{free} and R factor were 0.257 and 0.213, respectively. A Ramachandran plot was derived using PROCHECK [45] and showed all but one residue, His221, in the most favoured or allowed regions. This residue, which interacts with the binuclear metal centre, is part of a classic γ turn. Two residues, Pro21 and Phe56, have *cis* peptide conformations.

Figures for this paper were generated using MOLSCRIPT version 2.0.1 [46], RASTER 3D version 2 [47,48], SETOR [49], INSIGHT 97 (MSI Corporation) and IRIS SHOWCASE version 3.1 (Silicon Graphics Corporation).

Accession numbers

The refined coordinates and structure factors have been deposited in the Brookhaven Protein Data Bank, are on hold for six months, and have the accession codes 1ute and r1utesf, respectively.

Acknowledgements

We thank L Gahan for useful discussions. This work was supported by a grant-in-aid from Johnson and Johnson Research Pty Ltd, and by the National Health and Medical Research Council. JLM was the recipient of an ARC Queen Elizabeth II Fellowship.

References

1. Klabunde, T. & Krebs, B. (1997). The dimetal centre in purple acid phosphatases. *Structure and Bonding* 89, 177-198.

2. Hayman, A.R., *et al.*, & Cox, T.M. (1996). Mice lacking tartrate-resistant acid phosphatase (Acp5) have disrupted endochondral ossification and mild osteopetrosis. *Development* **122**, 3151-3162.
3. Marshall, K., *et al.*, & Hamilton, S. (1997). Recombinant human and mouse purple acid phosphatases: expression and characterization. *Arch. Biochem. Biophys.* **345**, 230-236.
4. Ek-Rylander, B., Flores, M., Wendel, M., Heinegård, D. & Andersson, G. (1994). Dephosphorylation of osteopontin and bone sialoprotein by osteoclastic tartrate-resistant acid phosphatase. *J. Biol. Chem.* **269**, 14853-14856.
5. Hayman, A.R. & Cox, T.M. (1994). Purple acid phosphatase of the human macrophage and osteoclast – characterization, molecular properties, and crystallization of the recombinant di-iron oxo protein secreted by baculovirus infected insect cells. *J. Biol. Chem.* **269**, 1294-1300.
6. Campbell, H.D., *et al.*, & Zerner, B. (1978). Iron-containing phosphatases: comparisons of enzymes from beef spleen and pig allantoic fluid. *Biochem. Biophys. Res. Commun.* **82**, 615-620.
7. Ek-Rylander, B., Bill, P., Norgård, M., Nilsson, S. & Andersson, G. (1991). Cloning, sequence and developmental expression of a type 5 tartrate-resistant, acid phosphatase of rat bone. *J. Biol. Chem.* **266**, 24684-24689.
8. Ketcham, C.M., Roberts, R.M., Simmen, R.C.M. & Nick, H.S. (1989). Molecular cloning of the type 5, iron-containing, tartrate-resistant acid phosphatase from human placenta. *J. Biol. Chem.* **264**, 557-563.
9. Lord, D.K., *et al.*, & Cox, T.M. (1990). Type 5 acid phosphatase: sequence, expression and chromosomal localization of a differentiation-associated protein of the human macrophage. *Eur. J. Biochem.* **189**, 287-293.
10. Sinn, E., O'Connor, C.J., de Jersey, J. & Zerner, B. (1983). Evidence for a binuclear iron site in pig allantoic fluid acid phosphatase. *Inorg. Chim. Acta* **78**, L13-L15.
11. Antanaitis, B.C., Aisen, P., Lilienthal, H.R., Roberts, R.M. & Bazer, F.W. (1980). The Novel 'g′ = 1.74' spectrum of pink and purple uteroferrin. *J. Biol. Chem.* **255**, 11204-11209.
12. Merkx, M. & Averill, B.A. (1998). The activity of oxidized bovine spleen purple acid phosphatase is due to an Fe(III)Zn(II) 'impurity'. *Biochemistry* **37**, 11223-11231.
13. Sibille, J., Doi, K. & Aisen, P. (1987). Hydroxyl radical formation and iron-binding proteins. *J. Biol. Chem.* **262**, 59-62.
14. Barford, D., Das, A.K. & Egloff, M.-P. (1998). The structure and mechanism of protein phosphatases: insights into catalysis and regulation. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 133-164.
15. Klabunde, T., Sträter, N., Frölich, R., Witzel, H. & Krebs, B. (1996). Mechanism of Fe(III)-Zn(II) purple acid phosphatase based on crystal structures. *J. Mol. Biol.* **259**, 737-748.
16. Wilcox, D.E. (1996). Binuclear metallohydrolases. *Chem. Rev.* **96**, 2435-2458.
17. Waller, B.J. & Lipscomb, J.D. (1996). Dioxygen activation by enzymes containing binuclear non-heme iron clusters. *Chem. Rev.* **96**, 2625-2657.
18. Sträter, N., Klabunde, T., Tucker, P., Witzel, H. & Krebs, B. (1995). Crystal structure of a purple acid phosphatase containing a dinuclear Fe(III)-Zn(II) active site. *Science* **268**, 1489-1492.
19. Beck, J.L., de Jersey, J., Zerner, B., Henrich, M.P. & Debrunner, P. (1988). Properties of the Fe(II)-Fe(III) derivative of red kidney bean purple acid phosphatase. Evidence for a binuclear iron centre in the native enzyme. *J. Am. Chem. Soc.* **110**, 3317-3318.
20. Klabunde, T., Sträter, N., Krebs, B. & Witzel, H. (1995). Structural relationship between the mammalian Fe(III)-Fe(II) and the Fe(III)-Zn(II) plant purple acid phosphatases. *FEBS Lett.* **367**, 56-60.
21. Roberts, R.M., Raub, T.J. & Bazer, F.W. (1986). The role of uteroferrin in transplacental iron transport in the pig. *Fed. Proc.* **45**, 2513-2518.
22. Keough, D.T., Beck, J.L., de Jersey, J. & Zerner, B. (1982). Iron-containing acid phosphatases: interaction of phosphate with the enzyme from pig allantoic fluid. *Biochem. Biophys. Res. Commun.* **108**, 1643-1648.
23. Cassady, A.I., King, A., Cross, N.C.P. & Hume, D.A. (1993). Isolation and characterization of the genes encoding mouse and human type-5 acid phosphatase. *Gene* **130**, 201-207.
24. Sträter, N., Lipscomb, W.N., Klabunde, T. & Krebs, B. (1996). Two-metal ion catalysis in enzymatic acyl- and phosphoryl-transfer reactions. *Angew. Chem. Int. Ed. Engl.* **35**, 2024-2055.
25. True, A.E., Scarrow, R.C., Randall, C.R., Holz, R.C. & Que, L. (1993). EXAFS studies of uteroferrin and its anion complexes. *J. Am. Chem. Soc.* **115**, 4246-4255.
26. Wang, X. & Que, L. (1998). Extended X-ray fine structure studies of the anion complexes of FeZn uteroferrin. *Biochemistry* **37**, 7813-7821.
27. Rusnak, F., Yu, L. & Mertz, P. (1996). Metalloenzymes and signal transduction: the protein serine/threonine phosphatases, a novel class of binuclear metal-containing enzymes. *J. Bioinorg. Chem.* **1**, 388-396.
28. Turowski, P.N., Armstrong, W.H., Liu, S., Brown, S.N. & Lippard, S.J. (1994). Synthesis and characterization of hydroxo-bridged diiron (III) complexes containing carboxylate or phosphate ester bridges: comparisons to diiron(III) proteins. *Inorg. Chem.* **33**, 636-645.
29. Wahnon, D., Lebuis, A.-M. & Chin, J. (1995). Hydrolysis of a phosphate diester doubly coordinated to a dinuclear cobalt(II) complex: a novel mechanism. *Angew. Chem. Int. Ed. Engl.* **34**, 2412-2414.
30. Aquino, M., Lim, J.S. & Sykes, A.G. (1994). Mechanism of the reaction of different phosphates with the iron(II) iron(III) form of purple acid phosphatase from porcine uteri (uteroferrin). *J. Chem. Soc. Dalton Trans.* **21**, 429-436.
31. Orlando, J.L., Zirino, T., Quirk, B.J. & Averill, B.A. (1993). Purification and properties of the native form of the purple acid phosphatase from bovine spleen. *Biochemistry* **32**, 8120-8129.
32. Beck, J.L., *et al.*, & Zerner, B. (1986). Properties of a purple acid phosphatase from red kidney bean: a zinc-iron metalloenzyme. *Biochim. Biophys. Acta* **869**, 61-68.
33. Schlosnagle, D.C., Bazer, F.W., Tsibris, J.C.M. & Roberts, R.M. (1974). An iron-containing phosphatase induced by progesterone in the uterine fluids of pigs. *J. Biol. Chem.* **249**, 7574-7579.
34. Nash, K., Feldmuller, M., de Jersey, J., Alewood, P. & Hamilton, S. (1993). Continuous and discontinuous assays for phosphotyrosyl protein phosphatases using phosphotyrosyl peptide substrates. *Anal. Biochem.* **213**, 303-309.
35. Orengo, C. (1994). Classification of protein folds. *Curr. Opin. Struct. Biol.* **4**, 429-440.
36. Murzin, A.G., Brenner, S.E., Hubbard, T. & Chothia, C. (1995). 'SCOP'; a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* **247**, 536-540.
37. Barford, D. (1996). Molecular mechanisms of the protein serine/threonine phosphatases. *Trends Biochem. Sci.* **21**, 407-412.
38. Goldberg, J., *et al.*, & Kuriyan, J. (1995). Three dimensional structure of the catalytic subunit of protein serine/threonine phosphatase-1. *Nature* **376**, 745-753.
39. Kissinger, C.R., *et al.*, & Villafranca, E. (1995). Crystal structure of human calcineurin and the human FKBP12–FK506-calcineurin complex. *Nature* **378**, 641-644.
40. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Meth. Enzymol.* **276**, 307-326.
41. Sheldrick, G.M. (1993). SHELXL93. A program for crystal structure refinement. University of Göttingen, Germany.
42. Furey, W. & Swaminathan, S. (1997). 'PHASES' - a program package for processing and analyzing diffraction data from macromolecules. *Methods Enzymol.* **277**, 590-620.
43. Jones, T.A., Zou, J.Y., Cowan, S.W. & 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.
44. Brünger, A.T. (1992) *X-PLOR*: A system for X-ray crystallography and NMR. Yale University, New Haven, CT.
45. Laskowski, R.A., Macarthur, M.W., Moss, D.S. & Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283-291.
46. Kraulis, P.J. (1991). MOLSCRIPT - a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946-950.
47. Merritt, E.A. & Murphy, M.E.P. (1994). Raster3D Version 2.0. A program for photorealistic molecular graphics. *Acta. Crystallogr. D* **50**, 869-873.
48. Bacon, D.J. & Anderson, W.F. (1988). A fast algorithm for rendering space filling molecular pictures. *J. Mol. Graph.* **6**, 219-222.
49. Evans, S.V. (1993). SETOR: Hardware lighted three dimensional solid model representations of macromolecules. *J. Mol. Graph.* **11**, 134-138.
50. Brünger, A.T. (1992). The free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472-474.
51. Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577-2637.